



Fundamental Limitations within the Selected Cryptographic Scenarios and Supra-Quantum Theories

By

M.SC.ENG. MAREK WINCZEWSKI

Supervisor: DR HAB. KAROL HORODECKI, PROF. UG

Institute of Theoretical Physics and Astrophysics,
Faculty of Mathematics, Physics and Informatics
UNIVERSITY OF GDAŃSK

Acknowledgement

This work would not have been possible without the mentorial and personal support of my supervisor, Karol Horodecki. In particular, I acknowledge and am grateful for support from the National Science Centre, Poland (NCN) grant Sonata Bis 5 (grant number: 2015/18/E/ST2/00327), and Foundation for Polish Science (IRAP project, ICTQT, contract No. 2018/MAB/5, co-financed by EU within Smart Growth Operational Programme).

I would also like to extend the deepest gratitude to Borhan Ahmadi, Stefan Bäuml, Jakub Jan Borkała, Marek Czachor, Siddhartha Das, Michał Horodecki, Paweł Horodecki, Michał Kamoń, Ryszard P. Kosteki, Stanisław Kryszewski, Marcin Łobejko, Antonio Mandarino, Adam Rutkowski, Omer Sakarya, Roberto Salazar, John H. Selby, and Bianka Wołoncewicz for many years of fruitful collaboration, support, and countless discussions. Last but certainly not least, I would like to especially thank Tamoghna Das for the long-lasting, close, and fruitful collaboration.

Finally, I would like to say thanks to all my friends and family for their support. In particular, I wish to thank members of ice hockey teams Marvin's Fighters and Ślizg Gdynia for the time spent on the ice. Special thanks go to my friends Michał Mischyszyn and Martyna Wygonna-Mischyszyn for their unconditional love and support.

Dedication

To my loving parents,

Dorota Winczewska & Piotr Winczewski

Abstract

Quantum theory provides the ground for the most significant technological advances of the twenty-first century. The quantum revolution that is about to happen will bring us not only quantum computers but also quantum communication between quantum hardware. The global network of quantumly communicating devices will constitute the so-called quantum Internet. The promise quantum theory brings us about quantum communication is higher than the classical level of security and, in some scenarios, even unconditional security. Namely, there exist protocols for secret key distillation having entangled quantum states as a resource such that the security of the produced secret key is guaranteed by the laws of physics and are independent of the (possibly malicious) inner-working of the involved cryptographic devices. The more bits of the secure key can be distilled, given some entangled resources, the better the performance of the protocol. In this way, particular protocols constitute the lower bounds on the key rate achievable in the scenario. On the other hand, determining the optimal protocol or justification for pursuing it requires knowledge of the upper bounds on the achievable key rate. Finding the upper bounds on the secure key rate is, therefore, of crucial importance regarding the construction of quantum communication networks of the future.

In the collection of articles constituting this dissertation, of which I am a coauthor, as our primary interest, we study the fundamental limitations within the selected quantum and supra-quantum cryptographic scenarios in the form of upper bounds on the achievable key rates. We investigate various security paradigms, bipartite and multipartite settings, as well as single-shot and asymptotic regimes. Specifically, our findings contribute to the secret key agreement scenario (SKA), device-dependent conference key agreement scenario (DD-CKA) both for quantum channels and quantum states in one-shot and asymptotic regimes, device-independent conference key-agreement scenario (DI-CKA) in both asymptotic and one-shot regime, and non-signaling device-independent secret key agreement scenario (NSDI). Our studies, however, extend beyond the derivations of the upper bounds on the secret key rates in the mentioned scenarios. In particular, we propose a novel type of rerouting attack on the quantum Internet for which we find a countermeasure and benchmark its efficiency. Furthermore, we propose several upper bounds on the performance of quantum (key) repeaters settings. We derive a lower bound on the secret key agreement capacity of a quantum network, which we tighten in an important case of a bidirectional quantum network. The squashed nonlocality derived here as an upper bound on the secret key rate is a novel non-faithful measure of nonlocality. Furthermore, the notion of the non-signaling complete extension arising from the complete extension postulate as a counterpart of purification of a quantum state allows us to study analogies between non-signaling and quantum key distribution (QKD) scenarios.

From a larger perspective, our results are not only technical ones describing selected cryptographic tasks. In some cases, our findings confront fundamental questions regarding the foundations of quantum theory. The frameworks we develop allow to investigate this underlying subject insightfully both from the inside and outside of the quantum theory.

Abstrakt

Teoria kwantowa stanowi podwaliny dla najbardziej znaczących osiągnięć technologicznych dwudziestego pierwszego wieku. Rewolucja kwantowa, która ma się wkrótce wydarzyć, przyniesie nam nie tylko komputery kwantowe, ale także kwantową komunikację między urządzeniami nowego typu. Globalna sieć urządzeń porozumiewających się między sobą z wykorzystaniem kwantowej komunikacji utworzy tzw. Internet kwantowy. Obietnica, jaką daje nam teoria kwantowa, w zakresie kwantowej komunikacji, gwarantuje poziom który jest wyższy niż klasyczny poziom bezpieczeństwa, a w niektórych scenariuszach kryptograficznych nawet bezwarunkowe bezpieczeństwo, niezależne od implementacji urządzeń generujących klucz kryptograficzny. Mianowicie, istnieją protokoły destylacji klucza kryptograficznego używające jako zasobu splątanych stanów kwantowych, takie że bezpieczeństwo wytworzonego klucza kryptograficznego gwarantowane jest prawami fizyki. Im więcej bitów bezpiecznego klucza można wydestylować, z określonego stanu kwantowego, tym wyższa wydajność protokołu. W ten sposób poszczególne protokoły wyznaczają granice dolne na ilość klucza możliwego do wydestylowania w danym scenariuszu kryptograficznym. Z drugiej strony określenie optymalnego protokołu, lub uzasadnienie potrzeby dalszego go szukania, wymaga znajomości ograniczeń górnych na możliwą do uzyskania ilość klucza. Znalezienie ograniczeń górnych na ilość bezpiecznego klucza kryptograficznego ma zatem istotne znaczenie dla budowy kwantowych sieci komunikacyjnych przyszłości.

W zbiorze artykułów składających się na niniejszą rozprawę doktorską, których jestem współautorem, naszym głównym celem jest określenie fundamentalnych ograniczeń górnych na osiągalną ilość bezpiecznego klucza kryptograficznego w wybranych kwantowych i supra-quantowych scenariuszach kryptograficznych. Zostały zbadane różne paradygmaty bezpieczeństwa, sytuacje dwu i wieloosobowe, a także reżimy jednorazowe i asymptotyczne. W szczególności nasze rezultaty dotyczą scenariusza uzgadniania klucza sekretnego (ang. secret key agreement scenario, SKA), scenariusza uzgadniania klucza konferencyjnego zależnego od urządzenia (ang. device-dependent conference key agreement, DD-CKA) zarówno dla stanów kwantowych jak i kanałów kwantowych, w reżimach jednorazowych i asymptotycznych, scenariusza uzgadniania klucza konferencyjnego niezależnego od urządzenia (ang. device-independent conference key agreement, DI-CKA) zarówno w reżimie jednorazowym jak i asymptotycznym, oraz scenariusza uzgadniania klucza sekretnego w obecności adwersarza ograniczonego jedynie więzami braku sygnalizacji (ang. device-independent non-signaling secret key agreement, NSDI). Nasze badania wykraczają jednak poza wyprowadzenie ograniczeń górnych na ilość klucza kryptograficznego osiągalnego we wspomnianych scenariuszach. W szczególności proponujemy nowy typ ataku przekierowującego (ang. rerouting attack) na kwantowy Internet. Dla tego ataku znajdujemy środek zaradczy oraz kwantyfikujemy skuteczność naszego rozwiązania. Proponujemy kilka ograniczeń górnych na wydajności układów kwantowych powtarzaczy oraz powtarzaczy klucza kwantowego. Dodatkowo wyprowadzamy ograniczenie dolne na pojemność uzgadni-

ania sekretnego klucza (ang. secret key agreement capacity) dla sieci kwantowej, które to ograniczenie to ulepszymy w ważnym przypadku sieci złożonej z dwukierunkowych kanałów kwantowych (ang. bidirectional quantum channel). Ściśnięta nielokalność (ang. the squashed nonlocality) wyprowadzona jako ograniczenie górne na ilość klucza kryptograficznego jest w istocie nową miarą nielokalności. Co więcej, pojęcie niesygnalizującego kompletnego rozszerzenia (ang. non-signaling complete extension) wynikające z postulatu kompletnego rozszerzenia jako odpowiednik kwantowej puryfikacji pozwala nam na badanie analogii między niesygnalizującymi i kwantowymi scenariuszami uzgadniania klucza kryptograficznego.

Patrząc z szerszej perspektywy, nasze wyniki nie są jedynie rezultatami technicznymi, opisującymi wybrane zadania kryptograficzne. W niektórych przypadkach nasze odkrycia dotyczą fundamentalnych kwestii dotyczących pryncypiów teorii kwantowej. Ramy, które opracowujemy, pozwalają wnikliwie badać ten podstawowy temat zarówno od wewnątrz, jak i od zewnątrz teorii kwantowej.

Publications Included in the Dissertation

- [A] Omer Sakarya, Marek Winczewski, Adam Rutkowski, Karol Horodecki. *Hybrid quantum network design against unauthorized secret-key generation, and its memory cost*. Phys. Rev. Research 2, 043022 (2020).
- [B] Siddhartha Das, Stefan Bäuml, Marek Winczewski, Karol Horodecki. *Universal Limitations on Quantum Key Distribution over a Network*. Phys. Rev. X 11, 041016 (2021).
- [C] Karol Horodecki, Marek Winczewski, Siddhartha Das. *Fundamental limitations on device-independent quantum conference key agreement*. Phys. Rev. A 105, 022604 (2022).
With erratum in Phys. Rev. A 107, 029902 (2023).
- [D] Marek Winczewski, Tamoghna Das, Karol Horodecki. *Limitations on device independent key secure against nonsignaling adversary via the squashed non-locality*. Phys. Rev. A 106, 052612 (2022).
- [E] Marek Winczewski, Tamoghna Das, John H. Selby, Karol Horodecki, Paweł Horodecki, Łukasz Pankowski, Marco Piani, Ravishankar Ramanathan. *Complete extension: the non-signaling analog of quantum purification*. Preprint arXiv:1810.02222. Published online 2018. Accepted for publication in Quantum 27.01.2023.

Other articles

- [E] Marek Winczewski, Antonio Mandarino, Michał Horodecki, Robert Alicki. *Bypassing the Intermediate Times Dilemma for Open Quantum System*. Preprint arXiv:2106.05776. Published online 2021.
- [F] Marek Winczewski, Robert Alicki. *Renormalization in the Theory of Open Quantum Systems via the Self-Consistency Condition*. Preprint arXiv:2112.11962. Published online 2021.
- [G] Marcin Łobejko, Marek Winczewski, Gerardo Suárez, Robert Alicki, Michał Horodecki. *Towards reconciliation of completely positive open system dynamics with the equilibration postulate*. Preprint arXiv:2204.00643. Published online 2022.

Table of Contents

	Page
1 Introduction	1
2 Summary of Dissertation	11
2.1 Preliminaries	11
2.1.1 Private States and Positive Partial Transpose States	11
2.1.2 Secret Key Agreement Tasks	14
2.1.3 Quantum Channels	22
2.1.4 Entanglement and Genuine Entanglement	23
2.1.5 Nonlocality and Genuine Nonlocality	24
2.1.6 Generalized Probabilistic Theories	25
2.2 Hybrid Quantum Network Design Against Unauthorized Secret-Key Generation, and Its Memory Cost [A]	31
2.3 Universal Limitations on Quantum Key Distribution over a Network [B]	39
2.4 Fundamental Limitations on the Device-Independent Quantum Conference Key Agreement [C]	53
2.5 Limitations on Device-Independent Key Secure Against Nonsignaling Adversary via the Squashed Nonlocality [D]	61
2.6 Complete Extension: the Non-Signalling Analog of Quantum Purification [E]	69
3 Outlook	81
Bibliography	85

Introduction

The desire to transfer messages in a confidential manner is probably as old as the history of humankind. It is known that ciphers were known already in ancient times. The motivations to hide information from untrusted parties vary but were always the same, ranging from personal to military purposes. In particular, the Caesar cipher owes its name to the Roman general, statesman, and dictator Gaius Julius Caesar who used it in his military campaigns [Doo18, Bau21]. The development of methods on how to pass information secretly led to the establishment of a new area of science called cryptography. The distinctive feature of cryptography as a field of science and engineering is the fact that one has to confront an intelligent enemy called the eavesdropper and not only the laws of nature. Up to the beginning of the twentieth century, cryptography was in the hands of linguists [Bau21], e.g., Auguste Kerckhoffs. Next, the engineers got involved, with Gilbert Vernam and his one-time pad encryption as a prominent example [Bau21]. In fact, one-time pad encryption was originally invented earlier by Frank Miller in 1882 [Mil82]. The later period belongs to mathematicians, who, amongst other successes, broke a German cipher device Enigma in the 1930s as the most famous achievement [Bau21]. In this way, the enormous efforts of Marian Rejewski, Jerzy Różycki, and Henryk Zygalski, continued by Alan Turing, contributed to the victory of the Allies in the Second World War. Later, in the 1940s, Claude Shannon [Sha01] heavily developed the theory of information established by Harry Nyquist and Ralph Hartley in the 1920s [Nyq24, Har28], which is crucial for modern cryptographic purposes. Besides formalizing the concepts of processing and transmitting information and introducing the concept of entropy of information, Claude Shannon proved that the one-time pad encryption scheme is information-theoretically secure [Sha01]. The development of computers and communication between them entailed involvement in the cryptography of computer scientists in the middle 1970s [Bau21]. The efforts of computer

scientists allowed for the construction of a global network of communicating electronic devices called the Internet and gave it its present shape. Finally, in the 1980s, physicists realized that the quantum theory, studied from the beginning of the twentieth century, has the potential to be applied in cryptography. Namely, in 1983 Stephen Wiesner proposed the concept of quantum money resistant to forgery [Wie83]. In fact, the idea of Stephen Wiesner goes back to the 1970s, but it remained unpublished until 1983. Next, in 1984, Charles Bennett and Gilles Brassard developed the first quantum communication protocol named BB84 after the authors [BB84], i.e., based on the formalism of quantum states and indeterminacy in their measurements. The BB84 protocol, under some assumptions, is provably secure as it produces a secure key ready for one-time pad encryption. Next, in 1991 Artur Ekert designed the E91 protocol [Eke91], i.e., the first quantum cryptography protocol based on the entangled quantum states, that allows for the detection of the eavesdropper's presence. From these times, the combined power and knowledge of scientists from all of the aforementioned fields were used to develop a plethora of protocols, schemes, scenarios, and security paradigms for quantum, supra-quantum, and post-quantum cryptography. The ultimate goal of the efforts put into the development of the field is to establish communication schemes, the security of which is guaranteed by the laws of physics. In this way, quantum cryptography and recent progress in building quantum computers give a promise for constructing a quantumly secure Internet in the not-too-distant future [DBCZ99, MLK⁺16, ZPD⁺18, WEH18].

In the past several decades, four major security paradigms were developed for scenarios in which the cryptographic task of the honest parties is to distill the secret key in the presence of a malicious eavesdropper. Here, we list them with respect to the increasing power of the eavesdropper, which also reflects the level of professional paranoia in the number of incorporated assumptions. Furthermore, in this description, for the sake of its simplicity and conciseness of the notation, we restrict ourselves to the situation in which there are only two honest parties, usually called Alice and Bob, and the eavesdropping malicious Eve. The first one is the secret key agreement scenario (SKA) [CK78, Mau93, GRTZ02, BB84], in which the honest parties share n -copies of $P(AB)$ which is a marginal of a classical tripartite probability distribution $P(ABE)$, where A , B , and E are the random variables describing the outputs of Alice, Bob and Eve, respectively. In order to distill the secret key, the honest parties process their data with so-called local operations and public communication (LOPC). At the same time, Eve listens to public communication and is in power to apply any stochastic map to her data. Furthermore, there are two distinct quantum key distribution (QKD) paradigms. The second paradigm to be described here is the quantum device-dependent (QDD) [BB84, Eke91, Ben92, ABB⁺06] scenario introduced at the beginning of quantum cryptography. In this scenario, the honest parties share a marginal state ρ_{AB} of (in the worst case from their perspective) a tripartite pure quantum state $|\psi_{ABE}\rangle$, i.e., $|\psi_{ABE}\rangle$ is the purification of ρ_{AB} to the eavesdropper's system E . Aiming to distill the secret key, the honest parties, Alice and Bob, process their subsystems

using local quantum operations and classical communication (LOCC). Here, we remark that another definition based on LOPC is equivalent in the mentioned worst-case. On the other hand, the malicious Eve is assumed to receive any quantum system discarded by Alice and Bob, listen to the communication between the honest parties, and be able to perform a quantum operation on her subsystem. The problem with the QDD scenario is that Alice and Bob have to trust the inner workings of the device they operate on, used to transfer quantum states and to produce a classical output. Precisely speaking, the honest parties have to trust the dimensionality of the quantum state and measurements performed on it. A sophisticated solution to the drawback mentioned above is provided in the quantum device-independent (QDI) [Eke91, BCP⁺14, MY04, AGM06, ABG⁺07, MPA11, AFRV19] scenario. In the QDI scenario, Alice and Bob share an untrusted device, described with a conditional probability distribution $P(AB|XY)$ originating from local measurements $M_{A|X}$, $M_{B|Y}$ on a quantum state ρ_{AB} , i.e., $P(AB|XY) \equiv \text{Tr} [M_{A|X} \otimes M_{B|Y} \rho_{AB}]$. Here, X and Y correspond to the choices of measurement settings (inputs) of Alice and Bob, respectively, and A and B to outcomes as before. The eavesdropping Eve is assumed to be restricted by the laws of quantum mechanics, and in particular, she may hold the purification $|\psi_{ABE}\rangle$ of the bipartite state ρ_{AB} . Therefore, the security in the QDI paradigm is based solely on the input-output statistic of the honest parties. In the last paradigm, called the non-signaling device-independent secret key agreement (NSDI) [BHK05, BCK12, SGB⁺06, AMP06, AGM06, Mas09, HRW10, MRC⁺14] scenario, the assumptions on the power of the eavesdropper are even more relaxed than in the QDI scenario. Namely, the eavesdropper is restricted solely by the non-signaling condition that prevents Eve from changing the statistics of the honest parties and excludes instantaneous communication. Furthermore, the NSDI scenario allows the honest parties to share possibly supra-quantum (stronger than quantum) correlations constrained only by the no-signaling condition [PR94]. Incidentally, the non-signaling conditional probability distributions constitute a generalized probabilistic theory (GPT) of non-signaling behaviors, wherein they have the role of states, see Ref. [Bar07] and references therein, and also Refs. [Plá21, Mül21, Lam18] for recent progress in this direction. Here, the device shared between parties is described as a tripartite conditional probability distribution $P(ABE|XYZ)$, with E and Z being the output and input of Eve, respectively, and A , B , X , Y are as before. The honest parties choose their measurement settings (inputs) and process their output data with LOPC operations. The same as before, the device is assumed to be provided by Eve, who can listen to public communication and perform a certain class of operations on her subsystem. Moreover, the tripartite device $P(ABE|XYZ)$ is assumed to be the worst-case extension of the marginal $P(AB|XY) \equiv \sum_e P(ABe|XYz)$ shared by the honest parties. Finally, all of the mentioned scenarios have their extensions to the case of multiple honest parties. The secret key agreement between multiple honest parties is called the conference key agreement [CL05, AH09b]. Similarly, QDD gives rise to the device-dependent conference key agreement (DD-CKA) [HHHO09, AH09b], QDI is lifted to

the device-independent conference key agreement (DI-CKA) [RMW18], and NSDI is naturally extended to the non-signaling conference key agreement (NS-CKA) [PKBW23]. Additionally, a modern research subject is to study secret (conference) key agreement capacities of quantum channels in scenarios that generalize the concept of the use of quantum states.

The most important property that quantifies the performance of the secret key distillation protocol is the number of bits of the secret key that is produced given some resource state (or device). The number of bits of the secret key produced in the protocol divided by the number of copies of the state (or device) used is called the protocol rate. In this way, the rates of particular protocols constitute the lower bounds on the maximal secret key rate achievable in the given scenario. The protocols of secret key distillation, and lower bounds on their rates, in the bipartite SKA [CK78, Mau93], QDD [DW04, KGR05, HHHO05, HHHO09], QDI [VV14, AFRV19], NSDI [BHK05, AGM06, AMP06, SGB⁺06, Mas09, HRW10, BCK12, MRC⁺14], and multipartite CKA [YHH⁺09], DD-CKA [AH09b], and DI-CKA [RMW18] scenarios have been widely studied. The definition of the secret key rate involves a supremum over achievable rates by any theoretically possible protocols. Therefore, in a generic case, the maximal achievable secret key rate is unknown, as we usually do not know if the best known protocol is optimal. The complementary method to study, in a protocol-independent way, secret key rates achievable in cryptographic scenarios is to determine the upper bounds on them. Furthermore, the upper bounds can be determined both in the one-shot (single-run) and asymptotic regime. The one-shot regime refers to the key distillation from a single copy of a quantum state (or device), and the asymptotic regime considers a situation in which the honest parties have access to an unlimited number of copies of the resource state (or device). The upper bounds help to determine whether the known protocol is optimal, assess how much the protocols can be improved, or decide if the pursuit for a better protocol is still justified. Moreover, the upper bounds are also important for practical reasons. If an actually implemented cryptographic device produces more key than is determined by the upper bound, then the key provided by the device can not be secure. These reasons make finding the upper bounds an important direction of research. Therefore, not only the lower bounds but also the upper bounds on the secret key rate in the bipartite SKA [MW99, MW97, RW03, RSW03], QDI and QDD [Chr02, CEH⁺07, YHH⁺09, Wil16, TGW14b, TGW14a, CFH21, AFL21, FBJL⁺21], NSDI [KWW20] and multipartite CKA [CMS02], DD-CKA [CMG22, AH09b, CMS02] scenarios have been studied. The same concerns multipartite DD-CKA scenario considering quantum channels [TGW14b, TGW14a, PLOB17]. However, apart from Ref. [AMP06], only recently the upper bounds in the DI-CKA [HWD22, PKBW23], NSDI [KWW20, WDH22, CFH21, AFL21], and NS-CKA [PKBW23] scenarios were studied.

The vast development of cryptography based on the principles of non-classical physical theories rather than solely on mathematics in the form of number theory, combinatorics, and linear algebra, is due to the physical phenomena invisible in the classical description of the

world stemming from intuitive perception. The advantage for the honest parties in secret key distillation tasks that comes with non-classical resources in the form of entanglement of quantum states or nonlocal correlations in devices comes with the flip side of the coin. Namely, the attack strategy of the eavesdropper possibly includes sharing of non-classical resources with the honest parties. Colloquially speaking, the classical eavesdropper is the weakest one (SKA), the quantum eavesdropper (QKD) is stronger than the classical, and the eavesdropper limited solely by the no-signaling principle (NS) is the most powerful one. Incidentally, the cryptography considering non-signaling adversary (NSDI) is still valid, even when some future theory will replace quantum theory as the most accurate description of reality, as long as it will be a non-signaling one, i.e., consistent with the special theory of relativity. The above hierarchy demonstrates, therefore, that the physical framework we adopt influences the notion of security and consequently determines the limitations of the physical theory concerning attainable quantities of secrecy. The converse statement is also true. The studies on the amounts of secret correlations achievable in different cryptographic paradigms can bring us new insights into quantum theory or non-signaling theories. In a similar manner, Peter Rastal [Ras85], Sandu Popescu, and Daniel Rohrlich [PR94] have shown that quantum theory is not maximally nonlocal, by inverting the logical order and making nonlocality an axiom. They demonstrated that there might exist theories that preserve relativistic causality but exhibit nonlocal correlations that violate some Bell inequalities stronger than quantum theory, for which the limit was proved by Boris Tsirelson [Cir80]. Having the above past example of indirect investigation in mind, we believe that a research on relations between entanglement, nonlocality, secrecy, and their quantification can improve the understating of physical reality or at least develop frameworks to study it.

The articles included in this dissertation [SWRH20, DBWH21, HWD22, WDH22, WDS⁺18], contribute to almost all, except until recently untouched NS-CKA scenario, mentioned cryptographic paradigms. Our focus is directed mainly but not solely on the upper bounds on the maximal secret (conference) key rates (and capacities) achievable within the selected cryptographic scenarios, in some cases both single-shot and asymptotic regimes. The other topics of our concern are, among others, hacking attacks on the Quantum Internet [Mak09, SNS⁺21] and possible countermeasures to them, the upper bounds on the performance of quantum (key) repeater schemes [BDCZ98, DBCZ99, MATN15, CZC⁺21], analogies between different cryptographic paradigms, the GPTs with a particular interest in the connection between the theory of non-signaling behaviors and NSDI cryptography, or the lower bounds on conference key agreement of selected quantum network schemes. We believe that the results presented in this dissertation display a considerably vast landscape of the fundamental limitations concerning the mentioned cryptographic scenarios, with some extra contributions. Here, we remark that by the fundamental, we mean limitations that stem solely from the formalism of quantum mechanics or the no-signaling principle. It is interesting, on its own, why Nature imposes

constraints on the transmission and processing of information. The main results of the articles integrated into this dissertation are briefly described in the following paragraphs. We invoke the published articles [SWRH20, DBWH21, HWD22, WDH22] in the chronological order of appearance, leaving the unpublished one, i.e., Ref. [WDS⁺18] preprint, for the end. A detailed description is left for the next Section.

The first article [SWRH20] in this dissertation considers a quantum network under a threat of a new type of rerouting attack of our proposal [SNS⁺21]. The idea of the attack is the use of malicious software by dishonest end-users that performs entanglement swapping protocol in the hub node of a quantum network. This attack leads to an unauthorized consumption of the resources of the network in favor of the malicious end-users that gain shared entangled states. The example of a quantum network under study exhibits star-like topology with a single (central) hub node and multiple end-users. The end-users want to exchange only classical data with the central node, however, in a quantumly secure way. As the main result, we devise a countermeasure to the proposed attack based on special classes of quantum states that carry directly accessible (device-dependent) secret key but possess a low repeatable key rate [BCHW15]. Precisely speaking, the countermeasure proposed by us is based on the fact that quantum security is not always transitive and employs either different types of the so-called private states [HHHO05, HHHO09] or positive partial transpose (PPT) states that approximate the mentioned private states. We benchmark our solution by quantifying the difference between the device-dependent secret key and repeatable key rates, i.e., the so-called gap of the scheme, as well as the amount of quantum memory spent solely for the security of the scheme, i.e., the so-called memory overhead of the scheme. This quantification is done by providing upper bounds on the repeatable key rate for the considered classes of states. At the same time, lower bounds on the gap of the scheme and memory overhead required us to identify both lower and upper bounds on the secret key rate. Moreover, we identify a low-dimensional example of a quantum state for which the scheme employing it exhibits a strict gap between secure and repeatable key rates [HHHO05, DKDD⁺11]. Our findings show that the protection of the scheme against the proposed attack has its cost in the amount of quantum memory required to be used in the scheme. In this way, our findings contribute to the development of the quantumly secure schemes of the Internet of the future.

In the second article [DBWH21], we study the upper bounds on the achievable key rates in the device-dependent conference key agreement (DD-CKA) scenario. The conference key is the secret key shared by more than two parties. We do so by first introducing the concept of a multiplex quantum channel and subsequently determining general upper bounds on the device-dependent secret key agreement capacities of quantum channels in the multipartite scenario. In principle, the uses of a multiplex quantum channel interleaved with local operations and classical communication (LOCC) can simulate any conference key distribution protocol. Entanglement measures based on the notion of the generalized divergence provide case-specific upper bounds

in terms of various relative entropy functions [Ume62, MLDS⁺13, Dat09b, Dat09a, MLDS⁺13, WWY14, BD10, WR12] on the plethora of secret-key distribution tasks both in asymptotic and one-shot (single-use) scenarios. In this way, the framework we developed has a unifying character as it provides a case-specific upper bound for a broad range of different scenarios within a common structure. We exemplify the use of our results in two specific cases, i.e., in measurement-device-independent quantum key distribution (MDI-QKD) [LCQ12, BP12] scenario and the setup of quantum key repeaters [BDCZ98, DBCZ99, MATN15, CZC⁺21]. In the case of the dual-rail scheme [RP10] for the MDI-QKD scenario, our upper bound is tight and hence outperforms the so-called repeaterless bound (RB). For the setups of quantum repeaters, the upper bound provided by us is at least comparable with those in Refs. [BCHW15, CF17]. However, in our case, more parameters of the scheme can be incorporated. Additionally, we provide a non-trivial lower bound on the secret key agreement capacity for a bidirectional quantum network. Furthermore, we provide upper bounds on the achievable secret key rate in the scenario employing quantum states. Here, as a byproduct, we obtain an upper bound on Greenberger-Horne-Zeilinger (GHZ) states distillation in both single-shot and asymptotic regimes. This result is possible since GHZ states are an example of private states that are ideal final states of any key distillation LOCC protocol. In order to obtain upper bounds on the rate of distillation of GHZ states from noisy GHZ and W states, we conduct an extensive numerical investigation. One of the overcome difficulties is finding biseparable states close enough to GHZ and W or their noisy versions required for the numerical calculations. In summary, our work substantially contributes to the development of the field of DD-CKA cryptography, mainly but not only via providing a unified framework for deriving upper bounds on conference key rates and capacities for various DD-CKA scenarios.

The third article [HWD22] is devoted to upper bounds on the device-independent conference key agreement (DI-CKA) rates. We remark here that our paper is the first to consider upper bounds on DI-CKA rates, as before, only lower bounds on DI-CKA rates had been studied (see also recent work in Ref. [PKBW23]). We first define the reduced c-squashed entanglement as the multipartite generalization of cc-squashed entanglement [AFL21, KHD22]. Next, we show that the reduced c-squashed entanglement upper bounds the device-independent conference key rate achievable by the standard protocols, i.e., the protocols that use a single input to generate the key [AFDF⁺18]. To be precise, we show the above with respect to two definitions of DI-CKA rates, i.e., the “dev” and the “par” definitions. We note that the independent and identically distributed (iid) setting is sufficient for our purposes. In the “dev” definition of DI-CKA secret key rate, the adversary is assumed to mimic the full statistics of the honestly implemented device. In contrast, in the “par” definition, the adversary has to mimic only some relevant parameters of the device, such as the level of violation of some Bell inequality or quantum bit error rate (QBER). Our goal is achieved by generalizing the upper bounds via intrinsic information studied in Ref. [CEH⁺07] to the case of an adversary which possibly can hold an infinite dimensional.

We then compare one of our upper bounds with the lower bound on the DI-CKA secret key rate in Ref. [RMW18]. Next, we provide a non-trivial upper bound on the device-independent key rate in the parallel measurement scenario in which all parties simultaneously set all values of their inputs. Furthermore, we provide a multipartite generalization of the reduced bipartite entanglement measures [KHD22]. Consequently, we show that the reduced regularized relative entropy of genuine entanglement [DBWH21] is an upper bound on the DI-CKA secret key rate. This relation brings us to the discussion on genuine nonlocality [BCP⁺14] and genuine entanglement in the context of the DI-CKA scenario [HHHH09]. Finally, we provide proof of the existence of a strict gap between the rates of the DI-CKA secret key and the DD-CKA secret key. The proof is constructive, as the gap is inherited from the bipartite case in which an example of states that exhibits the gap is known [CFH21]. To conclude, we are the first to provide upper bounds on the secret key rates in the DI-CKA cryptographic paradigm. At the same time, we provide several upper bounds on different scenarios and regimes.

In the fourth, recently published article [WDH22], we initiate a systematic study of upper bounds on the secret key rate in the non-signaling device-independent secret key agreement (NSDI) scenario [BHK05, BCK12, SGB⁺06, AMP06, AGM06, Mas09, HRW10, MRC⁺14]. Firstly, we show that the notion of security based on the so-called non-signaling norm (see also [CT09] in this topic), that we adopt, is equivalent to two security definitions present in the literature, i.e., to the one in Refs. [MRC⁺14, Mas09, MPA11] and to the second one in Refs. [HRW10, HRW13, HR10, Hän10]. To show the above-mentioned equivalence, we derive an explicit form of the non-signaling norm for the so-called \mathbf{c} -d states. Next, we show our main result, i.e., the squashing procedure. Namely, we show that any secrecy quantifier that is an upper bound on the secret key rate in the SKA scenario [Mau93, MW99, MW97, RW03, RSW03] can be lifted to give an upper bound on the secret key rate in the NSDI scenario. The squashing procedure is based on the notion of the non-signaling complete extension [WDS⁺18] and on the suitable optimization over measurements performed by the honest parties and the eavesdropper. The lift-up of the formalism is possible since we were able to rephrase the definition of the secret key rate in the SKA scenario in the form known from quantum key distribution scenarios (QKD). Subsequently, we focus on one of the squashed functions, i.e., the non-signaling squashed intrinsic information, also called the squashed nonlocality. As we show, the squashed nonlocality is not only an upper bound on the NSDI secret key rate but also a novel and nonfaithful measure of nonlocality. Next, we develop the convexification technique that provides even tighter upper bounds. Namely, in the convexification technique we combine different convex upper bounds to produce lower convex hull of their plots. The mentioned lower convex hull provides then tighter upper bound than individual plots on their own. Finally, we perform a numerical investigation that allowed us to compare our upper bound via the squashed nonlocality with the lower bounds given by the rates of Hänggi-Renner-Wolf [HRW10] and by Acín-Massar-Pironio [AMP06] protocols. In summary, we devised a technique that allows constructing upper bounds on the

secret key rate in the NSDI scenario from the upper bounds on the secret key rate in the SKA scenario. The upper bound via a nonfaithful measure of nonlocality shows that nonlocality is not always equivalent to secrecy. Our approach exposes analogies between different security paradigms. Therefore, our results contribute not only to the NSDI and SKA cryptographic paradigms but also to the fundamental topic of nonlocality.

In the fifth, recently accepted for publication manuscript [WDS⁺18], we study the concept of the complete extension postulate (CEP). We do so in the framework of the so-called generalized probabilistic theories (GPTs) [Bar07, Plá21, Mü121, Lam18] that allows us to study candidates for the beyond-quantum theory. By beyond-quantum theory, we mean the future theory of physics with stronger explanatory power than the quantum theory. As we discuss, the explanatory power of the quantum theory is limited, Refs. [LS18, SSC21, Nak20, WvdW22, Har05, Har07, OCB12, CDPV13, AFNB17] in this context. The CEP is devised to be a relaxation of the purification postulate, i.e., one of the postulates in terms of which quantum theory can be reformulated [CDP11]. In particular, contrary to the PP, the CEP does not require the existence, within the theory, of pure extensions for all systems and their states. Instead, the CEP requires the existence of extensions with the property of GENERATION, i.e., extensions that can be transformed into any other extension, of the extended system, with dynamics available within the theory. In the manuscript in Ref. [WDS⁺18], we study the properties of GPTs in which the PP is replaced with the CEP. We first show that the PP can not hold in any discrete convex theory, as within such theories, there are not enough pure states. On the contrary, the CEP holds both in the quantum theory and the classical probability theory, while the latter does not satisfy the PP. Next, we show that the property of GENERATION implies the property of ACCESS, i.e., the extending system of a generating extension has access to all statistical ensembles of the extended system via suitable choices of measurement settings. We show that replacing the PP with the CEP draws a demarcation line between the results that require the PP to hold and those for which the CEP is enough. In this context, we show that the impossibility of bit-commitment cryptographic task [May97, LC98], and its generalization to integer-commitment [SS18], holds within GPTs which satisfy the CEP. More precisely, we show that the lower bound on the product of the cheating probabilities of two parties in the integer-commitment task derived within the framework of quantum theory still holds if we replace the PP with the CEP. On the other hand, we show that the proof for the no-go for hyperdecoherence, which relies on the PP, no longer holds when the PP is replaced with the CEP [LS18]. The above suggests that the CEP might be a valid postulate for theories that hyperdecohere to the quantum theory. In order to show that the CEP is not an empty postulate and that complete extensions can actually be constructed outside quantum and classical theory, as our case study, we choose the theory of non-signaling behaviors [PR94, Bar07]. In this context, we first show that within the theory of non-signaling behaviors, the properties of ACCESS and GENERATION are equivalent. We then define the non-signaling complete extension (NSCE)

as an extension in which the extending system has direct access to all minimal ensembles upon suitable measurement choice. The minimal ensembles are simply ensembles that are not convex combinations of other ensembles. In fact, due to the dynamics available in the theory, access to minimal ensembles is equivalent to access to all ensembles by the extending system. Therefore, in analogy to quantum purifications, NSCEs satisfy both ACCESS and GENERATION. We then explicitly construct NSCEs in some cases. In particular, we show that the NSCE of maximally mixed behavior with a single binary input and single binary output is the famous Popescu-Rohrlich (PR) box [PR94]. We notice an analogy between NSCE described above and Bell's state, which is the purification of the maximally mixed state of a qubit. We see the above example as an independent derivation of the PR box in which we do not refer to any notion of Clauser-Horne-Shimony-Holt (CHSH) or other so-called Bell inequality. We also derive an upper bound on the dimension of the NSCE of an arbitrary behavior and show it has the lowest dimension amongst all extensions having the property of ACCESS. In this way, we show that the dimension of the NSCE of arbitrary behavior is always finite. This result is not only interesting on its own but also has consequences in the fields of device-independent cryptography against a non-signaling adversary [BHK05, Mas06, HRW10, Hän10] and private randomness [CR12, GMT⁺13, MGP15, BRG⁺16, RBH⁺16]. Namely, the adversary holding the extending system of NSCE possesses maximal operational power at the lowest memory cost required for it. To conclude, the CEP is an important relaxation of the PP in the context of studying possible beyond-quantum theories. Our findings reveal both the possibilities and the limitations of replacing the PP with the CEP. Our findings contribute not only to the field of GPTs but also to the field of device-independent cryptography against the non-signaling adversary.

Summary of Dissertation

2.1 Preliminaries

In this Section, we would like to provide formal definitions of the selected mathematical objects appearing in this dissertation. We aim here to provide a possibly high level of completeness of the descriptions of articles in the forthcoming Sections. Because of its formal and supplementary character, this Section can be skipped during the first reading of the dissertation. The descriptions of the articles composed in the dissertation contain suitable references to specific notions described below.

2.1.1 Private States and Positive Partial Transpose States

In Ref. [HHHO05], it was shown that the distillation of the secret key by two parties, A and B , that want to communicate secretly is nonequivalent to the distillation of Einstein-Podolsky-Rosen (EPR) states, i.e., states of the following form

$$(2.1) \quad |\psi_+\rangle_{AB} = \frac{1}{\sqrt{2}} (|00\rangle_{AB} + |11\rangle_{AB}).$$

Namely, there exist the so-called bound entangled states that can be used to obtain the secret key. To create a bound entangled state [HHH98], the resource of pure entanglement is required. However, no pure entanglement can be distilled from bound entangled states. The general form of quantum states that contain ideal secrecy was found in Ref. [HHHO05]. It was shown that a quantum state has an ideally secure key if and only if it has the form of the so-called private state $\gamma_{ABA'B'}$. Furthermore, private states were used in Ref. [HHHO09] to recast the theory of privacy, employing the notion of local operations and classical communication in terms of entanglement theory, where local operations and classical communication (LOCC)

are a valid set of operations. The LOCC paradigm does not require explicitly considering the eavesdropper's presence in the cryptographic scenario. Furthermore, in Ref. [HHHO09], equivalence between LOPC and LOCC secret key distillation paradigms was proved (see also Ref. [Hor09]). The above makes private states an important class of states, especially in the context of device-dependent quantum cryptography (DD-QKD) both in the bipartite and multipartite settings.

2.1.1.1 Bipartite Private States

In this Subsection, we provide the definitions of the bipartite private states and their special subclasses. In the definitions, we employ the notation used in Ref. [SWRH20] (see Sec. 2.2) contained in this dissertation rather than the original one in Refs. [HHHO05, HHHO09].

Definition 1 (Of a private state, cf. Ref. [HHHO09]). *A state $\rho_{ABA'B'}$ on a Hilbert space $\mathcal{C}^{d_A} \otimes \mathcal{C}^{d_B} \otimes \mathcal{C}^{d_{A'}} \otimes \mathcal{C}^{d_{B'}}$ with dimensions $d_A = d_b = d_k$, and $d_{A'} = d_{B'} = d_s$ of the form*

$$(2.2) \quad \gamma_{d_k, d_s} := \frac{1}{d_k} \sum_{i,j=0}^{d_k-1} |ii\rangle\langle jj|_{AB} \otimes X_{ij}, \quad X_{ij} := U_i \sigma_{A'B'} U_j^\dagger,$$

where the state $\sigma_{A'B'}$ is an arbitrary state of subsystem $A'B'$, U_i 's are arbitrary unitary transformations acting on $A'B'$ subsystem and $\mathcal{B}_A \equiv \{|i\rangle_A\}_{i=0}^{d_k-1}$, $\mathcal{B}_B \equiv \{|j\rangle_B\}_{j=0}^{d_k-1}$ are local (orthonormal) bases on A and B respectively, is called a private state or pbit. In the case of $d_k = 2$, the state γ_{2, d_s} is called - a pbit.

In the above definition, we refer to the subsystem AB as the key part because it contains directly accessible secret key via measurements in \mathcal{B}_A and \mathcal{B}_B bases. Similarly, the subsystem $A'B'$ is called the shielding system (the shield part). The shielding system has an important role in defining the LOCC secret key distillation paradigm. In the LOCC paradigm, it is assumed that no quantum system is discarded to outside of the laboratories of the honest parties. Instead, in the key distillation protocol, the shielding system serves as a dumpster, which is protected from the eavesdropper. The Definition 1, above, is constructed based on Definition 1 in Ref. [HCRS18]. Yet, we employ the notation used in Ref. [SWRH20] and restrict (with respect to Refs. [HHHO05, HHHO09, HCRS18]) to the case in which subsystems A' and B' are of the same dimension. See Sec. 2.1.1.3 for the most general definition of private states.

We now recall the definitions of special cases of private states. In the places where it does not lead to any confusion, we omit the subscript of γ_{d_k, d_s} and write simply γ . In the following definitions $K_D(\rho)$ is the secret key rate of state ρ (see Sec. 2.1.2).

Definition 2 (Of irreducible private state [HHHO09]). *Any pdit γ (with d_k -dimensional key part) for which $K_D(\gamma) = \log_2 d_k$ is called irreducible.*

Definition 3 (Of key-attacked private states [HCRS18]). *We call states of the following form key-attacked private states.*

$$(2.3) \quad \hat{\gamma}_{d_k, d_s} := \frac{1}{d_k} \sum_{i=0}^{d_k-1} |ii\rangle\langle ii| \otimes X_{ii},$$

where $X_{ii} = U_i \sigma U_i^\dagger$ for some state σ of $\mathcal{C}^{d_s} \otimes \mathcal{C}^{d_s}$ and U_i are some unitary transformations. They are private states which has been measured on their key part AB . We call X_{ii} conditional states of the shield.

Definition 4 (Of strictly irreducible private state [CF17, HCRS18]). *We say a private state γ_{d_k, d_s} is strictly irreducible, if its key-attacked version $\hat{\gamma}_{d_k, d_s}$ consists of separable states with respect to the systems on the shield part of the diagonal, i.e., $U_i \sigma U_i^\dagger \in \mathcal{SEP}$. We denote such states $\gamma_{\langle d_k, d_s \rangle}$, and their attacked versions by $\hat{\gamma}_{\langle d_k, d_s \rangle}$.*

We finish this Subsection with the following Remark.

Remark 1. *It follows directly from the definitions that $K_D(\gamma_{d_k, d_s}) \geq \log_2 d_k$, $K_D(\hat{\gamma}_{d_k, d_s}) = 0$, and $K_D(\gamma_{d_k, d_s}) = \log_2 d_k$ whenever γ_{d_k, d_s} is irreducible.*

2.1.1.2 Positive Partial Transpose States

In this Subsection, we introduce the definitions of the operation of partial transposition and positive partial transpose states (PPT). It follows directly from the definition of private states that they have negative partial transpose (NPT). In Ref. [HA06] it was shown that all private states allow for distillation of pure entanglement. On the other hand, all PPT states are bound entangled [HHH98]. Furthermore, there exist PPT states that are arbitrarily close, in trace norm, to private states, and they contain almost perfect secret key [HHHO05, HHHO09]. These feature of PPT states makes them an interesting class of states to study.

Definition 5 (Of partial transpose). *Let $\rho_{AB} = \sum_{ijkl} \rho_{kl}^{ij} |ik\rangle\langle jl|_{AB}$ be a bipartite quantum state acting on Hilbert space $\mathcal{H}_A \otimes \mathcal{H}_B$. The operation of partial transpose Γ_B with respect to the B subsystem is defined as follows*

$$(2.4) \quad \Gamma_B : \rho_{AB}^\Gamma \equiv (\mathbb{1} \otimes \Gamma_B) \rho_{AB} = \sum_{ijkl} \rho_{kl}^{ij} |il\rangle\langle jk|_{AB}.$$

Definition 6 (Of positive partial transpose (PPT) states). *Let ρ_{AB} be a bipartite quantum state acting on Hilbert space $\mathcal{H}_A \otimes \mathcal{H}_B$. State ρ_{AB} is called positive partial transpose (PPT) state, if and only if ρ_{AB}^Γ has only non-negative eigenvalues, i.e.,*

$$(2.5) \quad \rho_{AB}^\Gamma \geq 0,$$

where Γ is a partial transpose operation with respect to arbitrary subsystem.

2.1.1.3 Multipartite Private States

Private states can also be defined in the multipartite setting and help to define the LOCC paradigm for conference key agreement (DD-CKA). In this Section, we assimilate the notation used in Ref. [DBWH21] (see Sec. 2.3). Let us consider M parties $A \equiv A_1, \dots, A_M$ with $\vec{K} \equiv K_1, \dots, K_M$ key parts and $\vec{S} \equiv S_1, \dots, S_M$ shielding systems, defined on $\mathcal{H} \equiv \mathcal{H}_1 \otimes \dots \otimes \mathcal{H}_M$ and $\mathcal{H}' \equiv \mathcal{H}'_1 \otimes \dots \otimes \mathcal{H}'_M$, respectively.

In the following definition, $\Phi_{\vec{K}}^{\text{GHZ}}$ denotes the projection on M -partite Greenberger-Horne-Zeilinger (GHZ) state, i.e., the state of the following form

$$(2.6) \quad \vec{k} \equiv k_1, \dots, k_M : \forall_{i,j \in [M]} k_i = k_j \in \mathcal{K},$$

$$(2.7) \quad |\text{GHZ}\rangle := \frac{1}{K} \sum_{(2.6)} |\vec{k}\rangle.$$

Here, \mathcal{K} corresponds to the alphabet of outcomes for key parts of individual subsystems, and $K \equiv |\mathcal{K}| = |K_i|$ denotes the cardinality of \mathcal{K} .

Definition 7 (Of M -partite private state, cf. Refs. [AH09b, DBWH21]). *A state $\gamma_{\vec{SK}}$, with $|K_i| = K$ for all $i \in [M]$, is called a (M -partite) private state if and only if*

$$(2.8) \quad \gamma_{\vec{SK}} := U_{\vec{SK}}^{\text{tw}} \left(\Phi_{\vec{K}}^{\text{GHZ}} \otimes \omega_{\vec{S}} \right) \left(U_{\vec{SK}}^{\text{tw}} \right)^\dagger,$$

where $U_{\vec{SK}}^{\text{tw}} := \sum_{\vec{k} \in \mathcal{K}^{\times M}} |\vec{k}\rangle\langle \vec{k}|_{\vec{K}} \otimes U_{\vec{S}}^{\vec{k}}$ is called a twisting unitary operator for some unitary operator $U_{\vec{S}}^{\vec{k}}$ and $\omega_{\vec{S}}$ is an arbitrary finite-dimensional density operator.

The Definition 7, above, is more general than Definition 1, but also because we do not assume that different shielding systems are of the same dimension. It follows directly from the definition that $K_D(\gamma_{\vec{SK}}) \geq \log_2 K$. We finish with the following remark.

Remark 2. *In this place, we remark that the definitions 2, 3 and 4, defining different subclasses of private states, can be naturally extended to the multipartite case.*

2.1.2 Secret Key Agreement Tasks

The secret key agreement is a cryptographic task in which two or more honest allies aim to distill the secure key that can be used for one-time pad encryption. A scenario that involves more than two parties is called a conference key agreement. More precisely, the task of the honest parties is to use resources and operations available in the scenario in order to obtain perfectly correlated bit strings which remain unknown to the eavesdropper. The sequence of operations performed by the honest parties, which may also contain communication via an authenticated but insecure classical channel, is called a protocol. Similarly, the action of the malicious eavesdropper is called the attack strategy. The particular resources and operations

available in the cryptographic task depend on the considered cryptographic paradigm (see Chap. 1). The crucial mathematical objects quantifying the scenario are key rate and security conditions. The rate of a protocol is, roughly speaking, the number of bits of secret key distilled in protocol per one copy of the resource (state). The secret key rate is, therefore, the supremum over rates of all possible protocols. The secret key rate can be defined in both one-shot (single use) regimes, where the number of copies of the resource state is finite, as well as in the asymptotic regime, in which we assume that the number of copies of the resource state is arbitrarily large. The security conditions define what is understood by the secure key. The key distillation protocol is, therefore, called secure if the correlations shared by the honest parties at the end of the protocol satisfy the security conditions. The definitions of those are, again, scenario-dependent. In what follows, we assimilate the so-called “optimistic” definition of asymptotic key rate in which “ $\limsup_{n \rightarrow \infty}$ ” is used rather than “ $\liminf_{n \rightarrow \infty}$ ”, where n refers to the number of copies of the resource state, as it is done in, e.g., in Ref. [BCHW15]. For the discussion about the “optimistic” and the “pessimistic” definitions of the asymptotic key rate, see Refs. [Ahl06]. See also a different approach to ours in the case of entanglement theory, e.g., in [Wil21].

2.1.2.1 Secret Key Rate and Security

We first describe the definition of secret key rate in the secret key agreement scenario (SKA) [Mau93, MW00]. As described in Chapter 1, the resource states in the SKA scenario are bipartite probability distributions $P(AB)$, and A and B are random variables in possession of respective parties. The honest parties, also called A and B , process their marginal probability distributions using local operations in the form of local stochastic maps and public communication (LOPC). The eavesdropper E is explicitly present in the scenario and is assumed to share with the honest parties a tripartite probability distribution $P(ABE)$ that is an extension of $P^{\otimes N}(AB)$, i.e., $\sum_e P(ABe) = P^{\otimes N}(AB)$. The secret key rate $S(A : B||Z)$ in the SKA scenario is defined as follows

Definition 8 (Of the SKA secret key rate, cf. Ref. [Mau93, MW00]). *The secret key rate of A and B with respect to E , denoted $S(A : B||Z)$, is the maximal $R \geq 0$ such that for every $\varepsilon > 0$ and for all $N \geq N_0(\varepsilon)$ there exists a protocol, using public communication over an insecure but an authenticated channel, such that Alice and Bob, who receive $A^N = [A_1, \dots, A_N]$ and $B^N = [B_1, \dots, B_N]$, can compute keys S_A and S_B , respectively, with the following properties. First, $S_A = S_B$ hold with probability at least $1 - \varepsilon$, and second,*

$$(2.9) \quad \frac{1}{N} I(S_A : CE^N) \leq \varepsilon \quad \text{and} \quad \frac{1}{N} H(S_A) \geq R - \varepsilon$$

hold. Here, C denotes the collection of messages sent over the insecure channel by Alice and Bob.

In the Definition 8 above, $I(X : Y)$ is the classical mutual information function, and $H(X)$ is the Shannon entropy.

The SKA scenario can be naturally extended to the setting that involves more than two honest parties, i.e., conference key agreement CKA [YHH⁺09]. Following the notation in Ref. [YHH⁺09], the object shared by the m honest parties ($(m)A \equiv A_1 \dots A_m$) is m -partite probability distribution $P_{(m)A}$ defined on $\mathcal{X}_{(m)A} \equiv \mathcal{X}_{A_1} \times \dots \times \mathcal{X}_{A_m}$. Here, \mathcal{X}_{A_i} denotes the set values of random variable A_i in possession of the i -th party. The n copies of such distribution belong then to $\mathcal{X}_{A_1}^n \times \dots \times \mathcal{X}_{A_m}^n \equiv \mathcal{X}_{(m)A}^n$. In analogy to SKA scenario the extension $P_{(m)AE}^n$ of $P_{(m)A}^{\otimes n}$ to the eavesdropper system is defined as a probability distribution on $\mathcal{X}_{(m)AE}^n \equiv \mathcal{X}_{A_1}^n \times \dots \times \mathcal{X}_{A_m}^n \times \mathcal{X}_E$, where \mathcal{X}_E is the set of values of random variable E in possession of the eavesdropper. We first, provide the definition of the LOPC protocol in CKA scenario

Definition 9 (Of CKA protocol, cf. Ref. [YHH⁺09]). *An LOPC protocol \mathcal{P} is a family $\{\Lambda\}_n$ of classical channels $\Lambda_n : \left(\mathcal{X}_{(m)AE}^n\right) \rightarrow \mathcal{X}_{(m)A'E'}$ which are a finite number of concatenation of local operations (local channels) and public communications steps (communication between honest parties via authorised but insecure classical channel).*

The secret key rate $C_D^{(m)}$ in the m -partite CKA scenario is then defined as follows

Definition 10 (Of CKA secret key rate, cf. Ref. [YHH⁺09]). *We say that LOPC protocol \mathcal{P} is a classical key distillation protocol for a distribution $P_{(m)A} \in \mathcal{X}_{(m)A}$, if*

$$(2.10) \quad \lim_{n \rightarrow \infty} \left\| \Lambda^n \otimes \mathbb{1}_E \left(P_{(m)AE}^n \right) - K_{(m)AE}^{l_n} \right\|_1 = 0,$$

where $K_{(m)AE}^{l_n} = \frac{1}{l_n} (i_1 \dots i_m) \delta_{i_1, \dots, i_m} \otimes P_E$ is the ideal key distribution on $\mathcal{X}_{(m)AE}$, for some distribution P_E of the eavesdropper. The rate of the protocol is given by

$$(2.11) \quad \mathcal{R}(\mathcal{P}) = \limsup_{n \rightarrow \infty} \frac{\log_2 l_n}{n}.$$

The classical distillable key of a distribution $P_{(m)A}$ is defined as supremum of the rates

$$(2.12) \quad C_D^{(m)}(P_{(m)A}) = \sup_{\mathcal{P}} \mathcal{R}(\mathcal{P}).$$

Definition 8 and Definition 10 are seemingly different. The first one refers to the notion of mutual information and Shannon Entropy, whereas the second uses trace norm distance $\|\cdot\|_1$ between the output state of the protocol and the ideal state. In Ref. [WDH22] (see also Sec. 2.5), the equivalence between the mentioned definitions was proved in the bipartite case.

In the case of the device-dependent quantum key distribution (DD-QKD), the resources used in the cryptographic tasks are entangled quantum states. The set of operations available to the honest parties are, basically, transformations of quantum states and measurements on them. The secret key rate can then be defined both in LOPC and LOCC paradigms [HHHO09]. For

the sake of this dissertation, we provide the definition of DD-CKA secret key rate referring to the LOCC paradigm that employs the notion of private states (see Sec. 2.1.1). In the following definition of device-dependent conference key rate $K_D^{(m)}$, $\rho_{(m)A}$ is a multipartite quantum state shared by m -honest parties ($(m)A \equiv A_1 \dots A_m$), n copies of which are shared by the honest parties at the beginning of the conference key distillation protocol.

Definition 11 (Of DD-CKA protocol and conference key rate, cf. Ref. [AH09b, YHH⁺09]). *For any given state $\rho_{(m)A} \in \mathcal{B}(\mathcal{H}_{(m)A})$ let us consider a sequence P_n of LOCC operations such that $P_n(\rho_{(m)A}^{\otimes n}) = \sigma_n$. A set of operations $\mathcal{P} \equiv \bigcup_{n=1}^{\infty} \{P_n\}$ is called a pdit distillation protocol of state $\rho_{(m)A}$ if there holds*

$$(2.13) \quad \lim_{n \rightarrow \infty} \left\| \sigma_n - \gamma_{d_n}^{(m)} \right\|_1 = 0,$$

where $\gamma_{d_n}^{(m)}$ is a multipartite pdit whose key part is of dimension $d_n^{(1)} \times \dots \times d_n^{(m)}$. For a \mathcal{P} , its rate is given by

$$(2.14) \quad \mathcal{R}(\mathcal{P}) = \limsup_{n \rightarrow \infty} \frac{\log_2 d_n}{n}.$$

The distillable key of state $\rho_{(m)A}$ is given by

$$(2.15) \quad K_D^{(m)}(\rho_{(m)A}) = \sup_{\mathcal{P}} \mathcal{R}(\mathcal{P}).$$

Definitions of secret key rate 8, 10 and 11 have an asymptotic character. Namely, they assume that the available number of copies of the resource state is not limited. Hence, the supremum over protocols in definitions of the rates also incorporates protocols that require an arbitrary large number of copies. On the other hand, in practical situations, the number of copies of the resource state is limited. Therefore, the number of bits of the secret key that can be distilled (per copy of resource state) from a finite number of copies is an important quantity to study. The following definition of one-shot (single run) secret key rate in the DD-CKA scenario corresponds to the one implicitly used in Ref. [DBWH21].

Definition 12 (One-shot secret key rate in DD-CKA scenario, cf. Ref [AH09b, YHH⁺09]). *For any given state $\rho_{(m)A} \in \mathcal{B}(\mathcal{H}_{(m)A})$ let us consider a LOCC operation Λ_n such that $\Lambda_n(\rho_{(m)A}^{\otimes n}) = \sigma_n$. We call Λ_n a ε -secure key distillation protocol from n copies of state $\rho_{(m)A}$ if there holds*

$$(2.16) \quad F(\sigma_n, \gamma_{d_n}^{(m)}) \geq 1 - \varepsilon,$$

where $\gamma_{d_n}^{(m)}$ is a multipartite pdit whose key part is of dimension $d_n^{(1)} \times \dots \times d_n^{(m)}$, and $F(\cdot, \cdot)$ denotes the fidelity between quantum states. For a Λ_n , its rate is given by

$$(2.17) \quad \kappa_n^\varepsilon(\Lambda_n(\rho_{(m)A}^{\otimes n})) := \frac{\log_2 d_n}{n},$$

and the rate of ε -secure secure key distilled from n copies of state $\rho_{(m)A}$, by protocols satisfying the above security condition, is given by

$$(2.18) \quad K_D^{(n,\varepsilon)}(\rho_{(m)A}) := \sup_{\Lambda_n \in \text{LOCC}} \kappa_n^\varepsilon \left(\Lambda_n(\rho_{(m)A}^{\otimes n}) \right).$$

Finally, the one-shot ε -secure secret key is defined as

$$(2.19) \quad K_D^{s,\varepsilon}(\rho_{(m)A}) := K_D^{(1,\varepsilon)}(\rho_{(m)A}).$$

Eventually, let us remark that the following relation holds between asymptotic and one-shot DD-CKA secret key rates, namely

$$(2.20) \quad K_D^{(m)}(\rho) = \inf_{\varepsilon > 0} \limsup_{n \rightarrow \infty} K_D^{(n,\varepsilon)}(\rho),$$

where m denotes the number of honest parties. The Definition 12 is written in the LOCC paradigm, and the security condition is based on the fidelity between quantum states. On the contrary, the Definition 11 is given in the LOPC paradigm where the security condition is based on the trace norm. The relation in Eq. (2.20) follows from the equivalence between LOCC and LOPC paradigms [HHHO09] and Fuchs-van de Graaf inequalities [FvdG99].

In the device-independent conference key agreement scenario (DI-CKA), the N honest parties are assumed to share an untrusted device, described with input-output statistics $\{p(\mathbf{a}|\mathbf{x})\}_{\mathbf{a}|\mathbf{x}}$ originating from tensor product POVM (positive operator-valued measure) measurements $\mathcal{M} \equiv \{M_{a_1}^{x_1} \otimes M_{a_2}^{x_2} \otimes \dots \otimes M_{a_N}^{x_N}\}_{\mathbf{a}|\mathbf{x}}$ on quantum state $\rho_{N(A)}$, i.e., a pair $(\rho_{N(A)}, \mathcal{M})$. Here, $N(A) \equiv A_1 \dots A_N$, where A_i refers to the subsystems of the i -th party, and $\mathbf{x} := (x_1, x_2, \dots, x_N)$ and $\mathbf{a} := (a_1, a_2, \dots, a_N)$ refer to the inputs and outputs of the honest parties, respectively. The attack strategy of the adversary is to replace a device $(\rho_{N(A)}, \mathcal{M})$ with another device $(\sigma_{N(A)}, \mathcal{N})$ which yields the same attack statistics as the honest one. The power of the adversary is restricted then to the laws of quantum theory, and therefore, the adversary may hold the purifying system of $\sigma_{N(A)}$, i.e., the E subsystem of the joint state $|\psi_\sigma\rangle_{N(A)E}$. We further omit writing subscripts like $N(A)$ in places in which it does not lead to any confusion.

Consider the following relations that are a basis for two different security conditions in the DI-CKA scenario, i.e., “par” and “dev” definitions.

$$(2.21) \quad (\rho, \mathcal{M}) \approx_\varepsilon (\sigma, \mathcal{N}),$$

$$(2.22) \quad \omega(\rho, \mathcal{M}) \approx_\varepsilon \omega(\sigma, \mathcal{N}),$$

$$(2.23) \quad P_{err}(\rho, \mathcal{M}) \approx_\varepsilon P_{err}(\sigma, \mathcal{M}),$$

where $\omega(\rho, \mathcal{M})$ is the level of violation of some (chosen by honest parties) multipartite Bell inequality, and $P_{err}(\rho, \mathcal{M})$ is quantum bit error rate (QBER). Furthermore, relation in Eq. (2.21) means that two input-output statistics p and p' originating from devices (ρ, \mathcal{M}) and (σ, \mathcal{N}) respectively, satisfies

$$(2.24) \quad d(p, p') = \sup_{\mathbf{x}} \|p(\cdot|\mathbf{x}) - p'(\cdot|\mathbf{x})\|_1 \leq \varepsilon.$$

We restrict ourselves to the definitions of the DI-CKA secret key rates in the independent and identically distributed (iid) setting. See Definition 12 for the meaning of κ_n^ε .

Definition 13 (Of “dev” iid DI-CKA secret key rate, cf. Ref [CFH21]). *The (multipartite) device-independent quantum key distillation rate of a device (ρ, \mathcal{M}) with independent and identically distributed behavior is defined as*

$$(2.25) \quad K_{DI,dev}^{iid}(\rho, \mathcal{M}) := \inf_{\varepsilon>0} \limsup_{n \rightarrow \infty} \sup_{\hat{\mathcal{P}}} \inf_{(2.21)} \kappa_n^\varepsilon \left(\hat{\mathcal{P}}((\sigma, \mathcal{N})^{\otimes n}) \right),$$

where κ_n^ε is the rate of a key distillation protocol $\hat{\mathcal{P}}$ producing ε -secure output, acting on n copies of the state σ , measured with \mathcal{N} . Here $\hat{\mathcal{P}}$ is a protocol composed of classical local operations and public (classical) communication (CLOPC) acting on n identical copies of (σ, \mathcal{N}) which, composed with the measurement, results in a quantum local operations and public (classical) communication (QLOPC) protocol.

Definition 14 (Of “par” iid DI-CKA secret key rate, cf. Ref. [AFL21]). *The (multipartite) device-independent quantum key distillation rate of a device (ρ, \mathcal{M}) with independent and identically distributed behavior, Bell inequality violation $\omega(\rho, \mathcal{M})$, and QBER $P_{err}(\rho, \mathcal{M})$ is defined as*

$$(2.26) \quad K_{DI,par}^{iid}(\rho, \mathcal{M}) := \inf_{\varepsilon>0} \limsup_{n \rightarrow \infty} \sup_{\hat{\mathcal{P}}} \inf_{(2.22),(2.23)} \kappa_n^\varepsilon \left(\hat{\mathcal{P}}((\sigma, \mathcal{N})^{\otimes n}) \right).$$

Definition 15 (Of single-shot DI-CKA secret key rate, cf. Ref. [HWD22]). *The single-shot device-independent quantum key distillation rate of a device (ρ, \mathcal{M}) with independent and identically distributed behavior is defined as*

$$(2.27) \quad K_{DI,dev}^{\text{single-shot}}(\rho, \mathcal{M}, \varepsilon) := \sup_{\hat{\mathcal{P}}} \inf_{(2.21)} \kappa_n^\varepsilon \left(\hat{\mathcal{P}}(\sigma, \mathcal{N}) \right),$$

where κ_n^ε is the quantum key rate achieved for any security parameter ε and measurements \mathcal{N} . Here $\hat{\mathcal{P}}$ is a protocol composed of classical local operations and public (classical) communication acting on a single copy of (σ, \mathcal{N}) which, composed with the measurement, result in local quantum operations and public (classical) communication protocol.

Definition 16 (Of the reduced DD-CKA secret key rate, cf. [CFH21, HWD22]). *The reduced device-dependent conference key rate of an N -partite state $\rho_{N(A)}$ reads*

$$(2.28) \quad K^\downarrow(\rho_{N(A)}) := \sup_{\mathcal{M}} \inf_{(\sigma_{N(A)}, \mathcal{L}) = (\rho_{N(A)}, \mathcal{M})} K_{DD}(\sigma_{N(A)}).$$

In the Definition 16 above, K_{DD} correspond to $K_D^{(N)}$ in Definition 11. The reduced DD-CKA secret key rate is an important upper bound on the iid DI-CKA secret key rate [CFH21].

In the non-signaling device-independent secret key agreement scenario (NSDI), the two honest parties share N copies of the non-signaling probability distribution (non-signaling device)

$P(AB|XY)$, being the aforementioned resource, where X and Y denote the inputs and A and B correspond to outputs of the honest parties. In the NSDI scenario, the eavesdropper is limited solely with the no-signaling conditions [Bar07] and possesses an extending system of the device $P(AB|XY)$. The device shared by all three parties is, therefore, described with tripartite conditional probability distribution $Q(ABE|XYZ)$, which has N copies of $P(AB|XY)$ as its marginal state. Here, E and Z denote the output and input of the system held by the eavesdropper. Moreover, $Q(ABE|XYZ)$ can exhibit a supra-quantum correlation between its subsystems [PR94]. In the considerations in Ref. [WDH22], we assume that the extension held by the eavesdropper is the non-signaling complete extension (NSCE) of N independent and identically distributed (iid) copies of the device $P(AB|XY)$, i.e., $\mathcal{E}(P^{\otimes N})(ABE|XYZ)$ [WDS+18] (see also Sec. 2.6). From the perspective of the honest parties, the NSCE is the worst-case extension the eavesdropper may use. This power is because NSCE allows access to all statistical ensembles of the extended system via suitable measurement and processing on the extending system [WDS+18].

We now provide the necessary definitions of MDLOPC operations, secret key distillation protocol and ideal state, its security condition, and the secret key rate in the iid NSDI scenario.

Definition 17 (Of MDLOPC operations, cf. Ref. [WDH22]). *The MDLOPC operation consists of (i) direct measurement $\mathcal{M}_{x,y}^F$ on inputs X and Y of the honest parties, that changes the device into a probability distribution, followed by (ii) arbitrary bipartite LOPC operations. Direct measurement is a measurement that does not incorporate any external randomness on the measured input.*

Definition 18 (Of the MDLOPC protocol in the iid NSDI scenario, cf. Refs. [HRW10, WDH22]). *A protocol of key distillation is a sequence of MDLOPC operations $\Lambda \equiv \{\Lambda_N\}$, performed by the honest parties on N iid copies of the shared devices. Each of this Λ_N , consists of a measurement stage $\{\mathcal{M}_N\}$, followed by post-processing $\{\mathcal{P}_N\}$, on N iid copies of $P(AB|XY)$. Moreover, for each consecutive, complete extension of N copies of shared devices $\mathcal{E}(P^{\otimes N})(ABE|XYZ)$, the protocol outputs a probability distribution in part of Alice and Bob and a device in part of Eve, which is arbitrarily close to an ideal distribution, and satisfies*

$$(2.29) \quad \|P_{\text{out}} - P_{\text{ideal}}^{(d_N)}\|_{\text{NS}} \leq \varepsilon_N \xrightarrow{N \rightarrow \infty} 0,$$

$$(2.30) \quad P_{\text{ideal}}^{(d_N)}(s_A, s_B, q, e|z) := \frac{\delta_{s_A, s_B}}{|S_A|} \sum_{s'_A, s'_B} P_{\text{out}}(s'_A, s'_B, q, e|z).$$

Here, $P_{\text{out}} = \Lambda_N \left(\mathcal{E} \left(P^{\otimes N} \right) \right)$, and random variables S_A , S_B and Q correspond to the keys of the honest parties and communication respectively. Moreover, $\mathbf{A} = A_1 A_2 \dots A_N$, \mathbf{B} , \mathbf{X} and \mathbf{Y} are similarly defined.

In the Definition 18 above $\|\cdot\|_{\text{NS}}$ denotes the non-signaling norm described in Sec. 2.5 (see also Ref. [WDH22]).

Definition 19 (Of the secret key rate in the iid NSDI scenario, cf. Ref. [WDH22]). *Given a bipartite device $P \equiv P(AB|XY)$ the secret key rate of the protocol of key distillation Λ_N , on N iid copies of the device, denoted by $\mathcal{R}(\Lambda|_P)$ is a number $\limsup_{N \rightarrow \infty} \frac{\log d_N}{N}$, where $\log d_N$ is the length of a secret key shared between Alice and Bob, with $d_N = \dim_A \left(\Lambda_N \left(\mathcal{E} \left(P^{\otimes N} \right) \right) \right) \equiv |S_A|$. The device independent key rate of the iid scenario is given by*

$$(2.31) \quad K_{DI}^{(iid)}(P) = \sup_{\Lambda} \mathcal{R}(\Lambda|_P),$$

where the supremum is taken over all MDLOPC protocols $\{\Lambda_N\}$.

In Ref. [WDH22], we show that the Definition 19 of the iid NSDI secret key rate is equivalent in terms of security to the one present in the literature [MRC⁺14, Mas09, MPA11, HRW13, HR10, HRW10, Hän10].

2.1.2.2 Key Repeater Rate

Quantum repeaters and quantum key repeaters [BDCZ98, DBCZ99, MATN15, CZC⁺21] are devices that allow overcoming attenuation of (e.g., optical) signal, and decoherence of quantum states, in order to entangle quantum systems separated by arbitrarily large distances. The idea of quantum repeaters is based on the entanglement swapping protocol [BDCZ98], which is performed inside a quantum repeater. In the simplest setup with one intermediate station denoted here H , the joint measurement in “entangled” basis performed in central station H on subsystems $A'E'$ of two initially uncorrelated but entangled quantum states $\rho_{AA'}$ and $\rho_{EE'}$ can create entanglement between subsystems A and E that are spatially separated [BDCZ98, BDSW96]. The measurement outcome at hub station H is then communicated to stations A and E . Importantly, quantum repeaters can be combined in chains that have an arbitrary number of hub stations. Consequently, the distance at which entanglement can be created increases. The idea of quantum key repeaters generalizes the concept of quantum repeaters. The task of a quantum key repeater is to distill private states (see Sec. 2.1.1.1) between stations A and E using tripartite LOCC operations among nodes A , E , and H . The communication between the central station H and stations A and E can be considered both one- and two-way. The rate of a quantum key repeater is defined as follows

Definition 20 (Of quantum key repeater rate, cf. Ref. [BCHW15, CF17]). *The quantum key repeater rate with respect to arbitrary tripartite LOCC operations among A , E and H is defined as follows*

$$(2.32) \quad R^{A \leftrightarrow H \leftrightarrow E}(\rho, \rho') := \lim_{\epsilon \rightarrow 0} \limsup_{n \rightarrow \infty} \sup_{\Lambda_n^{LOCC, \gamma_{d_k, d_s}}} \left\{ \frac{\log d_k}{n} : \text{tr}_H \Lambda_n^{LOCC}((\rho \otimes \rho')^{\otimes n}) \approx_{\epsilon} \gamma_{d_k, d_s} \right\},$$

where Adam and Hub share state ρ while Hub and Eve share ρ' . $\Lambda := \{\Lambda_n^{LOCC}\}$ are tripartite LOCC protocols with two-way classical communication between nodes H , A , and E . In the case in which communication between central node and A, E systems is restricted to one-way from H to A and E , we denote this rate with $R^{H \rightarrow A: E}$.

In the above definition, γ_{d_k, d_s} refers to a private state distilled in a quantum key repeater (see Sec. 2.1.1.1).

2.1.3 Quantum Channels

Quantum channels, also known as quantum operations, are the most general form of transformations between quantum states (see Ref. [NC00] in this context). The (memoryless) quantum channel can be defined as follows

Definition 21 (Of quantum operations, aka quantum channels, cf. Ref [NC00]). *Let $\mathcal{D}(\mathcal{H})$ and $\mathcal{D}(\mathcal{H}')$ define the set of density operators acting on two arbitrary Hilbert spaces \mathcal{H} and \mathcal{H}' , respectively. Any completely positive and trace preserving (CPTP) linear map $\Lambda : \mathcal{D}(\mathcal{H}) \rightarrow \mathcal{D}(\mathcal{H}')$ is called a quantum operation (quantum channel).*

In Ref. [DBWH21] (see also Sec. 2.3), we specify the use of quantum channels in situations concerning conference key agreement, in which the spatial distribution of the subsystems after the action of a quantum channel does not correspond to the initial distribution. We distinguish the roles of the senders and receivers \mathbf{A}_a , senders \mathbf{B}_b , and receivers \mathbf{C}_c . We call such types of quantum channels multiplex quantum channels. In the following definition, we use a notation in which $\vec{\mathbf{X}} \equiv \mathbf{X}_1 \dots \mathbf{X}_K$, where K is the number parties of type $\mathbf{X} \in \{\mathbf{A}, \mathbf{B}, \mathbf{C}\}$.

Definition 22 (Of multiplex quantum channel, cf. Ref. [DBWH21]). *Consider multipartite quantum channel $\mathcal{N}_{\vec{\mathbf{A}}' \vec{\mathbf{B}} \rightarrow \vec{\mathbf{A}} \vec{\mathbf{C}}}$ where each pair A'_a, A_a is held by a respective party \mathbf{A}_a and each B_b, C_c are held by parties $\mathbf{B}_b, \mathbf{C}_c$, respectively. While \mathbf{A}_a is both sender and receiver to the channel, \mathbf{B}_b is only a sender, and \mathbf{C}_c is only a receiver to the channel. Such a quantum channel is referred to as the multiplex quantum channel. Any two different systems need not be of the same size in general.*

In principle, multiple uses of multiplex quantum channel interleaved by LOCC can simulate a wide range of QKD protocols, which are secure in the sense of the device-dependent paradigm. It is, therefore, useful to quantify conference key agreement rates achievable with particular multiplex quantum channels. The formal description of LOCC-assisted conference key agreement protocol over a multiplex quantum channel $\mathcal{N}_{\vec{A}'\vec{B}\rightarrow\vec{A}\vec{C}}$ and the definition of secret-key-agreement capacity $\hat{P}_{\text{LOCC}}(\mathcal{N})$ of quantum channel $\mathcal{N}_{\vec{A}'\vec{B}\rightarrow\vec{A}\vec{C}}$ are provided in Sec. 2.3 (see also Ref. [DBWH21]). The same concerns the definition of cppp-assisted (one-shot) secret-key-agreement capacities $\hat{P}_{\text{cppp}}^{(1,\varepsilon)}$ of multiplex quantum channels. Here, cppp refers to classical preprocessing and postprocessing, which are operations employed in the protocol.

2.1.4 Entanglement and Genuine Entanglement

The presence of entanglement in quantum theory is one of the basic features that differentiate quantum theory from the classical theory of Nature. It was already observed by A. Einstein, B. Podolsky, N. Rosen [EPR35] and E. Schrödinger [Sch35] that systems described with entangled quantum states exhibit correlations that can not be explained with the classical theory of probability. The discussion about the meaning of quantum entanglement in modern physics is beyond the scope of this dissertation (see Ref. [HHHH09]). The entangled quantum states are defined as states that are not fully separable, i.e.,

Definition 23 (Of fully separable and entangled states, cf., e.g., Ref. [AH09b]). *If a N -partite quantum state $\rho_{N(A)}$, where $N(A) = A_1 \dots A_N$, can be written as a convex combination of tensor product of states $\rho_{A_j}^{(i)}$ of individual subsystems*

$$(2.33) \quad \rho_{N(A)} = \sum_i p_i \rho_{A_1}^{(i)} \otimes \dots \otimes \rho_{A_N}^{(i)},$$

then we call it a full separable state. All separable states constitute a set of fully separable states (FS). All other states are entangled.

We remark here, that operationally entangled quantum states are quantum states that can not be prepared with (multipartite) LOCC operations. In the bipartite setting the set of (fully) separable states is denoted \mathcal{SEP} , and called the set of separable states.

According to Definition 23 a quantum state of the form $\rho_{A_1} \otimes \sigma_{A_2 A_3}$ is entangled, if and only if, $\sigma_{A_2 A_3}$ is entangled. However, system A_1 is not entangled with subsystems A_2 and A_3 . Indeed, it is “biseparable” as it is of product form in (at least) one “cut” between its subsystems. It is, therefore, convenient, also from the cryptographic perspective, to define quantum states in which each subsystem A_i is entangled with any other subsystem $A_{\neq i}$ or collection of them $A_{\neq i} \dots A'_{\neq i}$.

Definition 24 (Of biseparable state, cf. Ref. [HHHH09, DBWH21]). *If a N -partite quantum state $\rho_{N(A)}$ can be written as a convex combination of N -partite states that are separable with respect to at least one partition of its subsystems*

$$(2.34) \quad \rho_{N(A)} = \sum_i p_i \rho_{K_i(A)}^{(i)} \otimes \rho_{M_i(A)}^{(i)}, \quad K_i + M_i = N,$$

then we call it a biseparable state. All separable states constitute a set of biseparable states (BS). All other states are genuine entangled (GE).

In the Definition 24, the word genuine emphasizes that the state therein, indeed, can not be prepared by LOCC without the resource of N -partite entangled pure states. In this place, we remark that the sets of fully separable (FS) and biseparable (BS) states are convex sets. However, the set of biseparable states is not closed under the tensor product. The subset of biseparable states that is closed under the tensor product is called the set of tensor-stable biseparable states and is defined as follows

Definition 25 (Of tensor-stable biseparable states, cf. Ref. [PdV22, DBWH21]). *If any tensor power of some biseparable state $\rho_{N(A)}$ is a biseparable state, i.e.,*

$$(2.35) \quad \forall_{n \geq 1} \rho_{N(A)}^{\otimes n} \in \text{BS},$$

we call the state $\rho_{N(A)}$ tensor-stable biseparable state.

2.1.5 Nonlocality and Genuine Nonlocality

Nonlocality is a notion that refers to the violation of Bell inequalities by quantum systems [BCP⁺14]. In analogy to the case of the theory of entanglement (see Sec. 2.1.4), one can define the notions of local and genuine nonlocal quantum behaviors within the theory of nonlocality [BCP⁺14]. Quantum behaviors $P(\mathbf{a}|\mathbf{x})$ are conditional probability distributions that originate from measurement on quantum states (see Secs. 2.1.2.1 and 2.4 for more details).

Definition 26 (Of behavior local in a “cut”, cf. Ref. [HWD22]). *We say that N -partite behavior $P(\mathbf{a}|\mathbf{x})$ is local in a cut $(A_{i_1} \dots A_{i_k}) : (A_{i_{k+1}} \dots A_{i_N})$ for some $k \in \{1, \dots, N\}$, if it can be written as a product of two behaviors on systems $A_{i_1} \dots A_{i_k}$ and $A_{i_{k+1}} \dots A_{i_N}$, respectively, i.e.,*

$$(2.36) \quad P(\mathbf{a}|\mathbf{x}) = P_{1,k}(a_{i_1} \dots a_{i_k} | x_{i_1} \dots x_{i_k}) \otimes P_{k+1,N}(a_{i_{k+1}} \dots a_{i_N} | x_{i_{k+1}} \dots x_{i_N}),$$

for (i_1, \dots, i_N) being some permutation of indices $(1, \dots, N)$

There are various definitions of locality of behaviors in multipartite case [BCP⁺14]. In Ref. [HWD22] we use the following definition of locality.

Definition 27 (Of locally quantum behaviors, cf. Ref. [BCP⁺14, HWD22]). *The locally quantum behaviors are convex mixtures of behaviors that are a product in some cut and both behaviors in the product have quantum realization.*

The class of locally quantum behaviors in Definition 27 falls between the so-called TOBL and NSBL classes in Ref. [GWAN12] (see also Ref. [Sve87, BCP⁺14]). Definition 27 of locally quantum behaviors induces the following definition of genuinely nonlocal quantum behaviors

Definition 28 (Of genuinely nonlocal quantum behaviors, cf. Ref. [HWD22]). *The behavior $P(\mathbf{a}|\mathbf{x})$ is genuinely nonlocal if and only if it is not locally quantum, i.e., it is not a mixture of behaviors that are a product in at least one cut and have quantum realizations.*

2.1.6 Generalized Probabilistic Theories

The framework of generalized probabilistic theories (GPTs) [Har01, Bar07] is a formalism that describes the operational predictions of essentially arbitrary conceivable physical theories. In particular, the quantum theory, the classical probabilistic theory, and the theory of non-signaling behaviors (see Sec. 2.1.6.2) are known examples of theories that can be described and systematically studied within the framework of GPTs. Each generalized probabilistic theory (GPT) is defined by some set of postulates, the implications of which can be methodically investigated within the framework of GPTs. As discussed in Sec. 2.6 here, the framework of GPTs can help in finding the future theory of Nature that would be more fundamental and have greater explanatory power than the quantum theory. See Refs. [Plá21, Mül21, Lam18] for a more insightful introduction to the framework of GPTs.

2.1.6.1 Generalized Probabilistic Theories in a Nutshell

The primitive building blocks of any GPT, \mathcal{G} are the systems, denoted $A, B, \dots \in \text{Syst}[\mathcal{G}]$, that are equipped with an associative bilinear composition rule $\otimes : \text{Syst}[\mathcal{G}] \times \text{Syst}[\mathcal{G}] \rightarrow \text{Syst}[\mathcal{G}]$ which allow constructing composite systems out of the systems. Each system A is described as a vector in finite-dimensional real vector space V_A . In this place, we note that $V_{A \otimes B}$ is not necessarily equal to $V_A \otimes V_B$. The equality holds only under the assumption of tomographic locality [Har01]. The convex, compact, and closed set $\Omega_A \subset V_A$ describes the state space for the system A . Moreover, Ω_A must have an affine dimension at least one less than the linear dimension of the vector space such that the affine span of Ω_A does not intersect with the origin. One can also define the convex state cone $K_A := \{rs | r \in \mathbb{R}^+, s \in \Omega_A\}$ which includes also subnormalised ($r < 1$), normalized ($r = 1$) and supernormalised ($r > 1$) states.

Let V_A^* be the dual vector space to V_A , i.e., the vector space of all linear functionals on V_A . A specified proper convex subset \mathcal{E}_A of V_A^* , i.e., $\mathcal{E}_A \subset V_A^*$, describes the effect space for the system A . The subset \mathcal{E}_A is assumed to be full dimensional (with respect to V_A^*), compact and closed. The effects $e \in \mathcal{E}_A$ assign probabilities to measurement outcomes whenever the measurement is performed on an arbitrary state $s \in \Omega_A$, they, therefore, must satisfy $e(s) \in [0, 1]$. The effect space \mathcal{E}_A contains also the unique unit effect $u_A \in \mathcal{E}_A$ which is defined as $u_A(s) = 1$ for all

$s \in \Omega_A$. The uniqueness of the unit effect is built into the setup of the GPTs framework (see Ref. [CDP10] for an alternative approach).

In analogy to the case of GPT systems, both state and effect spaces are equipped with an associative composition rule, $\Omega_A \otimes \Omega_B := \Omega_{A \otimes B}$, $\mathcal{E}_A \otimes \mathcal{E}_B := \mathcal{E}_{A \otimes B}$ and $u_A \otimes u_B := u_{A \otimes B}$. This composition is assumed to be bilinear, and satisfy $e \otimes f(s \otimes t) = e(s)f(t)$ for all $e \in \mathcal{E}_A$, $f \in \mathcal{E}_B$, $s \in \Omega_A$ and $t \in \Omega_B$. The conditions above allow unambiguously defining a kind of “partial trace”. Consequently, they ensure that a GPT satisfies the no-signaling principle [CDP10, Coe14, KHC17] (see also Sec. 2.1.6.2). We remark here that, in general, \otimes is not the tensor product of vector spaces. The symbol “ \otimes ” is used in analogy to the notation of the tensor product in composing quantum systems.

Given any two GPT systems A and B , there is a space of transformations \mathcal{T}_A^B from system A to system B . The space of transformations forms a closed compact and convex set. There are two associative composition rules for transformations that are relevant (i) parallel composition, $\mathcal{T}_A^B \otimes \mathcal{T}_C^D := \mathcal{T}_{A \otimes C}^{B \otimes D}$, and (ii) sequential composition, $\mathcal{T}_B^C \circ \mathcal{T}_A^B := \mathcal{T}_A^C$. Both composition rules are bilinear and must satisfy the condition that $(T_1 \otimes T_2) \circ (T_3 \otimes T_4) = (T_1 \circ T_3) \otimes (T_2 \circ T_4)$. States can be viewed as transformations with the trivial input system, denoted \star , and effect as transformations with trivial output. The trivial system \star is a unit of system compositions, i.e., $A \otimes \star = A = \star \otimes A$. Consequently, one can write that $\Omega_A = \mathcal{T}_\star^A$ and $\mathcal{E}_A = \mathcal{T}_A^\star$. Finally, there is also an identity transformation $\mathbb{1}_A$ which for every $T : A \rightarrow B$ must satisfy $\mathbb{1}_B \circ T = T = T \circ \mathbb{1}_A$.

Withing the framework of GPTs, it is convenient to work with the convention in which every GPT contains classical systems, denoted Δ_I . These systems are used as registers in which measurement outcomes can be stored, and which control variables of “experiments” can be encoded [GS18, SSC21]. Let Δ_I be a classical system which corresponds to outcome degree of freedom on some measurement device where I labels the set of possible outcomes. Then, vector space V_{Δ_I} corresponds to the real vector space \mathbb{R}^I of real-valued functions from $I \rightarrow \mathbb{R}$. Consequently, the state space Ω_{Δ_I} is the space of probability distributions over I , i.e., real-valued functions $p : I \rightarrow \mathbb{R}$ such that $p(i) \in [0, 1]$ and $\sum_i p(i) = 1$ for all $i \in I$. In fact, geometrically Ω_{Δ_I} is a simplex in which the elements of I label the vertices. Furthermore, the vertices correspond to delta function probability distributions δ_i . The effect space \mathcal{E}_{Δ_I} is the vector space dual to V_{Δ_I} . However, due to the Riesz representation theorem via the inner product $\sum_{i \in I} f(i)g(i)$, the effects can be seen as the elements of \mathbb{R}^I . In this representation, the effects correspond to functions $e : I \rightarrow \mathbb{R}$ such that $e(i) \in [0, 1]$. Geometrically, the set of such functions forms a hypercube that contains the simplex of states. In Ref. [WDS⁺18], we denote the vertices of the simplex (when interpreted as effects via the Riesz representation theorem) with ϵ_i , then $\epsilon_i(\delta_j) = \sum_{k \in I} \delta_i(k)\delta_j(k) = \delta_{ij}$. The classical systems compose, therefore, via $\Delta_I \otimes \Delta_J := \Delta_{I \times J}$. In consequence, the transformations between classical systems are stochastic linear maps. Moreover, an important property of classical theory is that identity transformations $\mathbb{1}_{\Delta_I}$ can be expanded as $\sum_{i \in I} \delta_i \circ \epsilon_i$. Consequently, any measurement $M : A \rightarrow \Delta_I$ must satisfy

$M = \mathbb{1}_{\Delta_I} \circ M = \sum_{i \in I} \delta_i \circ \epsilon_i \circ M$. Now, because, $e_i := \epsilon_i \circ M \in \mathcal{T}_A^* = \mathcal{E}_A$ one can reexpress the above $\mathbb{1}_{\Delta_I}$ as $\sum_{i \in I} \delta_i \circ e_i$. Therefore, it is possible to construct an isomorphism between (i) measurements as a collection of effects and (ii) measurements as transformations to a classical system.

2.1.6.2 The Theory of Non-Signaling Behaviors

The theory of non-signaling behaviors (aka the Box-world theory) is an instance of a generalized probabilistic theory (GPT) based on the no-signaling principle wherein the states are multipartite conditional probability distributions $P_{\mathbf{A}|\mathbf{X}}$ called behaviors, which satisfy the so-called no-signaling conditions [Bar07, PR94]. Here, $\mathbf{X} \equiv X_1 X_2 \dots X_N$ and $\mathbf{A} \equiv A_1 A_2 \dots A_N$ refer to the inputs (measurement choices) and outputs (measurement outcomes) of N parties, respectively. Importantly, not all non-signaling behaviors have quantum realizations. The theory of non-signaling behaviors exhibits stronger correlations between subsystems than quantum theory and saturates the algebraic maximum of the Clauser-Horne-Shimony-Holt (CHSH) inequality [PR94]. The no-signaling principle that prohibits instantaneous communication (faster-than-light communication, colloquially) between spatially separated subsystems yields the no-signaling conditions on the behaviors $P_{\mathbf{A}|\mathbf{X}}$. In what follows, we refer to the systems described with the theory of non-signaling behaviors as to the non-signaling devices (see also Sec. 2.5). In the bipartite setting, the no-signaling conditions and non-signaling behaviors are defined as follows

Definition 29 (Of bipartite non-signaling behaviors, cf. Ref. [Bar07, BCP⁺14]). *Let $P_{AB|XY}$ be a behavior describing a bipartite device. If the following so-called no-signaling conditions hold*

$$(2.37) \quad \forall_{a,x,y,y'} \quad P_{A|X}(a|x) = \sum_b P_{AB|XY}(ab|xy) = \sum_b P_{AB|XY}(ab|xy'),$$

$$(2.38) \quad \forall_{b,x,x',y} \quad P_{B|Y}(b|y) = \sum_a P_{AB|XY}(ab|xy) = \sum_a P_{AB|XY}(ab|x'y),$$

we call $P_{AB|XY}$ a non-signaling behavior, and $P_{A|X}(a|x)$, $P_{B|Y}(b|y)$ the marginal distributions of A and B respectively.

The Definition 29 of the no-signaling conditions and non-signaling behaviors can be readily extended to the multipartite setting.

Definition 30 (Of N -partite non-signaling behaviors, cf. Ref. [HRW10, HGJ⁺15]). *Let $P_{\mathbf{A}|\mathbf{X}}$ be a behavior describing a N -partite device, where $\mathbf{A} \equiv A_1 A_2 \dots A_N$ and $\mathbf{X} \equiv X_1 X_2 \dots X_N$ are the outputs and inputs of N parties, respectively. If the following, condition hold*

$$(2.39) \quad \forall_{1 \leq i \leq N, \mathbf{a}^{\neq i}, \mathbf{x}^{\neq i}} \forall_{x_i, X'_i} P_{\mathbf{A}^{\neq i}|\mathbf{X}^{\neq i}}(\mathbf{a}^{\neq i}|\mathbf{x}^{\neq i}) = \sum_{a_i} P_{\mathbf{A}|\mathbf{X}}(\mathbf{a}|\mathbf{x}^{\neq i}, x_i) = \sum_{a_i} P_{\mathbf{A}|\mathbf{X}}(\mathbf{a}|\mathbf{x}^{\neq i}, x'_i),$$

where $\mathbf{a}^{\neq i} \equiv a_1 \dots a_{i-1} a_{i+1} \dots a_N$ and *mutatis mutandis*, for $\mathbf{A}^{\neq i}$, $\mathbf{x}^{\neq i}$ and $X^{\neq i}$. We call $P_{\mathbf{A}|\mathbf{X}}$ a N -partite non-signaling behavior, and $P_{\mathbf{A}^{\neq i}|\mathbf{X}^{\neq i}}(\mathbf{a}^{\neq i}|\mathbf{x}^{\neq i})$ the marginal distribution of $\mathbf{A}^{\neq i}$ subsystem.

The no-signaling condition in Definition 30 above implies that the no-signaling principle is satisfied between any two subsets of parties sharing a non-signaling device [HRW10].

The state spaces in the theory of non-signaling behaviors are non-signaling polytopes [Bar07]. A polytope is a multidimensional (to dimensions greater than 3) generalization of the notion of polyhedron. A convex polytope is a convex hull of a finite number of points in some real vector space. The set of all N -partite non-signaling behaviors $\{P_{\mathbf{A}|\mathbf{X}}\}$ with m_i being the number of inputs of i -th party, and v_{ij} being the number of outputs for the j -th input, is a proper subspace of \mathbb{R}^t , where $t = \prod_{i=1}^N \sum_{j=1}^{m_i} v_{ij}$ is the total number of outputs in the behavior [Pir05]. Generally, the polytope of N -partite non-signaling behaviors is the set of all behaviors that satisfy the no-signaling conditions in Definition 30. For the sake of Sec. 2.6 we provide a more insightful definition of non-signaling polytopes.

Definition 31 (Of non-signaling polytope, cf. Ref. [Pir05, ELS06, WDS⁺18]). *Let $x \equiv (p_{ijk})_{i=1, j=1, k=1}^{i=N, j=m_i, k=v_{ij}}$ be a vector in a real vector space \mathbb{R}^t corresponding to the outputs of some N -partite behavior $P_{\mathbf{A}|\mathbf{X}}$, where m_i is the number of inputs of i -th party, and v_{ij} is the number of outputs for the j -th input, and $t = \prod_{i=1}^N \sum_{j=1}^{m_i} v_{ij}$. The non-signaling polytope \mathcal{B} is then a convex region within \mathbb{R}^t , i.e.,*

$$(2.40) \quad \mathcal{B} = \left\{ x \in \mathbb{R}^t : Ax \leq b \right\},$$

where A is an $h \times t$ matrix of reals and let $b \in \mathbb{R}^h$ be a real vector, if and only if A and b encode (i) probabilistic constraints, i.e., $0 \leq p_{ijk} \leq 1$, (ii) normalization constraints, i.e., $\sum_k^{v_{ij}} p_{ijk} = 1$, (iii) non-signaling constraints (see Definition 30), that all together form h linearly independent constraints. The dimension $\dim \mathcal{B}$ of the polytope \mathcal{B} is given by [Pir05]

$$(2.41) \quad \dim \mathcal{B} = \prod_{i=1}^n \left(\sum_{j=1}^{m_i} (v_{ij} - 1) + 1 \right) - 1.$$

As it was noted in Ref. [PR94] (see also Refs. [Ras85]) the quantum theory is nonlocal but not maximally. Namely, there exist non-signaling behaviors that exhibits stronger violations of CHSH inequalities than allowed by quantum theory [Cir80].

Definition 32 (Of Popescu-Rohrlich (PR) boxes, see Ref. [PR94]). *The Popescu-Rohrlich (PR) box or PR behavior, and anti-PR behavior $\overline{\text{PR}}$ are bipartite conditional probability distributions defined as follows*

$$(2.42) \quad \text{PR}_{\text{AB|XY}}(ab|xy) = \begin{cases} 1/2 & \text{if } a \oplus b = xy \\ 0 & \text{otherwise,} \end{cases},$$

$$(2.43) \quad \overline{\text{PR}}_{\text{AB|XY}}(ab|xy) = \begin{cases} 1/2 & \text{if } a \oplus b = xy \oplus 1 \\ 0 & \text{otherwise,} \end{cases}$$

Alternatively, in form of input-output probability tables, the PR and $\overline{\text{PR}}$ are explicitly expressed as follows

$$(2.44) \quad \text{PR}_{\text{AB|XY}}(ab|xy) = \begin{array}{c|cc|cc} & \begin{array}{c} x \\ y \end{array} \begin{array}{c} a \\ b \end{array} & 0 & 1 & \begin{array}{c} 0 \\ 1 \end{array} & \begin{array}{c} 1 \\ 0 \end{array} \\ \hline 0 & 0 & \frac{1}{2} & 0 & \frac{1}{2} & 0 \\ & 1 & 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ \hline 1 & 0 & \frac{1}{2} & 0 & 0 & \frac{1}{2} \\ & 1 & 0 & \frac{1}{2} & \frac{1}{2} & 0 \end{array},$$

$$(2.45) \quad \overline{\text{PR}}_{\text{AB|XY}}(ab|xy) = \begin{array}{c|cc|cc} & \begin{array}{c} x \\ y \end{array} \begin{array}{c} a \\ b \end{array} & 0 & 1 & \begin{array}{c} 0 \\ 1 \end{array} & \begin{array}{c} 1 \\ 0 \end{array} \\ \hline 0 & 0 & 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ & 1 & \frac{1}{2} & 0 & \frac{1}{2} & 0 \\ \hline 1 & 0 & 0 & \frac{1}{2} & \frac{1}{2} & 0 \\ & 1 & \frac{1}{2} & 0 & 0 & \frac{1}{2} \end{array}.$$

The PR and anti-PR boxes in Definition 32 are extreme points in the polytope of bipartite devices with two binary inputs and two binary outputs (for each input) behaviors, i.e., (2,2,2,2) behaviors. Furthermore, they reach the algebraic maximum for the CHSH inequality and consequently do not have quantum realizations [Cir80]. In Ref. [WDH22], we refer to the convex mixtures of the PR and anti-PR behaviors as to isotropic behaviors.

Definition 33 (Of isotropic non-signaling behaviors, cf. Ref. [BCP⁺14, WDH22]). *The isotropic behaviors $P_{\text{iso}}(\varepsilon)$ are defined as probabilistic mixtures, i.e., convex combinations, of PR and $\overline{\text{PR}}$ behaviors*

$$(2.46) \quad P_{\text{iso}}(\varepsilon) := (1 - \varepsilon)\text{PR} + \varepsilon\overline{\text{PR}},$$

for any $\varepsilon \in [0, 1]$.

Finally, we remark that the $(2,2,2,2)$ polytope P_{HRW} of bipartite binary input-output devices [HRW10, Hän10] studied by us in Ref. [WDH22] (see also Sec. 2.6) consists of 24 extremal behaviors [BLM⁺05]. Among extremal behaviors, 16 are local or deterministic behaviors, and the remaining 8 are nonlocal. The local behaviors are given by

$$(2.47) \quad L_{\alpha\beta\gamma\sigma}(ab|xy) = \begin{cases} 1 & \text{if } a = \alpha x \oplus \beta, b = \gamma y \oplus \sigma \\ 0 & \text{otherwise.} \end{cases}$$

where $\alpha, \beta, \gamma, \sigma \in \{0, 1\}$. And the nonlocal devices are

$$(2.48) \quad B_{rst}(ab|xy) = \begin{cases} 1/2 & \text{if } a \oplus b = xy \oplus rx \oplus sy \oplus t \\ 0 & \text{otherwise,} \end{cases}$$

where $r, s, t \in \{0, 1\}$.

2.2 Hybrid Quantum Network Design Against Unauthorized Secret-Key Generation, and Its Memory Cost [A]

In the first article [SWRH20], we notice a possible vulnerability of the future Quantum Internet designs [DBCZ99, MLK⁺16, ZPD⁺18, WEH18] and propose a countermeasure to it. In this way, we first design a rerouting attack on the classical parts (classical computers) of the devices that would constitute the Quantum Internet and discuss the possible consequences for the Quantum Internet service provider [SNS⁺21]. The introduced countermeasure is based on the hybrid quantum network design, which employs the fact that quantum security is not always transitive, i.e., there exist so-called nonrepeatable secure states [BCHW15, CF17]. Namely, our secure network scheme employs quantum states that possess directly accessible secret key but small amount of repeatable key, i.e., specific kinds of private and positive partial transpose (PPT) states (see Sec. 2.1.1). We refer to the difference between the secret and repeatable keys of a state as to the gap of the scheme for which we find lower bounds. Moreover, we show an example of a quantum state for which the gap of the scheme is strictly larger than zero. Application of the countermeasure brings an expense in terms of the required amount of quantum memory, i.e., the memory overhead. Consequently, we study the performance of our solution quantitatively in terms of lower bound on the amount of quantum memory needed to apply the proposed scheme.

In parallel to the advances in quantum computing that is about to enter the so-called NISQ (noisy intermediate scale quantum) era [Pre18], a great effort has been put into the development of the so-called Quantum Internet [DBCZ99, MLK⁺16, ZPD⁺18, WEH18] in which qubits rather than the classical bits would be exchanged between NISQ processing units. The main advantage of the future Quantum Internet would be its inherent security of the sent signals guaranteed by the laws of physics [Wie83, BB84]. In the first generation of the Quantum Internet [MLK⁺16], the security would be based on the quantum correlations called entanglement and their beneficial property of transitivity. Namely, two disconnected nodes of the network can obtain a mutual, unconditionally secure connection if only they share maximally entangled states with a common node. This outcome is achieved via performing the so-called entanglement swapping protocol [ŻZHE93, BBP⁺96], which allows for the initially disconnected nodes to share the maximally entangled state after the protocol is executed. On the other hand, the monogamy of quantum correlations reflected in the no-cloning theorem [WZ82] guarantees the secrecy of the newly established link. By these means, an extensive network of nodes that can perform the entanglement swapping protocol (quantum repeaters [DBCZ99], see also Sec. 2.1.2.2) connecting the end-nodes (end-users) is about to establish the first generation of the Quantum Internet in the near future [WEH18].

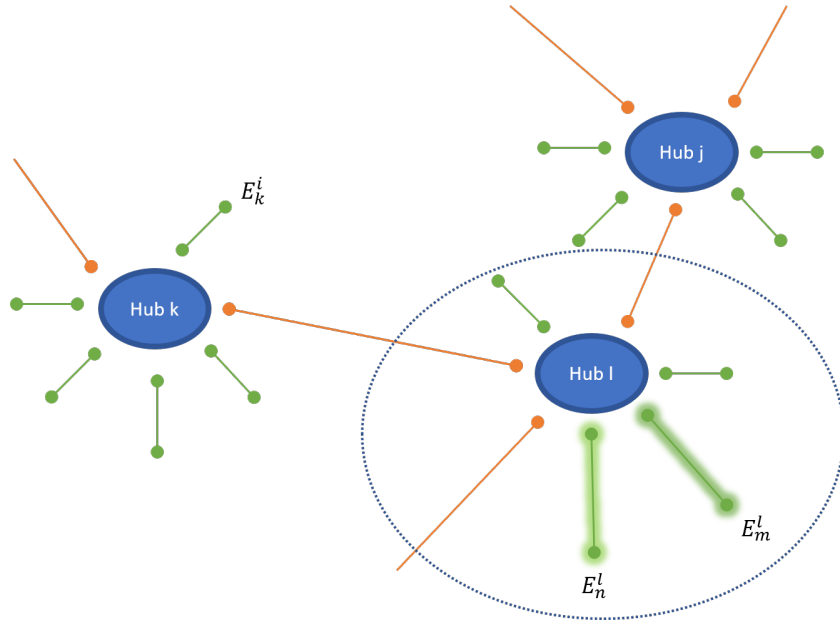


Figure 2.1: Design of the hybrid quantum network [SWRH20]. Thin green lines represent the connections between end-users and the hubs. In these, only classical data can be transferred. Thick orange lines represent the connection between the hubs being the routing nodes. In these lines also, quantum states can be transmitted. Shaded green lines connect the hub node with two end-users communicating classical data with the former. The selected region is the star-shaped network we focus on in our considerations.

As we discuss in the article, the new quantum technology that would come in the form of the Quantum Internet would not only allow for new possibilities but would also open new threats unknown from the era of the classical Internet [SNS⁺21]. In order to exemplify the possible danger, we focus on the quantum network with a star-shaped topology, i.e., a single, centrally placed quantum repeater (the hub node) connected to multiple end-users (end nodes) E_i with $i \in [1, 2, \dots, N]$ (see Fig. 2.1). The secure connections in the network are established due to entangled states shared between the hub and each end-user. Moreover, the hub is a processing unit with a classical and possibly a quantum computer inside. The critical observation is that if the network is based on pure entanglement (e.g., maximally entangled states), then the topology of the network can be altered by performing the entanglement swapping protocol incorporating the hub and two end-users (see Fig. 2.2). The described situation realizes the so-called rerouting attack [SNS⁺21]. In this manner, we consider a situation in which two malicious end-users, call them Adam and Eve, perform a Trojan horse attack on the classical computer of the hub by installing malware (malicious software) on it. In this way, Adam and Eve take control over the functions of the central (hijacked) node of the network, and, via the entanglement swapping protocol, they change the topology of the network into a (locally) disconnected graph. Therefore, Adam and Eve gain access to unconditionally secure communication with each other illegally,

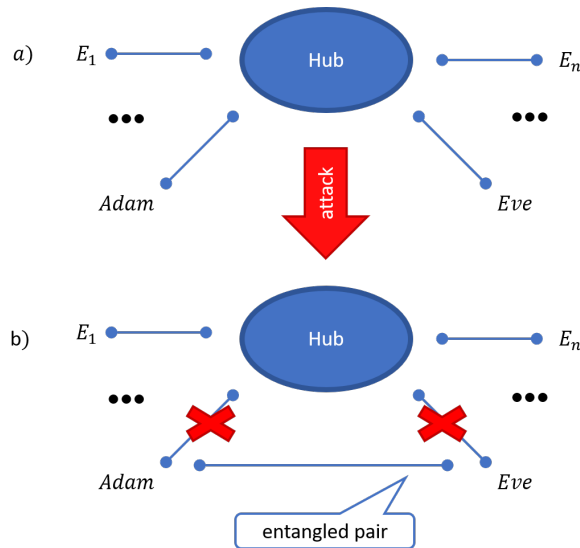


Figure 2.2: The main idea of the proposed attack [SWRH20]: (a) the hub shares pure entangled states with end users E_i , Eve and Adam, amongst others. Adam and Eve can attack the hub via malware which makes the hub perform the entanglement swapping protocol in their favor. (b) The malicious users share an entangled state after a successful attack.

i.e., without permission and at the cost of the owner of the hub. Another possibility is when the dishonest administrator of the hub sells the connections by themselves on the expense of the owner. In both situations, any physical resources needed for the functioning of the hub (e.g., energy) would be stolen. In the article, we focus on the former situation of the Trojan horse attack and study cases in which the hijacked node can perform the three-party classical communication but also when the communication is limited to one-way from the central node to the hackers.

We further propose a countermeasure to the above threat based on the recently proved no-go theorem (impossibility) [BCHW15]. Namely, *there exist quantum states that allow for point-to-point security of classical data against a quantum adversary, and in spite of this fact, they can not be effectively used in quantum key repeaters* [BCHW15]. The results in Ref. [BCHW15] show that quantum security is not always transitive, i.e., for certain states, call them nonrepeatable secure quantum states when an agent in node A has a secure link (via state ρ) with an agent in node B, and B has a secure link with an agent in node C, there is no possibility to (efficiently) create a link with non-negligible security between A and C with the help of B via three-partite local quantum operations and classical communication (3-LOCC), that is protected against B as well (see Fig. 2.3). The above feature can be realized with certain bound entangled states [HHH98] (from which pure entanglement can not be distilled via local operations and classical communication [BDSW96]). Moreover, highly noisy private states [HHHO05] also fit the above scheme in the case of 3-way and one-way classical communication

for node B to nodes A and C (with nodes A and C communicating freely) [BCHW15, CF17]. In that way, we propose a specially designed hybrid quantum network that is more robust against the aforementioned kind of attacks than the original design of a quantum network using solely quantum repeaters. The hybrid design is based on both quantum repeaters, and special relay stations called here hubs. Our approach applies in the scenario in which

- (i) the hub nodes can be connected via quantum repeaters,
- (ii) only classical data is transferred between the end nodes and the hub node,
- (iii) the distance between the end nodes and the hub node is up to the distance achievable for repeaterless quantum networks [TGW14a, PLOB17, TLWL18],
- (iv) a single instance of the attack incorporates a single hub node and its two adjacent nodes
- (v) hackers perform the “honest but curious” attack, which means that only the functioning of the classical processor is altered by the malware, while the sensitive data remain unread.

The above scheme fits the realistic use of quantum-secure Internet, in which the end-users, as in the traditional Internet, want to exchange anonymously classical data rather than quantum states, but in a quantum-secure way. In this scheme, the role of the hub nodes is to provide the registered users (end-nodes) access to the online services rather than to provide a connection between the users. The examples of the possible application of the scheme considered in the paper are online baking, data clouds, access to medical data from medical laboratories, online shopping, and possibly many other applications in which the hub node takes a role similar to the server in the traditional Internet.

As it is usually in real life, any good also comes indispensably with a price. In the above scheme that implements the proposed countermeasure, the price is the number of qubits needed to be stored and processed in the quantum memory of the nodes. Because, in the NISQ era, there is no technology to store qubits coherently for a long time, the amount of quantum memory used in the quantum network architecture is of primary importance. Therefore, in the paper, we study the lower bounds on the memory cost of the implementation of the hybrid quantum network in which every link is represented by the same quantum state ρ . Moreover, we show that it can be realized with quite modest memory requirements. We quantify the memory needed to realize the proposed scheme S_ρ in terms of the memory overhead

$$(2.49) \quad V(S_\rho) := M(\rho)(1 - \mathcal{D}(\rho)),$$

where $M(\rho)$ is the total memory of the scheme based on quantum states ρ , and $\mathcal{D}(\rho)$ is the density of the secure-key (cf. Refs. [BCHW15, HHHO09, BHH⁺14]), i.e., the ratio between the distillable key of ρ and the number of qubits (of the memory at the hub-node) needed to store the state ρ . The efficiency of the scheme is quantified by the difference between the amount of

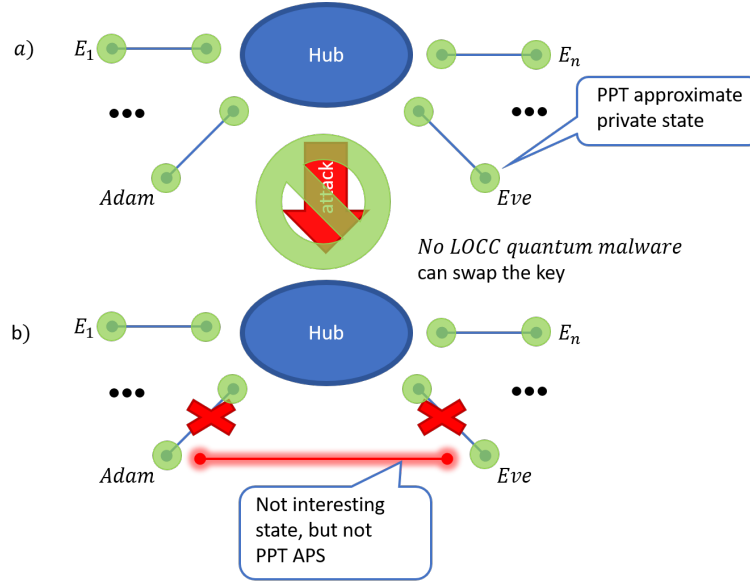


Figure 2.3: The main idea of the proposed countermeasure: (a) the connection between the end-users E_i and the hub is realized with bound entangled states (each having nearly 1 bit of secret-key at least), in particular, with Eve and Adam (shaded lines). No malware based on LOCC can efficiently swap the key. (b) Eve and Adam can not share a state with a non-negligible amount of the key (shaded red line).

the key that can be repeated R and the initial key in the link K_D . The repeatable key in our approach is the hackable one, so we want to diminish its amount while keeping the distillable key K_D possibly large. We call the difference between these two quantities ($K_D - R \geq 0$) the gap of the scheme. Furthermore, we say that a scheme is (θ, η) -good when $K_D \geq \eta$ but $R \leq \theta$ (assuming $\eta > 0$). The above notions were used for the assessment of the performance of the scheme in terms of the number of qubits (or their percentage) intended in use solely for the security of the scheme against the considered attack.

In the paper, we recognize that states especially useful for the realization of the scheme proposed by us are, among others, the so-called private states [HHHO05, HHHO09, HĆRS18]. These states are equipped with secret-key directly accessible via measurement on the key part. Moreover, private states can also possess a shielding part (shield) that is responsible for their property of having a low repeatable key [CF17] (see also Secs. 2.1.1 and 2.1.2.2). The shield protects the key but costs the quantum memory needed to implement it. Additionally, if the shielding system of a private state does not possess any secrecy content we call such a private state irreducible private state (denoted $\gamma_{\langle d_k, d_s \rangle}$, see also Sec. 2.1.1.1). Moreover, as we show in the paper, the states having positive partial transpose (PPT states, see also Sec. 2.1.1.2) that approximate private states prove their usefulness in the proposed anti-malware solution.

For the quantitative results, we firstly derive a lower bound on the overhead of the $(\theta, \log_2 d_k)$ -good secure network scheme $S_{\gamma_{d_k, d_s}}^{\rightarrow}$ employing irreducible private states γ_{d_k, d_s} , and one-way classical communication from the hub to the attacking end-nodes. The lower bound reads (Theorem 1 in Ref. [SWRH20])

$$(2.50) \quad V \left(S_{\gamma_{d_k, d_s}}^{\rightarrow} \right) \geq \Delta \log_2(d_k d_s) \left(1 - \frac{1}{2 - \frac{\theta}{\log_2 k_k}} \right) \approx_{\theta \approx 0} \frac{1}{2} M(\gamma_{d_k, d_s}),$$

where Δ is the degree of the central node (hub) equal to the number of connections between the hub and the end nodes. We observe that the requirement for the rate of the repeatable key to be approximately zero implies that at least half of the memory must be spent on the shielding system that protects the scheme. A similar result is proved for any quantum states (Theorem 2 in Ref. [SWRH20]). In the latter case, we observe a general property of secure schemes in which $\theta \approx 0$, i.e., at least half of the memory, must be spent to protect the scheme and does not store the secure-key K_D .

In the further part, we develop a lower bound on the memory overhead in the case of a secure scheme that employs strictly irreducible private states $\gamma_{\langle d_k, d_s \rangle}$ hardly distinguishable from their attacked versions $\hat{\gamma}_{\langle d_k, d_s \rangle}$ [HĆRS18] (see Sec. 2.1.1.1). The degree to which a strictly irreducible private state differs from its attacked version is quantified with the aid of the trace distance for which we assume the following

$$(2.51) \quad \left\| \gamma_{\langle d_k, d_s \rangle}^{\Gamma} - \hat{\gamma}_{\langle d_k, d_s \rangle}^{\Gamma} \right\|_1 \leq \epsilon,$$

where Γ stands for partial transpose. In this way, we show that special private states that satisfy the above, and for which conditional shield states X_{ii} are positive partial transpose operators, i.e., $X_{ii}^{\Gamma} \geq 0$, satisfy $d_s \geq \frac{d_k - 1}{\epsilon}$ (Lemma 1 in Ref. [SWRH20]). Finally, we find a lower bound on the overhead of the (θ, η) -good scheme (with one-way classical communication) which employs strictly irreducible private states (Theorem 3 in Ref. [SWRH20]) which satisfy the condition in Eq. (2.51) with separable conditional shield states ($X_{ii} \in \text{SEP}$)

$$(2.52) \quad V \left(S_{\gamma_{\langle d_k, d_s \rangle}}^{\rightarrow} \right) \geq M \left(\gamma_{\langle d_k, d_s \rangle} \right) \left(1 - \frac{\log_2 d_k}{\log_2 d_k + \log_2 \frac{d_k - 1}{\epsilon}} \right) \approx_{\epsilon \rightarrow 0} M \left(\gamma_{\langle d_k, d_s \rangle} \right),$$

for $\theta = 2 \log_2(1 + \epsilon) \approx_{\epsilon \ll 1} \frac{2}{\ln 2} \epsilon$ and $\eta = \log_2 d_k$. Here, the upper bound θ on the repeatable key rate is due to Observation 2 in Ref. [SWRH20], and the lower bound η (saturated here) is due to the properties of the strictly irreducible private states.

We observe that the parameter ϵ appears both in the formula for the lower bound on the overhead and the upper bound on the repeater rate (see Eq. (2.52) and lines below it). Therefore, one can not diminish the repeater rate to zero as desired while keeping the memory overhead at a considerably low level. Instead, one should decide on an acceptable level of the repeater rate for which the memory overhead is still reasonable. In this way, we identify low-dimensional

examples of states for which such a tradeoff is controllable [HHHO05, DKDD⁺11]. The block matrix representation of these states reads

$$(2.53) \quad \Omega_{d_s} = \frac{1}{2} \begin{bmatrix} \frac{I}{d_s^2} & 0 & 0 & \frac{F}{d_s^2} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \frac{F}{d_s^2} & 0 & 0 & \frac{I}{d_s^2} \end{bmatrix},$$

where $F = \sum_{i,j=0}^{d_s-1} |ij\rangle\langle ji|$ is a matrix of swap quantum logic gate of dimension d_s^2 . In this way, we find the first non-trivial case in which a secure network scheme has an advantage over malicious parties. Namely, we show that already for $d_s = 3$ case of Ω_{d_s} , the gap between the secure-key and repeatable key is strictly larger than zero.

The later results concern secure network schemes that employ positive partial transpose (PPT) states ρ approximating private states and two-way classical communication. In Theorem 4 of Ref. [SWRH20], we derive a lower bound on the memory overhead in the case of PPT states approximating strictly irreducible private bit (pbits, $d_k = 2$). Without loss of generality, we assume PPT states, of dimension $d_k^2 \times d_s^2$, are of the following form $\rho = \sum_{i,j,k,l=0}^{d_k-1} |ij\rangle\langle kl| \otimes A_{ij,kl}$, where $A_{ij,kl}$ are blocks of dimension d_s^2 . We obtain even more interesting results in the case of the higher dimensions. Firstly, we show the following implication (Lemma 2 in Ref. [SWRH20])

$$(2.54) \quad \|\rho - \gamma_{d_k, d_s}\|_1 \leq \epsilon \implies d_s \geq \left(\frac{d_k - 1}{\epsilon} \right) (1 - \epsilon d_k).$$

The above implication has an important corollary, namely

$$(2.55) \quad \|\rho - \gamma_{d_k, d_s}\|_1 \geq \frac{d_k - 1}{d_s + d_k(d_k - 1)}.$$

The above inequality not only holds for any dimension but, in the case of $d_k = 2$ is tighter than the known results [BHH⁺14, DKDD⁺11]. The above findings are the basis for our further developments. In Proposition 2 of Ref. [SWRH20], we show (under some mathematical conditions) an upper bound on the two-way repeater rate for PPT states approximating strictly irreducible private dit (pdit, $d_k > 2$). Finally, we derive a lower bound on the memory overhead for a (θ, η) -good secure scheme involving two-way classical communication, and $\rho \in \text{PPT}$, such that $\|\rho - \gamma_{(d_k, d_s)}\| \leq \epsilon$ for $\frac{d_k - 1}{d_s + d_k(d_k - 1)} \leq \epsilon < \frac{1}{d_k}$, $\sum_{i \neq j} \|A_{ij,ji}^\Gamma\| \leq \epsilon$, and its conditional shield states are separable (Theorem 5 in Ref. [SWRH20])

$$(2.56) \quad V(S_\rho) \geq M(\rho) \left(1 - \frac{\epsilon}{2} - f(d_k, \epsilon) \right),$$

$$(2.57) \quad f(d_k, \epsilon) := \frac{\log d_k + (1 + \frac{\epsilon}{2})h(\frac{\frac{\epsilon}{2}}{1 + \frac{\epsilon}{2}})}{\log d_k + \log \left(\frac{d_k - 1}{\epsilon} \right) + \log(1 - \epsilon d_k)},$$

with $\eta = \log d_k - 8\epsilon \log d_k - 4h(\epsilon)$ (where $h(\cdot)$ is the binary Shannon entropy), and $\theta = 2(\sqrt{\epsilon} + \epsilon) \log \dim_H(\rho) + (1 + 2\sqrt{\epsilon} + 2\epsilon) h(\frac{\sqrt{\epsilon} + \epsilon}{\frac{1}{2} + \sqrt{\epsilon} + \epsilon})$. In Fig. 2.4, we illustrate the performance of the lower bounds in Theorem 5 in Ref. [SWRH20] (here, Eq. (2.56) and lines below it).

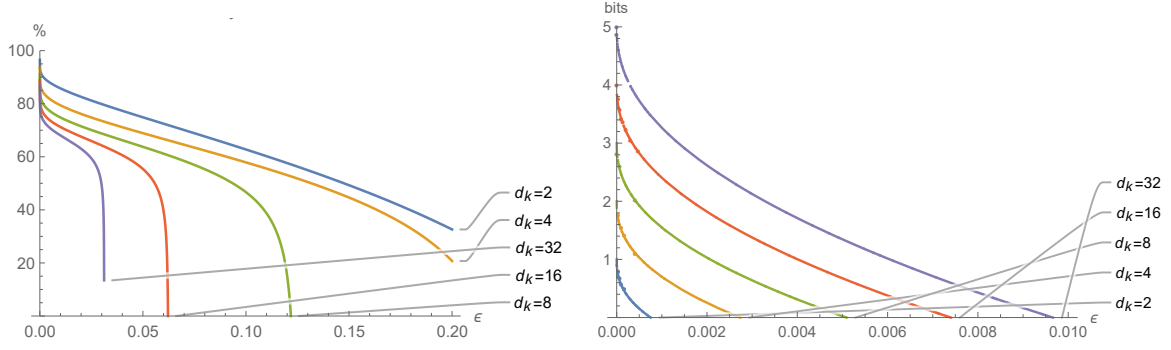


Figure 2.4: Sample results from Theorem 5. in Ref. [SWRH20]. Plots of lower bounds on the memory overhead (left) and the gap of the (θ, η) -good secure scheme (right) for $\rho \in \text{PPT}$ approximating strictly irreducible private dit (pdit) $\gamma_{\langle d_k, d_s \rangle}$, such that $\|\rho - \gamma_{\langle d_k, d_s \rangle}\|_1 \leq \epsilon$ for $\frac{d_k-1}{d_s+d_k(d_k-1)} \leq \epsilon < \frac{1}{d_k}$, $\sum_{i \neq j} \|A_{ij,ji}^\Gamma\|_1 \leq \epsilon$, for different values of the key part, i.e., d_k .

All of the above results contribute either directly or indirectly to the analysis of the efficiency of the hybrid quantum network design, as a countermeasure to the proposed rerouting attack, in terms of its memory cost. Finally, we note that we close the problem of the threat in the form of the malware attack proposed by us. Our findings fit the novel field of attacks on quantum internet [SNS⁺21] (Ref. [SWRH20] was already cited therein). We believe our results can help to design and construct the future Quantum-secure Internet that will meet the expectations of the rapidly developing information-era society.

2.3 Universal Limitations on Quantum Key Distribution over a Network [B]

In Ref. [DBWH21], we consider the distribution of secret key over a network both in bipartite and multipartite (conference) settings. Our primary achievement is establishing a unified framework capable of providing case-specific upper bounds on the achievable rates in the asymptotic and single-shot (single-use) regimes of device-dependent conference key agreement (QDD-CKA). We aim to describe a general network setting. We do this by employing the idea and framework of quantum channels. In a nutshell, any operation that transforms one quantum state into another can be considered a quantum channel (see Sec. 2.1.3 for more details), with unitary evolution or decoherence as prominent examples. In this way, we introduce the notion of a multiplex quantum channel, which links an arbitrary finite number of honest parties. In a multiplex quantum channel, each party can have the role of the sender to a channel, a receiver from a channel, or both sender and receiver. Within this scheme, we define asymptotic and single-shot (single-use) local operations and classical communication-assisted (LOCC-assisted) conference key agreement (CKA) secret capacities for multiplex quantum channels. The results of Ref. [DBWH21] consist of weak and strong converse (upper) bounds on the LOCC-assisted SKA capacities of quantum channels. To achieve our goal, we first show that any output state of a multipartite protocol distilling secret key amongst the honest parties must be genuinely multipartite entangled. The protocols we consider manifest an adaptive strategy to the secret key and entanglement distillation (in the form of private states and GHZ states, respectively) over an arbitrary multiplex quantum channel. Therefore, the structure of the protocols we consider is generic. In particular, our approach allows us to study the performance of quantum key repeaters, measurement-device-independent quantum key distribution (MDI-QKD) setups, or teleportation-covariant multiplex quantum channels. For the latter upper bounds on the SKA, capacities are given in terms of the entanglement measures of their Choi states. The bounds we provide are given in terms of the generalized relative entropies, the form of which has a topology-dependent character. Moreover, our approach allows us to obtain upper bounds on the rates at which tripartite Greenberger-Horne-Zeilinger (GHZ) states [GHZ89] can be distilled from a limited number of copies of an arbitrary multipartite quantum state. Additionally, we provide lower bounds on the SKA rate of a multiplex channel achievable in the sense of [DW05] by classical preprocessing and postprocessing (cppp) as a generalization of [PGPBL09] and show a non-trivial lower bound for a specific setup of a bidirectional quantum network.

Quantumly secure communication is one of the greatest promises of the Quantum Internet that is currently in the design phase [DBCZ99, MLK⁺16, ZPD⁺18, WEH18, DM03, Kim08, WEH18]. The formidable task of quantum communication over a network raises both fundamental and application-focused questions [BB84, Eke91, DM03, Ren05, CLL⁺09, VBD⁺15, ZXC⁺18]. The development of the technology [HBD⁺15, PDK⁺19, BRA⁺19, CZC⁺21] together

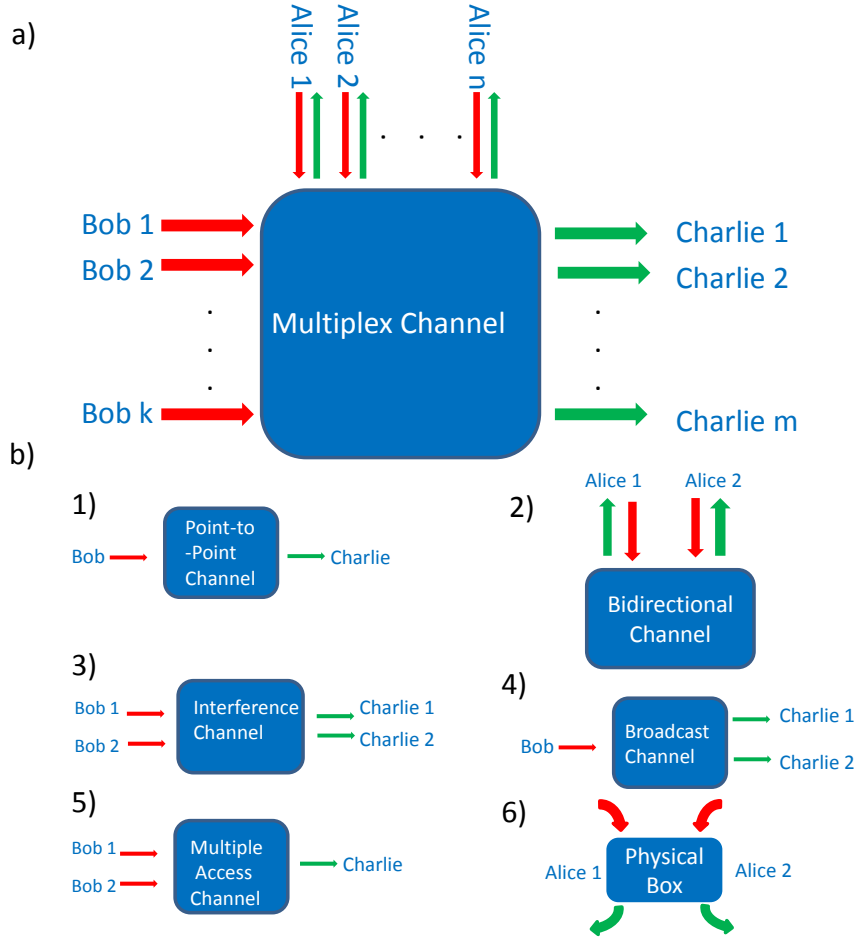


Figure 2.5: Pictorial illustration of multiplex quantum channel in Ref. [DBWH21]. The illustration depicts the universal nature of multiplex quantum channels from which all other (memoryless) quantum network channels arise. Here, red arrows represent the inputs, and green arrows represent the outputs to channels. See section IIIB of Ref. [DBWH21] for more details.

with the concerns about privacy [Sho94, ZXC⁺18] create a need for designing protocols and determining criteria for secure communication between multiple (trusted) parties in the network [CL05, AH09b]. A complex structure of a quantum network consists of various quantum channels, with possibly complex spatial alignment, due to environmental conditions. Moreover, the exponential attenuation of the signal, which can not be amplified by cloning or broadcasting, along an optical fibre [ATL15] and the difficulties in preserving entanglement for a long time due to the interaction with the environment [BRA⁺19] make the task of building a network even more complicated. On the other hand, the advancements in the technology of quantum repeaters [BDCZ98, DBCZ99, MATN15, CZC⁺21] give hope to overcome these issues. Given broad interest in the application of technologies such as quantum networks, quantum repeaters, and measurement-device-independent QKD (MDI-QKD) protocol [LCQ12, BP12] makes the

understanding of the fundamental limitations on the achievable key rates in the mentioned scenarios an important task. The seminal papers [HHHO05, CW04] on secret key distillation from states together with results in Refs. [BDSW96, VPRK97, VP98, HHH99, Dat09a] open a pathway to studies in the mentioned direction, in the case of point-to-point LOCC-assisted quantum channels [TGW14b, PLOB17, WTB17, CMH17]. Consequently, further progress has been made in the restricted network settings, for example, the case of quantum repeaters [BCHW15, CMH17], between two parties over bidirectional [DBW20, BDW18, Das18], multiple access, and interference quantum channels [LP17], and networks consisting of point-to-point [AML16, RKB⁺18, Pir19] or broadcast channels [BA17].

In Ref. [DBWH21], we focus on providing a unifying framework for obtaining upper bounds on the conference key rates achievable in the general network setting also in the one-shot scenario. To achieve this task, we introduce a multiplex quantum channel (see Fig. 2.5) as the most general form of memoryless transformation between multipartite quantum states, where each of M honest parties can have a role of the sender (Bob), receiver (Charlie) or both sender and receiver (Alice). We introduce a concise notation in which systems that are input to multiplex quantum channel of multiple Alice parties are denoted $\vec{A}' := \{A'_a\}_{a \in \mathcal{A}}$, and similarly outputs of multiple Alice parties are denoted by $\vec{A} := \{A_a\}_{a \in \mathcal{A}}$. The systems sent to multiplex quantum channel by multiple Bob parties are denoted by $\vec{B} = \{B_b\}_{b \in \mathcal{B}}$, and systems received from the channel by multiple Charlie parties are denoted $\vec{C} = \{C_c\}_{c \in \mathcal{C}}$. The reference systems kept by Alice, Bob, and Charlie parties are denoted \vec{L} , \vec{R} , and \vec{P} , respectively. Furthermore, we use $:\vec{A}:$ to denote the partition with respect to all systems in the set \vec{A} as spatially separated parties keep them. For example $:\vec{L}\vec{A}:\vec{R}\vec{B}:$ denotes systems of spatially separated Alice and Bob parties, for which system L_a are not separated from A_a , and similarly R_b are not separated from B_b . Additionally, $\vec{K} = \{K_i\}_{i=1}^M$ denotes the systems holding the secret key at the end of a protocol, and $\vec{S} = \{S_i\}_{i=1}^M$ the shielding (see Sec. 2.1.1 for the definition) systems ($\vec{S}\vec{K}$ altogether). Furthermore, we introduce secret key agreement LOCC-assisted protocols over multiplex quantum channels that provide a unifying framework for many seemingly different QKD scenarios. The unification is achieved by constructing a multiplex quantum channel in a way that its uses interleaved by LOCC simulates the protocol. Our upper bounds are based on entanglement measures called sandwiched Rényi relative entropies [WWY14, MLDS⁺13], for which relative entropy is a special case. One of the sandwiched Rényi relative entropies, i.e., the relative entropy of entanglement and its regularisation, have already proved their usefulness in providing upper bounds on the secret key rate in the bipartite case [HHHO05]. As we show, these entanglement measures are topology-dependent, as the upper bound depends on the partition of roles between the honest parties.

In Ref. [DBWH21] we consider LOCC-assisted conference key agreement protocol amongst M honest parties employing multiplex quantum channel $\mathcal{N}_{\vec{A}'\vec{B} \rightarrow \vec{A}\vec{C}}$ (see Fig. 2.6). We assume that environmental part E of the isometric extension $U_{\vec{A}'\vec{B} \rightarrow \vec{A}\vec{C}E}^{\mathcal{N}}$ of the quantum channel \mathcal{N} is

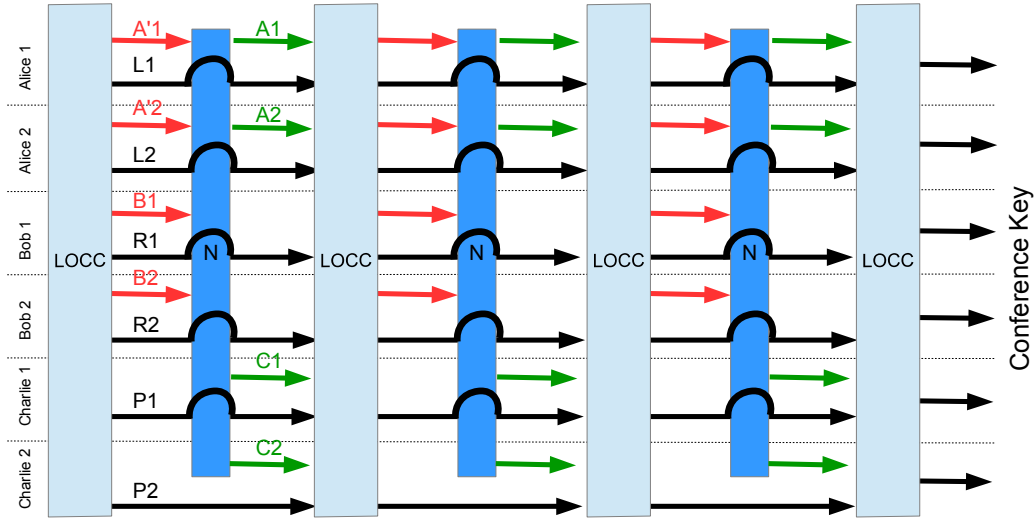


Figure 2.6: Illustration in Ref. [DBWH21] of a sample LOCC-assisted conference key agreement protocol incorporating six parties. Inputs to quantum channel \mathcal{N} are represented by red arrows, outputs are represented by green arrows, and reference systems by black arrows. Alice1 and Alice2 both send and receive quantum systems from channel \mathcal{N} , Bob1 and Bob2 only send their systems, and Charlie1 and Charlie2 only receive systems as an output of channel \mathcal{N} . In the end, all parties share a six-partite conference key.

accessible to the eavesdropper together with all classical information communicated between the honest parties while performing LOCC. All other quantum systems that are locally available to the honest parties are assumed to be secure against the eavesdropper. Fortunately, in the paradigm based on the private states, we do not need to consider the eavesdropper explicitly [HHHO05]. In a LOCC-assisted CKA-DD, the uses of a multiplex quantum channel \mathcal{N} are interleaved with the uses of LOCC channels. In the first round, the honest parties perform LOCC channel \mathcal{L}^1 to generate a fully separable state $\rho_1 \in \text{FS}(\overrightarrow{L^{(1)}}A^{(1)'} : \overrightarrow{R^{(1)}}B^{(1)} : \overrightarrow{P^{(1)}})$ that subsequently enter multiplex quantum channel $\mathcal{N}^1_{A^{(1)'}B^{(1)} \rightarrow A^{(1)}C^{(1)}}$ to yield output state $\tau_1 := \mathcal{N}^1(\rho)$. Next, the LOCC channel \mathcal{L}^2 acts on τ_1 , which enters the multiplex quantum channel again. After the final (n -th) round, the decoding \mathcal{L}^{n+1} LOCC channel generates the final state $\omega_{\overrightarrow{SK}}$, where K_i and S_i denote the key and shielding systems of the honest parties respectively. It is assumed that the eavesdropper has access to purification $\omega_{\overrightarrow{SK}Y^{n+1}E^n}$ of $\omega_{\overrightarrow{SK}}$ as it possesses all environmental systems from the isometric extension $U^{\mathcal{N}}$ and copies of classical data Y exchanged among the honest parties. The protocol for which $F(\gamma_{\overrightarrow{SK}}, \omega_{\overrightarrow{SK}}) \geq 1 - \varepsilon$ is called (n, K, ε) LOCC-assisted secret key agreement protocol (see Sec. 2.1.1.3 for more details upon multipartite private states $\gamma_{\overrightarrow{SK}}$). The rate P of the protocol is equal to the number of the conference (secret) bit per use of multiplex quantum channel, i.e., $P := \frac{1}{n} \log_2 K$ (here, $K = \dim \mathcal{H}_{K_i}$). A rate P is (weak converse) achievable if for $\varepsilon \in (0, 1)$, $\delta > 0$, and n

sufficiently large, there exists an $(n, 2^{n(P-\delta)}, \varepsilon)$ LOCC-assisted secret key agreement protocol. The supremum over all achievable rates is called the LOCC-assisted secret-key-agreement capacity $\hat{P}_{\text{LOCC}}(\mathcal{N})$ of a multiplex quantum channel \mathcal{N} . Furthermore, a rate P is called a strong converse rate for LOCC-assisted secret key agreement if for all $\varepsilon \in [0, 1), \delta > 0$, and n sufficiently large, there does not exist a $(n, 2^{n(P+\delta)}, \varepsilon)$ LOCC-assisted secret key agreement protocol. The strong converse LOCC-assisted secret-key-agreement capacity of quantum channel $\tilde{P}_{\text{LOCC}}(\mathcal{N})$ is defined as the infimum of all strong converse rates. It follows directly from the definitions that

$$(2.58) \quad \hat{P}_{\text{LOCC}}(\mathcal{N}) \leq \tilde{P}_{\text{LOCC}}(\mathcal{N}).$$

Moreover, we call a protocol in which the honest parties are allowed only for classical preprocessing and postprocessing (cPPP) communication, i.e., one use of LOCC channel for encoding and a second for decoding, a cPPP-assisted secret key agreement protocol over multiplex quantum \mathcal{N} . The $(1, K, \varepsilon)$ LOCC-assisted secret key agreement protocol is identical to $(1, K, \varepsilon)$ cPPP-assisted secret key agreement protocol. Moreover, the LOCC-assisted secret-key-agreement capacity \hat{P}_{LOCC} of the channel \mathcal{N} is always greater or equal to \hat{P}_{cPPP}

$$(2.59) \quad \hat{P}_{\text{cPPP}}(\mathcal{N}) \leq \hat{P}_{\text{LOCC}}(\mathcal{N}).$$

Finally, let $\hat{P}_{\text{cPPP}}^{\mathcal{N}}(n, \varepsilon)$ denote the maximum rate such that $(n, 2^{nP}, \varepsilon)$ cPPP-assisted secret key agreement is achievable, for any given multiplex quantum channel \mathcal{N} .

Our first main result has a technical significance and constitutes a background for further investigation. Here, we focus on the multipartite private states [AH09b] (see Sec. 2.1.1.3) that constitute the most general class of states that provide quantum conference key directly by local measurements with no further distillation required

$$(2.60) \quad \gamma_{\vec{SK}} := U_{\vec{SK}}^{\text{tw}}(\Phi_{\vec{K}}^{\text{GHZ}} \otimes \omega_{\vec{S}})(U_{\vec{SK}}^{\text{tw}})^{\dagger}.$$

Here, $\vec{K} = K_1, \dots, K_N$ denotes the key part, that is, the systems that the N honest parties have to measure in order to obtain the conference key, and $\vec{S} = S_1, \dots, S_N$ denotes the so-called shield that has to be kept secure from the eavesdropper. Furthermore, each multipartite private state $\gamma_{\vec{SK}}$ has at least $\log_2 K$ secret bits (key) [HHHO09], where $K = \dim(\mathcal{H}_i)$, for all $i \in \{1, \dots, N\}$. We prove that multipartite private states are necessarily genuine multipartite entangled. To show this we construct a multipartite privacy test (γ -privacy test), defined as a $\{\Pi^{\gamma}, \mathbb{1} - \Pi^{\gamma}\}$ such that any ε -approximate multipartite private state ρ with fidelity $F(\rho, \gamma) \geq 1 - \varepsilon$ passes the test with success probability $\text{Tr}[\Pi^{\gamma}\rho] \geq 1 - \varepsilon$. Consequently, in Theorem 1 of Ref. [DBWH21] we show that any biseparable state $\sigma_{\vec{SK}} \in \text{BS}(\vec{SK})$ can not pass any multipartite privacy test with probability greater than $1/K$, i.e.,

$$(2.61) \quad \text{Tr}[\Pi_{\vec{SK}}^{\gamma}\sigma_{\vec{SK}}] \leq \frac{1}{K}.$$

The above result is our starting point for showing how achievable rates of quantum channels capacities can be upper-bounded with various divergence measures.

We observe that interleaving uses of a multiplex quantum channel with LOCC amongst the honest parties provide the general framework to describe a number of different conference key agreement protocols. The critical point is the idea that the protocol can be simulated by a suitably constructed multiplex channel laced with LOCC operations. In this manner, we generalize the results for point-to-point [PLOB17, WTB17, CMH17] and bidirectional [Das18, DBW20, BDW18] channels. Namely, we show that secret-key-agreement capacities of multipartite generalizations of the mentioned channels are upper bounded with divergence-based measures of entangling abilities of multiplex quantum channels. The measures of entanglement and genuine entanglement (see Sec. 2.1.4 for details) we provide are of the form

$$(2.62) \quad \mathbf{E}_r(\mathcal{N}) := \sup_{\tau \in \text{FS}(\vec{L}\vec{A}':\vec{R}\vec{B}')} \mathbf{E}_r(\vec{L}\vec{A}:\vec{R}:\vec{C})_{\mathcal{N}(\tau)}.$$

Here, $r = \text{E}$ or $r = \text{GE}$ (E and GE denote entanglement and genuine entanglement, respectively), and FS denotes the set of fully separable states (see Sec. 2.1.4 for more details). Furthermore, we define the divergence \mathbf{E}_r form the convex sets \mathbf{S}_E and \mathbf{S}_{GE} of separable and biseparable states respectively for any partition $:\vec{X}:$, measured by some generalized divergence \mathbf{D}

$$(2.63) \quad \mathbf{E}_r(\vec{X})_{\rho} := \inf_{\sigma \in \mathbf{S}_r(\vec{X})} \mathbf{D}(\rho \parallel \sigma).$$

We call a quantity a generalized divergence if it satisfies the data processing inequality. Namely, for any channel \mathcal{N} the following inequality must hold

$$(2.64) \quad \mathbf{D}(\rho \parallel \sigma) \geq \mathbf{D}(\mathcal{N}(\rho) \parallel \mathcal{N}(\sigma)).$$

The examples of generalized divergence are therefore quantum relative entropy $D(\rho \parallel \sigma)$ [Ume62], the max-relative entropy $D_{\max}(\rho \parallel \sigma)$ [MLDS⁺13, Dat09b, Dat09a], the sandwiched Rényi relative entropy $\tilde{D}_{\alpha}(\rho \parallel \sigma)$ [MLDS⁺13, WWY14], the ε -hypothesis-testing divergence $D_{\tilde{h}}^{\varepsilon}(\rho \parallel \sigma)$ [BD10, WR12], trace distance $\|\rho - \sigma\|_1$ and negative of fidelity of quantum states $-F(\rho, \sigma)$.

The main results of Ref. [DBWH21] are upper bounds on the secret-key-agreement capacities of multiplex quantum channels. The upper bounds on these capacities are, therefore, upper bounds on the maximum rates at which multipartite private states can be distributed amongst the honest parties by using the multiplex quantum channel as well as free operations of classical preprocessing and postprocessing (cppp). In the one-shot (single use) case of multiplex quantum channel \mathcal{N} use, for any fixed $\varepsilon \in (0, 1)$, in Theorem 2 of Ref. [DBWH21], we have the following weak converse bound on the achievable region of cppp-assisted secret-key-agreement capacity $\hat{P}_{\text{cppp}}^{(1,\varepsilon)}(\mathcal{N})$ of \mathcal{N}

$$(2.65) \quad \hat{P}_{\text{cppp}}^{(1,\varepsilon)}(\mathcal{N}) \leq E_{h,\text{GE}}^{\varepsilon}(\mathcal{N}).$$

Here, $E_{h,\text{GE}}^\varepsilon(\mathcal{N})$ denotes the ε -hypothesis testing relative entropy of genuine multipartite entanglement for the multiplex quantum channel. $E_{h,\text{GE}}^\varepsilon(\mathcal{N})$ is based on the ε -hypothesis testing divergence [BD10], see Eqs. (2.62) and (2.63). Further, in the case of multiple uses of multiplex quantum channel interleaved with LOCC operations in Corollary 3 of Ref. [DBWH21] (see also Theorem 3 therein), we show the following strong converse bound on asymptotic capacity $P_{\text{LOCC}}(\mathcal{N})$ of multiplex quantum channel \mathcal{N}

$$(2.66) \quad \tilde{P}_{\text{LOCC}}(\mathcal{N}) \leq E_{\max,E}(\mathcal{N}).$$

Here, $E_{\max,E}(\mathcal{N})$ denotes the max-relative entropy of entanglement of the multiplex channel \mathcal{N} . $E_{\max,E}(\mathcal{N})$ is based on the max-relative entropy [Dat09a], see Eqs. (2.62) and (2.63). Moreover, in the case of finite-dimensional Hilbert spaces (associated with subsystems of honest parties) in Theorem 4 of Ref. [DBWH21], we show another strong converse upper bound in terms of regularized relative entropy of entanglement

$$(2.67) \quad \tilde{P}_{\text{LOCC}}(\mathcal{N}) \leq E_E^\infty(\mathcal{N}).$$

Additionally, we observe that if $\mathcal{N}_{\vec{A}\vec{B}\rightarrow\vec{A}\vec{C}}$ is teleportation-simulable [TGW14b, PLOB17, TSW16, WTB17, LP17, PBL⁺18], the upper bounds on $P_{\text{LOCC}}(\mathcal{N})$ are reduced to the relative entropy of entanglement of the resource state $\theta_{\vec{L}\vec{A}\vec{R}\vec{C}}$ (see Theorems 5-7 and Corollary 5 in Ref. [DBWH21]).

$$(2.68) \quad \hat{P}_{\text{LOCC}}(\mathcal{N}) \leq E_{GE}^\infty(:\vec{L}\vec{A}:\vec{R}:\vec{C}:)_\theta,$$

$$(2.69) \quad \tilde{P}_{\text{LOCC}}(\mathcal{N}) \leq E_E(:\vec{L}\vec{A}:\vec{R}:\vec{C}:)_\theta,$$

where E_{GE}^∞ is the regularized relative entropy of genuine entanglement and E_E is relative entropy of entanglement. Finally, the derived upper bounds on the secret-key agreement capacities are also upper bounds on channel capacities, where the goal is to distill GHZ states, because GHZ state is an example of a multipartite private state [AH09b].

The second main result of Ref. [DBWH21] is the application of the upper bounds described above to the specific setups. Firstly, we identify protocols like MDI-QKD [BP12, LCQ12, FYCC15, WZG⁺16, OLLP19], and quantum key repeaters [BCHW15, CMH17, CF17] are in fact special cases of LOCC-assisted SKA protocols realized with some particular multiplex quantum channel. First, we discuss that there are multiple ways in which multiplex quantum channel can describe quantum key repeater depending on the specific repeater protocol in use. In the simplest situation two honest parties Alice and Bob, send subsystems of their locally prepared maximally entangled (singlet) states Φ_{ARA}^+ and Φ_{BRB}^+ via quantum channels $\mathcal{N}_1^{A\rightarrow C_A}$, $\mathcal{N}_2^{B\rightarrow C_B}$ to Charlie who is assumed to be cooperative but not trusted. In the next step, Charlie performs a joint measurement $\mathcal{M}_{C_A C_B \rightarrow XY}$ on $C_A C_B$ subsystems, in this way realizing a step in entanglement swapping protocol [ŻZHE93], and reports the outcome to Bob who performs unitary on his reference system R_B . This procedure creates entanglement between

Alice and Bob, which can be used for cryptographic purposes. The net multiplex quantum channel describing the quantum repeater is therefore

$$(2.70) \quad \mathcal{N}_{AB \rightarrow XY}^{\text{repeater}} := \mathcal{M}_{C_A C_B \rightarrow XY} \circ \mathcal{N}_1^{A \rightarrow C_A} \otimes \mathcal{N}_2^{B \rightarrow C_B}.$$

We further discuss the possibility in which the channels of Alice and Bob are noisy, and the scheme requires to incorporate many rounds combined with error correction for entanglement distillation [BDCZ98, DBCZ99, MATN15]. The schemes which incorporate many relay stations and allow for distributing entanglement at arbitrarily large distances are also taken into account [DBCZ99, MATN15] (we refer to these as chain schemes). The upper bounds on the rates at which the key can be distributed in key repeaters setting were widely studied in the literature [BCHW15, CMH17, CF17]. Using Theorem 4 in Ref. [DBWH21] (or alternatively using results in Refs. [DBW20, BDW18, Das18]), we obtain upper bounds on setups considered in Refs. [BCHW15, CF17] that involve only one-way classical communication from Charlie to Alice and Bob. Our upper bounds are given by $\min\{E_{\max, E}(\mathcal{N}^{\text{repeater (chain)}}), E_E^\infty(\mathcal{N}^{\text{repeater (chain)}})\}$. The new bounds in Ref. [DBWH21] are capable of taking into account imperfect measurements performed by Charlie or imperfect error correction, which makes it more useful in the description of practical implementations. Furthermore, in some cases of practical interest, our upper bounds are comparable and perform better than results in Refs. [BCHW15, CF17].

Let us now describe the results in Ref. [DBWH21] regarding the MDI-QKD protocol (see Fig. 2.7). The MDI-QKD protocol is a form of QKD in which the honest parties, Alice and Bob, trust the quantum state they are supplied with but do not trust the detectors (measurements) [BP12, LCQ12]. The MDI-QKD scheme is important because it solves the problem of imperfect detectors, which can be attacked [Mak09]. For this reason, MDI-QKD protocol has drawn enormous theoretical and experimental attention over the last few years [POS⁺15, FYCC15, LYDS18, MZZ18, TLWL18, LL18, CYW⁺19, CAL19, LWW⁺19, MPR⁺19, PLG⁺19, XMZ⁺20]. In the typical MDI-QKD setup [BP12, LCQ12], the honest parties prepare quantum states, which are sent to the relay station via quantum channels $\mathcal{N}_{A' \rightarrow A}^1$ and $\mathcal{N}_{B' \rightarrow B}^2$. The relay station where a joint measurement $\mathcal{M}_{AB \rightarrow x}$ is performed might be in control of eavesdropping Eve that communicates the outcome of the measurement to Alice and Bob via classical broadcast channel $\mathcal{B}_{X \rightarrow Z_A Z_B}$.

$$(2.71) \quad \mathcal{N}_{A'B' \rightarrow Z_A Z_B}^{\text{MDI}} := \mathcal{B}_{X \rightarrow Z_A Z_B} \circ \mathcal{M}_{AB \rightarrow X} \circ \mathcal{N}_{A' \rightarrow A}^1 \otimes \mathcal{N}_{B' \rightarrow B}^2,$$

where Z_A and Z_B are classical registers of Alice and Bob, respectively. In the virtue of Theorem 3 and Theorem 4 of Ref. [DBWH21] (alternatively, using results in Ref. [DBW20, BDW18, Das18]), we obtain an upper bound on the achievable key rate in terms of $E_{\max, E}(\mathcal{N}_{A'B' \rightarrow Z_A Z_B}^{\text{MDI}})$ and $E_E^\infty(\mathcal{N}_{A'B' \rightarrow Z_A Z_B}^{\text{MDI}})$ in case of general systems and finite-dimensional systems respectively. Moreover, we discuss in detail a photon-based prototype of MDI-QKD protocol in the form of the dual-rail scheme in which qubits are encoded in two orthogonal modes of a single

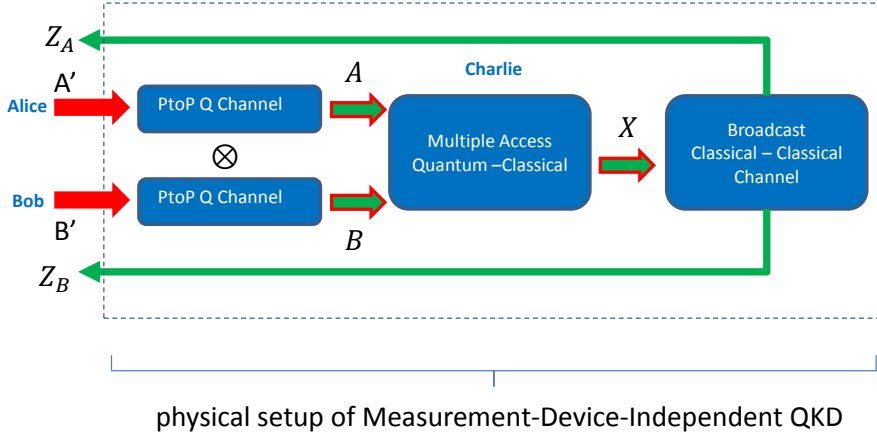


Figure 2.7: Graphical illustration of quantum classical multiplex channel $\mathcal{N}_{A'B' \rightarrow Z_A Z_B}^{\text{MDI}}$ in Ref. [DBWH21]. The channel $\mathcal{N}_{A'B' \rightarrow Z_A Z_B}^{\text{MDI}}$ is a composition of three types of elementary multiplex channels, i.e., a pair of point-to-point channels from Alice to Charlie and from Bob to Charlie composed with quantum measurement (multiple access quantum to classical channel) in the hands of Charlie followed by classical broadcast channel back to Alice and Bob. The red arrows represent the inputs to the channel, and the green arrows represent the outputs. The green arrows with red boundaries are the outputs of the multiplex channel which are also inputs to the consecutive channels.

photon [RP10]. For the noise model present in the quantum channel, we chose the pure-loss bosonic channel with transmissivity η [DKD18]. We observe that in the considered case, the action of the pure-loss bosonic channel is identical to the erasure channel [GBP97], and therefore the former channel is tele-covariant as the latter one is. We further assume that the central relay station operated by Charlie (or Eve) performs a perfect measurement in the Bell's basis with probability q and fails with probability $1 - q$ of what the honest parties are informed. Finally, using Theorem 7 in Ref. [DBWH21] we obtain the following capacity for the considered bipartite prototype of MDI-QKD protocol $\mathcal{N}^{\text{MDI}, \{\mathcal{E}_i\}_{i=1}^2}$

$$(2.72) \quad \tilde{P}_{\text{LOCC}}(\mathcal{N}^{\text{MDI}, \{\mathcal{E}_i\}_{i=1}^2}) = q\eta_1\eta_2.$$

The equality in Eq. (2.72) above is because $q\eta_1\eta_2$ is known to be an achievable rate for the given setup (what follows from results in Refs. [GEW16, PLOB17, WTB17]). The direct application of results from Refs. [GEW16, PLOB17] gives a repeaterless upper bound (RB) on the MDI-QKD capacity to be $\min\{\eta_1, \eta_2\}$ [CMH17, Pir19] what is never less than $q\eta_1\eta_2$. See Fig. 2.8 for a comparison between the performance of the RB upper bound and bound derived in Ref. [DBWH21].

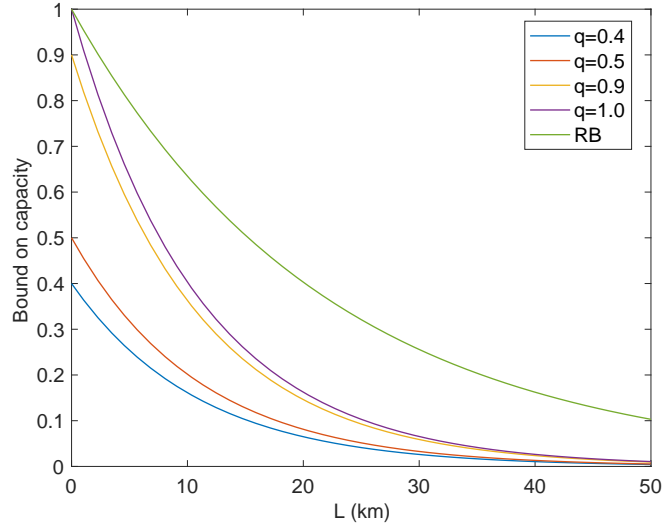


Figure 2.8: Plot in Ref. [DBWH21] depicting the rate-distance tradeoff comparison between different upper bound on MDI-QKD scheme for considered photon-based prototype. The values of parameters chosen are $\eta_1 = \eta_2 = \exp(-\alpha L)$, and $\alpha = \frac{1}{22\text{km}}$. The result of Ref. [DBWH21], here given by (2.72), are blue, red, yellow, and purple curves (plotted for different values of q), and the repeaterless bound (RB) is here green curve.

Our next main result concerns upper bounds on the rates of conference key distillation from quantum states rather than the secret-key-agreement capacity of quantum channels. First, we state the security condition for the output of (n, K, ε) LOCC-assisted conference key agreement protocol

$$(2.73) \quad F(\mathcal{L}_{A^{\otimes n} \rightarrow SK}(\rho_A^{\otimes n}), \gamma_{KS}) \geq 1 - \varepsilon,$$

where $\mathcal{L}_{A^{\otimes n} \rightarrow SK}$ is the LOCC (operation) channel performed by M honest parties. In Theorem 8 of Ref. [DBWH21] we show that the one-shot secret-key distillation rate $K_D^{(1,\varepsilon)}$ from a single copy of multipartite quantum state is upper bounded by

$$(2.74) \quad K_D^{(1,\varepsilon)}(\rho) \leq E_{h,\text{GE}}^\varepsilon(\vec{A} : \rho) := \inf_{\sigma \in \text{BS}(\vec{A})} D_h^\varepsilon(\rho \| \sigma).$$

Further, in Proposition 2 of Ref. [DBWH21] we show that the secret key rate K_D [AH09b] obtained directly from the definition $K_D(\rho) = \inf_{\varepsilon > 0} \limsup_{n \rightarrow \infty} \frac{1}{n} K_D^{(n,\varepsilon)}(\rho^{\otimes n})$ is upper bounded with the regularized relative entropy of the genuine entanglement

$$(2.75) \quad K_D(\rho_{\vec{A}}) \leq E_{GE}^\infty(\rho_{\vec{A}}).$$

The above result generalizes Theorem 9 in Ref. [HHHO09]. Moreover, we observe that (see Corollary 6 in Ref. [DBWH21]) $K_D(\rho_{\vec{A}}) = 0$ provided $\rho_{\vec{A}}$ is a tensor-stable biseparable state (see

Sec. 2.1.4 for details). In this place, we discuss that already in tripartite setting ($M = 3$), there exist two nonequivalent classes of genuinely entangled states, i.e., the Φ_M^{GHZ} class and Φ_M^{W} class of states [DVC00, BPR⁺00, HHHH09, AFOV08, HSD15, SdVSK17]. Despite the fact, that both classes contain the states that are pure entangled they can not be transformed between each other at unit rate [FL07, FL08, CHL10, VC15, VC19, SME20]. Since Φ_M^{GHZ} has a role of perfect implementation of CKA protocols, the task of distillation of Φ_3^{GHZ} from Φ_3^{W} has been studied [SVW05, FL07, KT10, CCL11, VC15, VC19]. In particular, it is known that a single Φ_3^{W} state can not be transformed into Φ_3^{GHZ} state even in a probabilistic manner [SVW05]. On the other hand, a lower bound on the asymptotic conversion rate is known to be ≈ 0.643 (see Theorem 2 in Ref. [SVW05]). However, as we exemplify, the Φ_3^{W} to Φ_3^{GHZ} conversion in the one-shot regime is still possible. Namely, two Φ_3^{W} states can be transformed into a single Φ_3^{GHZ} state with probability arbitrarily close to $\frac{2}{3}$. Since, distillation of Φ_M^{GHZ} is only an instance of conference key distillation strategy [Cab00, SG01, CL05, AH09b, AH09a, GKB19, VC15, VC19], and more general scenario incorporates distillation of private states, i.e., twisted Φ_M^{GHZ} states [HLPFH08, HHHH09, AH09b, PH10, BA17] Eq. (2.74) (Theorem 8 in Ref. [DBWH21]) constitutes an upper bound on Φ_M^{GHZ} distillation rate. In this way, the lower bound of $\frac{2}{3}$ can be compared with the upper bound in Eq. (2.74). In order to obtain non-trivial upper bounds on CKA distillation rates from single or multiple copies of noiseless, dephased, and depolarized Φ_3^{GHZ} and Φ_3^{W} we devise two families of biseparable states

$$(2.76) \quad \pi_{\text{GHZ}}^{n,M} := \frac{1}{M} \sum_{i=1}^M \left(\mathcal{S}_{1,i} \left(\frac{I}{2} \otimes \Phi_{M-1}^{\text{GHZ}} \right) \right)^{\otimes n},$$

$$(2.77) \quad \pi_{\text{W}}^{n,M} := \frac{1}{M} \sum_{i=1}^M \left(\mathcal{S}_{1,i} \left(|0\rangle\langle 0| \otimes \Phi_{M-1}^{\text{W}} \right) \right)^{\otimes n},$$

here the operator $\mathcal{S}_{1,i}$ swaps the qubit of the first system (party) with the qubit of the i -th system. The number-of-copies-sensitive definition of biseparable states $\pi_{\text{GHZ}}^{n,M}$ and $\pi_{\text{W}}^{n,M}$ is due to non-closure of the set of biseparable states under the tensor product. Interestingly, we point out that the state $\pi_{\text{W}}^{1,3}$ devised by us in Ref. [DBWH21] is closer in the Hilbert-Schmidt norm to Φ_3^{W} than the state in Ref. [VDM02] that was found to be the closest (in Ref. [VDM02] alternative definition of biseparability was employed). See Fig. 2.9 for the performance of the upper bound in Eq. (2.74) (Theorem 8 in Ref. [DBWH21]) employing biseparable states in Eqs. (2.76) and (2.77). Finally, in Proposition 3 and Corollary 7 of Ref. [DBWH21], we show a method of constructing a family of (in general) nonequivalent upper bounds on the asymptotic key rate K_D . The method is based on Proposition 2 in Ref. [DBWH21] and the observation that $M - 1$ -partite key is no less than M -partite key if two parties unite. This fact is a consequence of the strict inclusion of the set of M -partite LOCC within the set of $M - 1$ -partite LOCC. In this way, we obtain

$$(2.78) \quad K_D \left(\Phi_3^{\text{W}} \right) \leq K_D \left(\Phi_{2+1}^{\text{W}} \right) \leq E_{\text{GE}}^{\infty} \left(\Phi_{2+1}^{\text{W}} \right) = h_2 \left(\frac{1}{3} \right) \approx 0.9183 \text{ bit},$$

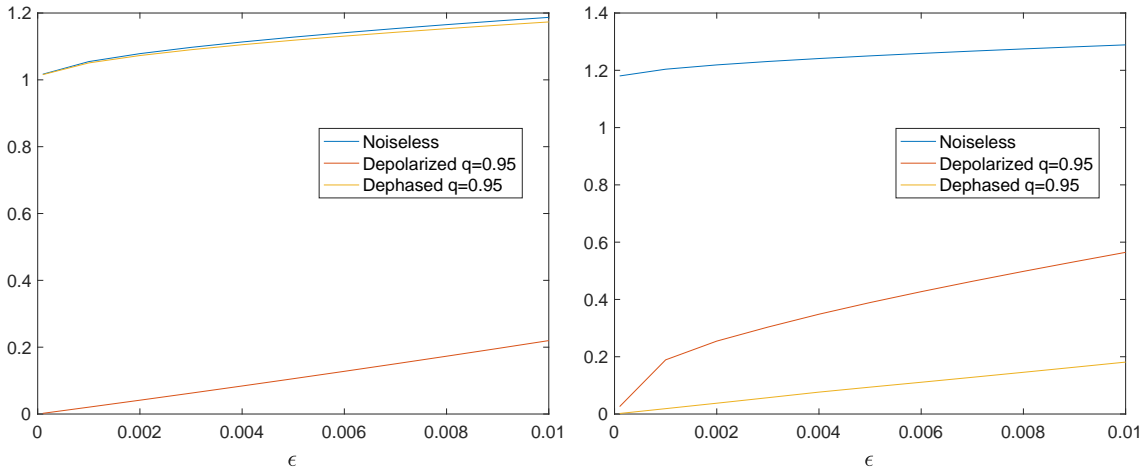


Figure 2.9: Plots in Ref. [DBWH21] depicting upper bounds in Theorem 8 (see Eq. (2.74)) on one-shot conference key distillation rates from a single copy of Φ_3^{GHZ} state (left) and two copies of Φ_3^{W} state (right). The dephasing ($\mathcal{D}_{deph.}^q(\omega) = q\omega + (1-q)\sigma_z\omega\sigma_z$, where σ_z is the Pauli Z operator) and depolarizing ($\mathcal{D}_{depol.}^q(\omega) = q\omega + (1-q)\frac{\mathbb{1}}{2}$) noise act separately on each qubit.

where $h_2(x)$ is the binary entropy function, and Φ_{2+1}^{W} denotes Φ_3^{W} state in which arbitrary two parties united, identical to one of Bell's states. The above upper bound can be compared with the lower bound of ≈ 0.643 bit obtained in Ref. [SVW05].

Additionally, as another contribution of Ref. [DBWH21], we provide lower bounds on the secret-key-agreement rates of multiplex quantum channels that can be achieved with the assistance of cppo operations. The multipartite protocols we devise are based on the Devetak-Winter protocol [DW05]. The Devetak-Winter protocol is a point-to-point protocol that incorporates one-way classical communication from Alice to Bob [DW05]. In this way, we generalize the lower bound on secret-key-agreement rates for multipartite quantum states derived in Ref. [AH09b] along with the lower bound for point-to-point quantum channels shown in Ref. [PGPBL09] to a network of bidirectional quantum channels. The first lower bound we derive is a direct extension of the result obtained for quantum states in Ref. [AH09b]. Here, the idea is to choose the so-called distributing party that performs the (directed) Devetak-Winter protocol with each of the other parties. In this case, the achievable rate is given by the worst-case Devetak-Winter protocol rate between the distributing party and any other party. As an enhancement of the lower bound, we maximize over possible choices of the distributing party. Our second protocol considers a chain of parties in which each party performs the Devetak-Winter protocol with the subsequent party. The achievable rate is given then by the lowest Devetak-Winter protocol rate between all links in the chain. In this case, we optimize with respect to all possible permutations of the parties in the chain. Furthermore, we consider a bidirectional quantum network, i.e., a network in which each node is connected to all of

its neighbors via a product of point-to-point quantum channels in opposite directions. Our technique is based on some facts from the graph theory [Wil96]. In particular, one can represent a bidirectional quantum network as a weighted, directed multigraph in which nodes represent parties, and each edge, having some weight, represents a product bidirectional channel. In order to find the best lower bound, in Ref. [DBWH21], for each spanning tree, we attribute its lowest rate of Devetak-Winter protocol [Dev05] achieved among any pair of parties connected by an edge in the tree. We further maximize this rate over all spanning trees. Finally, we exemplify that the last approach is advantageous with respect to the two described before.

To summarize, by providing universal limitations on the rates at which quantum conference key can be distributed over multiplex quantum network, we created a unified approach for the treatment of a diverse class of secure communication setups. In this way, we contributed to a better understanding of the limitations of the aforementioned class of protocols and, consequently, to a better understanding of the whole CKA-DD security paradigm. The derived upper bounds provide benchmarks on the present experimental setups and forthcoming quantum Internet of the future.

2.4 Fundamental Limitations on the Device-Independent Quantum Conference Key Agreement [C]

In Ref. [HWD22], we initiate the study of the upper bounds in the device-independent quantum conference key agreement (DI-CKA) scenario. We firstly provide a multipartite generalization of the cc-squashed entanglement introduced in Ref. [AFL21] and further developed in Ref. [KHD22]. With a little abuse of notation, we call the new measure reduced c-squashed entanglement $E_{sq,dev}^c$. We show that $E_{sq,dev}^c$ upper bounds the device-independent conference key rate $K_{DI,dev}^{iid,\hat{x}}$ achieved by (standard) protocols (see, e.g., [AFDF+18]) that use a single input \hat{x} to generate the key. This relation holds in the independent and identically distributed setting as we aim at the upper bounds on the rate of DI-CKA (see Sec. 2.1.2.1). Here, the subscript dev refers to the fact that the adversary mimics the full statistics of the honestly implemented device. We generalize further the scenario to the case in which the eavesdropper has to mimic only some relevant parameters of the device. The latter case is denoted with subscript par (e.g., $K_{DI,par}^{iid,\hat{x}}$). To achieve our goal, we generalize the upper bounds via intrinsic information studied in Ref. [CEH+07] to the case in which the system of the adversary can be infinite-dimensional. In that way, we closed a possible loophole in the proofs of Corollary 3 and Corollary 4 of Ref. [KHD22]. Our upper bounds are then compared with the lower bound on the DI-CKA secret key rate given in Ref. [RMW18] (see Fig. 2.10). We also provide a non-trivial upper bound on the device-independent key rate in the parallel measurement scenario in which all parties simultaneously set all values of inputs x_i . Moreover, we generalize reduced bipartite entanglement measures (see Ref. [KHD22]) to the multipartite case. In this way, we show that the reduced regularized relative entropy of genuine entanglement [DBWH21] upper bounds the DI-CKA rate for multipartite quantum states. We also discuss the issue of genuine nonlocality [BCP+14] and genuine entanglement (see Secs. 2.1.4 and 2.1.5) in the context of the DI-CKA [HHHH09]. Finally, we show a strict gap between the rate of the quantum device-independent conference key K_{DI} and quantum device-dependent conference key rate K_{DD} . This result is achieved by providing a method to lift the gap in bipartite case to the multipartite case. In this manner, we construct multipartite states that exhibit a strict gap using bipartite states for which the gap was proven in Ref. [CFH21].

Building a quantumly secured Internet would provide worldwide information-technologically secure communication [DM03, WEH18]. Quantum repeaters [DBCZ99, MLK+16, ZXC+18] give hope that this formidable task will be achieved in the not-too-distant future. Unfortunately, the level of quantum security proposed in the seminal paper of C. H. Bennett and G. Brassard [BB84] appears to be insufficient. This insufficiency is because, for example, a malicious eavesdropper can change the inner workings of a quantum device on the way of the device from the trusted manufacturer to the honest users, making the cryptographic device totally insecure [PAB+09]. Indeed, attacks on the quantum cryptographic devices and quantum internet are already considered [BRPB13, Mak09, SWRH20, SNS+21]. Fortunately,

the idea of device-independent cryptography overcomes this obstacle [Eke91, PAB⁺09] and importantly seems to be feasible in practise [ZvLR⁺22, NDN⁺22, LZZ⁺22]. Simultaneously, the limitations of the device-independent approach in terms of the upper bounds on the distillable key were considered [KWW20, CFH21, FBJL⁺21, KHD22]. However, present studies focus on the point-to-point device-independent secure communication. In Ref. [HWD22], we initiate the study of upper bounds in the device-independent conference key agreement (DI-CKA) [MGKB20, RMW18], where the task is to distribute the same secure key between $N > 2$ honest parties, security of which does not depend on the inner working of the devices of the honest parties. The key is used later for one-time-pad encryption. A protocol achieving the DI-CKA task has been shown in [RMW18]. Our considerations set upper bounds on the performance of any DI-CKA generating protocol, including the mentioned one, in the network setting.

In the considered setup, N trusted spatially separated allies intend to extract secret-key against a quantum adversary. Aiming at the upper bounds it is enough to consider that the honest parties share n identical devices. The honest implementation of the device, which consists of state and measurement, is denoted (ρ, \mathcal{M}) . However, the adversary can replace the honest device, provided by the manufacturer, with a different device (σ, \mathcal{N}) , such that it yields the same input-output statistics as the honest one. The (coarse-grained) statistics tested by the honest parties are usually the level of violation of some multipartite Bell inequality $\omega(\rho, \mathcal{M})$ (see Refs. [Mer90, Ard92, BK93, SS02, ŻB02, WW01, YCZ⁺12, HSD15, Luo22]), and the quantum bit error rate (QBER) $P_{err}(\rho, \mathcal{M})$. The QBER is the probability that the outputs of the honest parties' device are not equal to each other, given that the input for the key generating round has been chosen. Still, in some cases, the honest parties would like to test the whole statistics of the (ρ, \mathcal{M}) device. The honest measurement is described as a family of tensor product of POVMs for each of the honest party $\mathcal{M} \equiv \{M_{a_1}^{x_1} \otimes M_{a_2}^{x_2} \otimes \dots \otimes M_{a_N}^{x_N}\}_{\mathbf{a}|\mathbf{x}}$, where $\mathbf{x} := (x_1, x_2, \dots, x_N)$ and $\mathbf{a} := (a_1, a_2, \dots, a_N)$ correspond to the settings of inputs and outputs for some number $N \in \mathbb{N}$ of the honest parties (denoted $N(A) \equiv A_1 \dots A_N$). The set $\{a_i\}_{i=1}^N$ denotes a finite set of measurement outcomes for measurement choices x_i . In that way, the device yields a probability distribution

$$(2.79) \quad p(\mathbf{a}|\mathbf{x}) = \text{Tr}[M_{a_1}^{x_1} \otimes M_{a_2}^{x_2} \otimes \dots \otimes M_{a_N}^{x_N} \rho_{N(A)}],$$

for a measurement \mathcal{M} on a N -partite state $\rho_{N(A)}$. Provided the statistics $\{p(\mathbf{a}|\mathbf{x})\}_{\mathbf{a}|\mathbf{x}}$ obtained from devices (ρ, \mathcal{M}) and (σ, \mathcal{N}) are identical, we write $(\sigma, \mathcal{N}) = (\rho, \mathcal{M})$. Consequently, $(\sigma, \mathcal{N}) = (\rho, \mathcal{M})$ implies $\omega(\sigma, \mathcal{N}) = \omega(\rho, \mathcal{M})$ and $P_{err}(\sigma, \mathcal{N}) = P_{err}(\rho, \mathcal{M})$. Furthermore, if the statistic p and p' corresponding to the devices (σ, \mathcal{N}) and (ρ, \mathcal{M}) satisfy

$$(2.80) \quad d(p, p') = \sup_{\mathbf{x}} \|p(\cdot|\mathbf{x}) - p'(\cdot|\mathbf{x})\|_1 \leq \varepsilon,$$

we say that (σ, \mathcal{N}) and (ρ, \mathcal{M}) are ε -close to each other and denote the ε -proximity by $(\sigma, \mathcal{N}) \approx_\varepsilon (\rho, \mathcal{M})$. In that way, we consider the relations

$$(2.81) \quad (\rho, \mathcal{M}) \approx_\varepsilon (\sigma, \mathcal{N}),$$

$$(2.82) \quad \omega(\rho, \mathcal{M}) \approx_\varepsilon \omega(\sigma, \mathcal{N}),$$

$$(2.83) \quad P_{err}(\rho, \mathcal{M}) \approx_\varepsilon P_{err}(\sigma, \mathcal{M}),$$

that are used in the definitions of the DI-CKA key rate. Finally, for the sake of upper bounds, it is enough to consider the scenario that contains independent and identically distributed (iid) copies of a device. This simplification is because the iid scenario automatically constitutes an upper bound on the general scenario [CFH21]. In this, we work with two different notions of DI-CKA, iid maximal distillation rates for a device (ρ, \mathcal{M}) , i.e., dev and par key rates (see Sec. 2.1.2.1)

$$(2.84) \quad K_{DI,dev}^{iid}(\rho, \mathcal{M}) := \inf_{\varepsilon>0} \limsup_{n \rightarrow \infty} \sup_{\hat{\mathcal{P}}} \inf_{(2.81)} \kappa_n^\varepsilon \left(\hat{\mathcal{P}}((\sigma, \mathcal{N})^{\otimes n}) \right),$$

$$(2.85) \quad K_{DI,par}^{iid}(\rho, \mathcal{M}) := \inf_{\varepsilon>0} \limsup_{n \rightarrow \infty} \sup_{\hat{\mathcal{P}}} \inf_{(2.82),(2.83)} \kappa_n^\varepsilon \left(\hat{\mathcal{P}}((\sigma, \mathcal{N})^{\otimes n}) \right),$$

where $\hat{\mathcal{P}}$ is a protocol that consists of classical local operations and public (classical) communication (CLOPC), κ_n^ε is the ε -secure key rate of a protocol $\hat{\mathcal{P}}$ achieved for any fixed value of security parameter $\varepsilon > 0$, number of copies n , and measurements \mathcal{N} . Because, Eq. (2.81) implies Eq. (2.82) and Eq. (2.83) we automatically have $K_{DI,dev}^{iid}(\rho, \mathcal{M}) \geq K_{DI,par}^{iid}(\rho, \mathcal{M})$. Furthermore, we observe that DI-CKA maximal key distillation rates are upper bounded by the device-dependent distillation key rates (DD-CKA) and also by the entanglement measures that upper bound the latter.

In order to obtain our main result, we first generalize the notion of the cc-squashed entanglement [AFL21, KHD22] to the multipartite setting based on the notion of multipartite squashed entanglement [YHH⁺09]. We call the newly introduced quantity the c-squashed entanglement that reads

$$(2.86) \quad E_{sq}^c(\rho_{N(A)}, \mathbf{M}) := \inf_{\Lambda: E \rightarrow E'} I(A_1 : \dots : A_N | E')_{\mathbf{M}_{N(A)} \otimes \Lambda \psi_{N(A)E}^\rho}.$$

Here, $\mathbf{M}_{N(A)} \equiv M_{A_1}, \dots, M_{A_N}$ is an N -tuple of positive-operator-valued measures (POVMs) and state $|\psi_{N(A)E}^\rho\rangle$ is a purification of $\rho_{N(A)}$. Furthermore,

$$(2.87) \quad I(A_1 : \dots : A_N | E)_{\rho_{N(A)}} = \sum_{i=1}^N S(A_i | E)_{\rho_{N(A)}} - S(A_1, \dots, A_N | E)_{\rho_{N(A)}},$$

where $S(A_i | E)_{\rho_{N(A)}} = S(A_i E)_{\rho_{N(A)}} - S(E)_{\rho_{N(A)}}$ is the conditional entropy, with $S(AB) := -\text{Tr}[\rho_{AB} \log_2 \rho_{AB}]$ and $S(A) := -\text{Tr}[\rho_A \log_2 \rho_A]$ being von Neumann entropies. Having the new object defined, we provide the first upper bound (Theorem 1 of Ref. [HWD22])

$$(2.88) \quad K_{DD}(\mathbf{M}_{N(A)} \otimes \text{id}_E \psi_{N(A)E}^\rho) \leq \frac{1}{N-1} E_{sq}^c(\rho_{N(A)}, \mathbf{M}_{N(A)}).$$

Here, K_{DD} is the quantum device-dependent conference key agreement (DD-CKA) rate, and id_E is the identity channel (quantum operation). The above upper bound is a multipartite generalization (and improvement) of Theorem 5 in Ref. [KHD22]. The improvement is because in Ref. [HWD22], we additionally closed the problem of an eavesdropper holding an infinite-dimensional system. This is achieved directly by allowing the dimension of Eve's system in Eq. (2.88) to be infinite dimensional which is natural in the case of nonlocality. We then move to our main goal, i.e., to study the upper bounds in the DI-CKA scenario. In this case we concentrate on the standard protocols [AFDF⁺18] which use a single tuple of measurements $(\hat{x}_1, \dots, \hat{x}_N) \equiv \hat{\mathbf{x}}$ applied to \mathcal{M} of a device $(\rho_{N(A)}, \mathcal{M})$ yielding $\mathcal{M}(\hat{\mathbf{x}})$ (Theorem 2 and Theorem 3 in Ref. [HWD22])

(2.89)

$$K_{DI,dev}^{iid,\hat{\mathbf{x}}}(\rho_{N(A)}, \mathcal{M}) \leq E_{sq,dev}^c(\rho_{N(A)}, \mathcal{M}(\hat{\mathbf{x}})) := \frac{1}{N-1} \inf_{(\sigma_{N(A)}, \mathcal{L}) \equiv (\rho_{N(A)}, \mathcal{M})} E_{sq}^c(\sigma_{N(A)}, \mathcal{L}(\hat{\mathbf{x}})),$$

(2.90)

$$K_{DI,par}^{iid,\hat{\mathbf{x}}}(\rho_{N(A)}, \mathcal{M}) \leq E_{sq,par}^c(\rho_{N(A)}, \mathcal{M}(\hat{\mathbf{x}})) := \frac{1}{N-1} \inf_{\substack{\omega(\sigma_{N(A)}, \mathcal{L}) = \omega(\rho_{N(A)}, \mathcal{M}) \\ P_{err}(\sigma_{N(A)}, \mathcal{L}) = P_{err}(\rho_{N(A)}, \mathcal{M})}} E_{sq}^c(\sigma_{N(A)}, \mathcal{L}(\hat{\mathbf{x}})).$$

We refer to the quantities in the right hand sides of the Eqns. (2.89) and (2.90) above, as to the reduced c-squashed entanglement, “dev” and “par” respectively. We then observe that in the case $N = 2$, the upper bound in Eq. (2.90) recovers the results from Ref. [KHD22]. Moreover, we notice that in the definition of the reduced c-squashed entanglement $E_{sq,par}^c$, one can obtain a weaker upper bound by considering solely classical extensions to Eve's system [YHH⁺09]. In that case a for a single tuple of measurements $\hat{\mathbf{x}}$ the bound reads (Corollary 2 in Ref. [HWD22])

$$\begin{aligned} K_{DI,dev}^{iid,\hat{\mathbf{x}}}(\rho_{N(A)}, \mathcal{M}) &\leq \inf_{(\sigma_{N(A)}, \mathcal{L}) = (\rho_{N(A)}, \mathcal{M})} \frac{1}{N-1} I(N(A) \downarrow E)_{P(A_1 : \dots : A_N | E)} \\ (2.91) \quad &\equiv \inf_{(\sigma_{N(A)}, \mathcal{L}) = (\rho_{N(A)}, \mathcal{M})} \inf_{\Lambda_{E \rightarrow F}} \frac{1}{N-1} I(N(A) | F)_{P(A_1 : \dots : A_N | \Lambda(E))}, \end{aligned}$$

where measurements $\mathcal{L}(\hat{\mathbf{x}})$ performed on the purification of $\sigma_{N(A)}$ yield the distribution $P(A_1 : \dots : A_N | E)$. The above is especially useful for numerical investigation. Furthermore, we notice that the reduced c-squashed entanglement E_{sq}^c can be defined for multiple measurements in analogy to results in Ref. [KHD22]

$$(2.92) \quad E_{sq}^c(\rho_{N(A)}, \mathcal{M}, p(\mathbf{x})) := \sum_{\mathbf{x}} p(\mathbf{x}) E_{sq}^c(\rho_{N(A)}, \mathcal{M}_{\mathbf{x}}).$$

From the above, we construct an upper bound on the DI-CKA rate for the protocols in which parties broadcast their inputs (see Proposition 2 in Ref. [HWD22])

$$\begin{aligned} K_{DI,dev}^{iid,broad}(\rho_{N(A)}, \mathcal{M}, p(\mathbf{x})) &\leq E_{sq,dev}^c(\rho_{N(A)}, \mathcal{M}, p(\mathbf{x})) \\ (2.93) \quad &:= \frac{1}{N-1} \inf_{(\sigma_{N(A)}, \mathcal{N}) = (\rho_{N(A)}, \mathcal{M})} E_{sq}^c(\sigma_{N(A)}, \mathcal{N}, p(\mathbf{x})). \end{aligned}$$

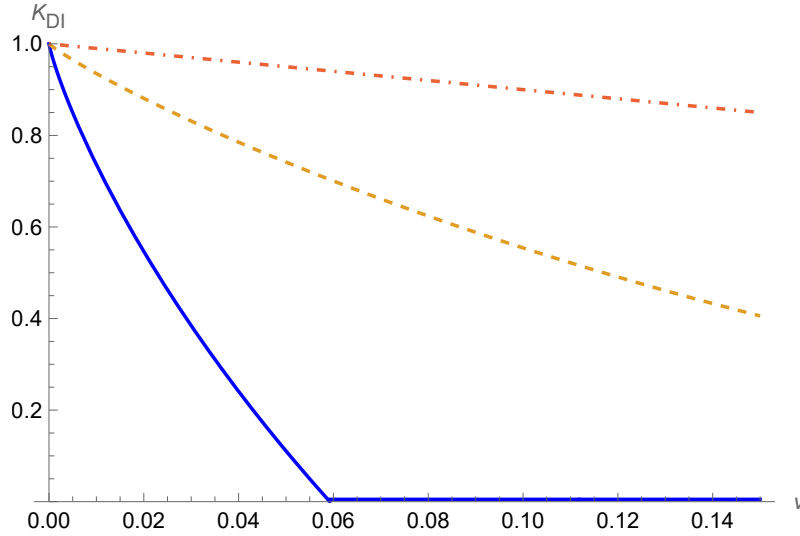


Figure 2.10: Modified plot from Ref. [HWD22] of upper bounds introduced in Ref. [HWD22] and lower bounds on the DI-CKA scenario in Ref. [RMW18]. Here, ν is the parameter of the noise in the depolarizing channel ($\mathcal{D}_\nu^{depol.}(\rho) = (1 - \nu)\rho + \nu\frac{\mathbb{1}}{2}$) that acts on each qubit. The yellow dashed curve represents an upper bound (not fully optimized) on the upper bound $\frac{1}{N-1}I(N(A) \downarrow E)$ from Eq. (2.91) (Corollary 2 in [HWD22]) with eavesdropper post-processing channel in Eq. (2.98). The red dash-dotted curve is the upper bound obtained in Corollary 6 of Ref. [HWD22] based on the relative entropy of entanglement bound $(1 - \nu)$. The solid blue line represents a lower bound, i.e., the rate of the protocol in Ref. [RMW18].

The investigation on the upper bounds constructed with the reduced c-squashed entanglement contains two significant technical results. Firstly we show that $E_{sq,par}^c$ is a convex function with respect to mixtures of quantum states that constitute the device (see Proposition 1 in Ref. [HWD22]), i.e.,

$$(2.94) \quad E_{sq,par}^c(\bar{\rho}, \mathcal{M}(\hat{\mathbf{x}})) \leq p_1 E_{sq,par}^c(\rho_1, \mathcal{M}(\hat{\mathbf{x}})) + p_2 E_{sq,par}^c(\rho_2, \mathcal{M}(\hat{\mathbf{x}})),$$

where $\bar{\rho} = p_1\rho_1 + p_2\rho_2$ and $p_1 + p_2 = 1$ with $0 \leq p_1 \leq 1$. The above results are useful because of the possible application of the convexification technique [WDH22] when studying numerical upper bounds. Secondly, we do not restrict the eavesdropper system to be finite-dimensional (see Lemma 6 in Ref. [HWD22]). This improvement closes a possible loophole in the proofs of Corollary 3 and Corollary 4 of Ref. [KHD22].

Our second main result is the numerical investigation that compares in the tripartite ($N = 3$) setting the upper bound derived by us with the rate of the parity Clauser-Horne-Shimony-Holt (parity CHSH) based protocol considered in Ref. [RMW18]. We restrict ourselves to the case with a classical eavesdropper in which the channels acting on Eve's system have only classical outputs. In this way, we exemplify the upper bound in Eq. (2.91) (see Corollary 2 in Ref. [HWD22]). To compare our upper bounds with the known lower bound, we consider the case of tripartite Greenberger-Horne-Zeilinger (GHZ) state $|\Phi_3^{\text{GHZ}}\rangle \langle \Phi_3^{\text{GHZ}}|$, on which depolarizing noise $\mathcal{D}_\nu^{depol.}$

parameterized with ν acts locally on each from the three qubits [RMW18]. For such a setting of the DI-CKA protocol in Ref. [RMW18], the statistics of the honest device reads

$$(2.95) \quad \begin{aligned} P_\nu(a, b_1, b_2 | x, y_1, y_2) &= \text{Tr} \left[M_{a|x} \otimes M_{b_1|y_1} \otimes M_{b_2|y_2} \mathcal{D}_\nu^{\otimes 3} (|\Phi_3^{\text{GHZ}}\rangle \langle \Phi_3^{\text{GHZ}}|_{AB_1B_2}) \right] \\ &= (1 - \nu)^3 P_{\text{GHZ}}(a, b_1, b_2 | x, y_1, y_2) + (1 - (1 - \nu)^3) P_\nu^{\text{L}}(a, b_1, b_2 | x, y_1, y_2), \end{aligned}$$

where the ranges of inputs and outputs are $x \in \{0, 1\}$, $y_1 \in \{0, 1, 2\}$, $y_2 \in \{0, 1\}$, and $a, b_1, b_2 \in \{0, 1\}$. The setting $(x, y_1, y_2) = (0, 2, 0)$ in the key-generating round is associated with measurements of σ_z observable. Furthermore, P_{GHZ} arises from local measurements on the GHZ state, and P_ν^{L} arises from the same measurements (for σ_z observable) on the fully separable state

$$(2.96) \quad \begin{aligned} \chi_\nu := \frac{1}{1 - (1 - \nu)^3} &\left((1 - \nu)^2 \nu \kappa_{AB_1} \otimes \frac{\mathbb{1}_{B_2}}{2} + (1 - \nu)^2 \nu \kappa_{AB_2} \otimes \frac{\mathbb{1}_{B_1}}{2} \right. \\ &\left. + (1 - \nu)^2 \nu \kappa_{B_1B_2} \otimes \frac{\mathbb{1}_A}{2} + (3 - 2\nu) \nu^2 \frac{\mathbb{1}_{AB_1B_2}}{2} \right), \end{aligned}$$

where $\kappa_{X_1X_2} = \frac{1}{2} (|00\rangle\langle 00|_{X_1X_2} + |11\rangle\langle 11|_{X_1X_2})$. The eavesdropper's strategy is based on the fact that P_ν^{L} can be expressed as a convex combination of deterministic behaviors. Therefore, Eve prepares a convex combination attack [AGM06, AMP06, FBJL+21]

$$(2.97) \quad \begin{aligned} P_\nu^{\text{CC}}(a, b_1, b_2, e | x, y_1, y_2) &= (1 - \nu)^3 P_{\text{GHZ}}(a, b_1, b_2 | x, y_1, y_2) \delta_{e,?} \\ &+ [1 - (1 - \nu)^3] P_\nu^{\text{L}}(a, b_1, b_2 | x, y_1, y_2) \delta_{e,(a,b_1,b_2)}, \end{aligned}$$

where the label $e = ?$ means that Eve is not correlated to the nonlocal part of the honest parties, i.e., Alice, Bob1, and Bob2, and $\delta_{e,(a,b_1,b_2)}$ means that Eve is correlated to all event associated with the local device. The above attack does not need to be optimal since it uses a particular decomposition of P_ν . In order to find a better strategy, Eve prepares a channel $E \rightarrow F$ that decomposes P_ν^{L} and does not diminish the weight of local behaviors $1 - (1 - \nu)^3$. In analogy to the attack given in bipartite case in Ref. [FBJL+21], we consider only the distribution coming from a key-generating measurement (see above)

$$(2.98) \quad \begin{aligned} P_\nu^{\text{ATTACK}}(a, b_1, b_2, f | 020) &= \Lambda_{E \rightarrow F} P_\nu^{\text{CC}}(a, b_1, b_2, e | 020) \\ &= (1 - \nu)^3 P_{\text{GHZ}}(a, b_1, b_2 | 020) \delta_{f,?} \\ &+ (1 - (1 - \nu)^3) P_\nu^{\text{L}}(a, b_1, b_2 | 020) (\delta_{a,b_1,b_2} \delta_{f,a} + (1 - \delta_{a,b_1,b_2}) \delta_{f,?}), \end{aligned}$$

where δ_{a,b_1,b_2} is 1 when $a = b_1 = b_2$ and 0 otherwise. The above attack strategy is thus a tripartite generalization of attack proposed in Ref. [FBJL+21] in which the eavesdropper aims to be correlated only to the $a = b_1 = b_2$ events associated with P_ν^{L} behavior. By applying the above attack, we plot the upper bound on $K_{DI,dev}^{iid,\hat{x}}(\rho_{N(A)}, \mathcal{M})$ in Fig. 2.10 using Eq. (2.91) (see Corollary 2 in Ref. [HWD22]).

For the third main result, we show that there exists a gap between device-dependent and device-independent conference key agreement rates. To obtain the result, we first generalize the reduced device-dependent conference key rate to the multipartite setting. The reduced device-dependent key agreement rate of state $\rho_{N(A)}$ is given by (see also Sec. 2.1.2.1)

$$(2.99) \quad K^\downarrow(\rho_{N(A)}) := \sup_{\mathcal{M}} \inf_{(\sigma_{N(A)}, \mathcal{L}) = (\rho_{N(A)}, \mathcal{M})} K_{DD}(\sigma_{N(A)}),$$

what yields (see Theorem 6 in Ref. [HWD22]) as an analog of Theorem 6 in Ref. [CFH21]

$$(2.100) \quad K_{DI}(\rho_{N(A)}) \equiv \sup_{\mathcal{M}} K_{DI}(\rho_{N(A)}, \mathcal{M}) \leq K^\downarrow(\rho_{N(A)}).$$

To show the existence of the strict gap between K_{DI} and K_{DD} in the multipartite setting (Theorem 7 in Ref. [HWD22]), we use the fact that there exist bipartite states ρ_{AB} for which $K^\downarrow(\rho_{AB}) < K_{DD}(\rho_{AB})$ as described in Ref. [CFH21]. Using these states we construct multipartite states $\rho_{N(A)}$ with the analogous property, i.e., $K_{DI}(\rho_{N(A)}) < K_{DD}(\rho_{N(A)})$. Our construction constitutes a path pointed by a chain of states $\rho_{A_k^1 A_k^2}$ shared between spatially separated N parties A_k , $k \in \{1, \dots, N-1\}$. Here, by the means of notation $A_1^1 \equiv A_1$ and $A_1^2 A_2^1 \equiv A_2, \dots, A_{N-1}^2 \equiv A_N$. In this way, we show in Theorem 7 of Ref. [HWD22] that a multipartite state

$$(2.101) \quad \tilde{\rho}_{N(A)} := \rho_{A_1^1 A_1^2} \otimes \rho_{A_2^1 A_2^2} \otimes \rho_{A_3^1 A_3^2} \otimes \dots \otimes \rho_{A_{N-1}^1 A_{N-1}^2},$$

where $\rho_{A_k^1 A_k^2} = \rho_{AB}$, $k \in \{1, \dots, N-1\}$, satisfies $K_{DI}(\rho_{N(A)}) < K_{DD}(\rho_{N(A)})$ as desired, provided all $\rho_{A_{k-1}^1 A_{k-1}^2}$ exhibit the gap, i.e., $K_{DD}(\rho_{A_{k-1}^1 A_{k-1}^2}) - K^\downarrow(\rho_{A_{k-1}^1 A_{k-1}^2}) \geq c > 0$. Indeed, such states do exist [CFH21], and so do $\tilde{\rho}_{N(A)}$. We remark here that the proof of Theorem 7 in Ref. [HWD22] can be relaxed to the case in which at least one state in the chain exhibits a bipartite gap and the theorem still holds.

Moreover, we discuss the topic of genuine nonlocality in analogy to the notion of genuine entanglement (see Secs. 2.1.4 and 2.1.5). We first define genuine quantum nonlocality by saying which devices are local and all that are not local are genuine nonlocal. Specifically, we assume that local are mixtures of devices that (i) admit quantum realization and (ii) are product in at least one partition into subsystems. In literature, there are various definitions of locality in multipartite case. Our class of local devices falls between the so-called TOBL and NSBL devices in Ref. [GWAN12] (cf. Ref. [Sve87, BCP⁺14]). In this way, we first say that behavior $P(\mathbf{a}|\mathbf{x})$ is local in a cut $(A_{i_1} \dots A_{i_k}) : (A_{i_{k+1}} \dots A_{i_N})$ for some $k \in \{1, \dots, N\}$, if and only if it can be written as a product of two behaviors of the systems correspondingly to the cut. Here, (i_1, \dots, i_N) is an arbitrary permutation of indices $(1, \dots, N)$. The set of all behaviors that are a product in some cut and have quantum realization is denoted by LQ, which means locally quantum (see Sec. 2.1.5). Furthermore, any distribution that is not locally quantum can be treated as nonlocal, although other definitions can be found in literature [BCP⁺14]. Consequently, we say that the behavior $P(\mathbf{a}|\mathbf{x})$ is genuinely nonlocal if and only if it is not

a mixture of behaviors that are local in at least one cut. As discussed, any behavior used by honest parties to obtain a conference key in a single-shot (single run) must exhibit genuine nonlocality. We consider a single-shot device-independent key distillation rate obtained by LOPC post-processing of a distribution obtained from some behavior $P(\mathbf{a}|\mathbf{x})$ when all inputs \mathbf{x} of all parties are set simultaneously

$$(2.102) \quad K_{DI,dev}^{\text{single-shot}}(\rho_{N(A)}, \mathcal{M}, \varepsilon) := \sup_{\hat{P}} \inf_{(2.81)} \kappa_n^\varepsilon \left(\hat{P}(\sigma_{N(A)}, \mathcal{N}) \right).$$

Here, κ_n^ε is the quantum key rate achieved for any measurements \mathcal{N} and security parameter ε . In Theorem 8 of Ref. [HWD22] we show that if $K_{DI,dev}^{\text{single-shot}}(\rho_{N(A)}, \mathcal{M}, \varepsilon) > 0$ then behavior $P(\mathbf{a}|\mathbf{x}) \equiv (\rho_{N(A)}, \mathcal{M})$ is not LQ. Furthermore, in Theorem 9 of Ref. [HWD22] we show the following upper bounds on DI-CKA rates

$$(2.103) \quad K_{DI,dev}(\rho_{N(A)}, \mathcal{M}) \leq \inf_{(\sigma_{N(A)}, \mathcal{L})=(\rho, \mathcal{M})} E_{GE}^\infty(\sigma_{N(A)}),$$

$$(2.104) \quad K_{DI,par}(\rho_{N(A)}, \mathcal{M}) \leq \inf_{\substack{\omega(\sigma_{N(A)}, \mathcal{L})=\omega(\rho_{N(A)}, \mathcal{M}) \\ P_{err}(\sigma_{N(A)}, \mathcal{L})=P_{err}(\rho_{N(A)}, \mathcal{M})}} E_{GE}^\infty(\sigma_{N(A)}).$$

Here, $E_{GE}^\infty(\varsigma)$ is the regularized relative entropy of genuine entanglement [DBWH21] for a state $\varsigma_{A_1 A_2 \dots A_N}$, with

$$(2.105) \quad E_{GE}^\infty(\varsigma) = \inf_{\varphi \in \text{BS}(N(A^{\otimes n}))} \lim_{n \rightarrow \infty} \frac{1}{n} D(\varsigma^{\otimes n} \| \varphi),$$

where $D(\rho \| \sigma) = \text{Tr } \rho \log_2 \rho - \text{Tr } \rho \log_2 \sigma$ (if $\text{supp } \rho \subseteq \text{supp } \sigma$ and ∞ otherwise [Ume62]) is the relative entropy between two states ρ and σ , and BS denotes the set of biseparable states (see Sec. 2.1.4 for details) [DBWH21].

To conclude, we have introduced several of upper bounds on the quantum secure conference key. Some of the introduced upper bounds generalize the relative entropy-based upper bound in Ref. [KHD22], and the reduced c-squashed entanglement provides another upper bounds as the generalization of results in Ref. [FBJL+21]. Moreover, we have shown constructive proof that a fundamental gap between the device-dependent and the device-independent key is inherited from the bipartite setting (see Ref. [AFL21, CFH21]) and holds in a multipartite case. We remark here, that Ref. [HWD22] is the first published manuscript considering upper bounds on DI-CKA key rates.

Note: We prove analogous upper bounds based on a different definition of the multipartite squashed entanglement. We refer to these bounds as dual bounds, as described in Sec. IV of Ref. [HWD22]. After publishing the paper, it was pointed out by M.E. Shirokov that the aforementioned definitions are equivalent due to Theorem 7 of Ref. [DSW18]. Therefore, our dual upper bounds are equivalent as well. The erratum to the article in Ref. [HWD22] has been published in Ref. [HWD23].

2.5 Limitations on Device-Independent Key Secure Against Nonsignaling Adversary via the Squashed Nonlocality [D]

In the recently published paper, [WDH22], we initiate a systematic study of upper bounds on the secret key rate in the non-signaling device-independent secret key agreement (NSDI) scenario [BHK05, BCK12, SGB⁺06, AMP06, AGM06, Mas09, HRW10, MRC⁺14]. We begin, by showing that the security definition, that we adopt, based on the so-called non-signaling norm, is equivalent to two notions of security already present in the literature in Refs. [MRC⁺14, Mas09, MPA11] and Refs. [HRW10, HRW13, HR10, Hän10]. In order to prove the above equivalence, we derive the explicit form of the non-signaling norm for the so-called \mathbf{c} -d states. Our main result is the so-called squashing procedure. We show any secrecy quantifier that is an upper bound on the secret key rate in the SKA scenario [Mau93, MW99, MW97, RW03, RSW03] subjected to the squashing procedure becomes an upper bound on the secret key rate in the NSDI scenario (see Sec. 2.1.2.1). The squashing procedure is based on two elements, i.e., on the notion of the non-signaling complete extension [WDS⁺18] and suitable optimization over measurements performed by the honest parties and the eavesdropper. Lifting the formalism by means of the squashing procedure was possible as we rephrased the definition of the secret key rate in the SKA scenario in the form known from quantum key distribution scenarios (QKD). Next, we concentrate on one of the squashed upper bounds, i.e., the non-signaling squashed intrinsic information hereafter called the squashed nonlocality. We prove that the squashed nonlocality is a novel unfaithful measure of nonlocality to what it owes its name. Moreover, we develop a convexification technique that allows to combine different upper bounds into a single tighter one. Finally, we perform a numerical investigation in which we compare our upper bound via the squashed nonlocality with the lower bounds given by the rates of Hänggi-Renner-Wolf [HRW10] and by Acín-Massar-Pironio [AMP06] protocols.

The NSDI scenario has the most relaxed assumptions from all security paradigms considered in this thesis. Here, the eavesdropper is constrained only with the no-signaling conditions that make it more powerful than the classical or quantum adversary. Similarly, the honest parties can possibly share supra-quantum correlations constrained only with the no-signaling conditions as well. In a nutshell, the no-signaling conditions restrict the input-output statistics of the devices shared between the parties in a manner that excludes the possibility of instantaneous communication [Bar07]. The possible advantage of the NSDI scenario over other cryptographic paradigms (e.g., SKA, QDD, QDI) is that it assures security even if a new theory of physics is established that would replace the Quantum Theory (QT) and allow for stronger violation of Bell inequalities than QT allows, provided the new theory is non-signaling one as well.

In our considerations, we assume the presence of two honest parties, Alice and Bob, and malicious Eve being the eavesdropper. The object shared by them is a tripartite non-signaling device $P(ABE|XYZ)$ (tripartite normalized, conditional probability distribution), with X, Y

and Z being inputs, A , B and E being outputs, of Alice, Bob and Eve respectively. On the device, the honest parties perform direct measurements $\mathcal{M}_{x,y}^F$ (MD) (X, Y) and post-process their output data (A, B) with some local operations and public communication (LOPC) in order to produce the secure-key. We refer to this class of operations used for the distillation of the secret-key as to MDLOPC class (see Sec. 2.1.2.1). The device is assumed to be provided by Eve, who can listen to public communication and is correlated with the subsystem shared by Alice and Bob as strongly as possible under the considered assumptions. This situation constitutes the worst-case scenario for the honest parties.

In the context of the NSDI scenario, mainly lower bounds on the secret key rate have been studied so far [BHK05, AGM06, AMP06, SGB⁺06, Mas09, HRW10, BCK12, MRC⁺14]. To our best knowledge, the only upper bound considered so far was in Ref. [AMP06]. This situation stands in contrast with the SKA [CK78, Mau93], QDD [BB84, Eke91, Ben92, GRTZ02, ABB⁺06], and QDI [Eke91, BCP⁺14, MY04, ABG⁺07, MPA11, AFRV19] scenarios in which both lower and upper bounds are known [CK78, Mau93, DW04, KGR05, DW05, CEH⁺07, HHHO05, AH09b, HHHO09, Chr02, CEH⁺07, YHH⁺09, Wil16, TGW14b, TGW14a, PLOB17, KWW20, Kau20, CFH21, AFL21, FBJL⁺21, KHD22, HWD22]. Most importantly, for the NSDI scenario, it was shown that in the presence of a collective eavesdropper's attack, a non-zero key rate can be obtained [HRW10, Mas09, MRC⁺14] under the fully no-signaling constraints. By the fully no-signaling constraints, one means that none of the $2N + 1$ subsystems of the device (N for each of the honest parties and 1 for the eavesdropper) can signal to each other. This assumption is vital because if the subsystems of an honest party can signal between each other, or the device has a memory and can perform the so-called forward signaling between the runs, then no hash function that can achieve privacy amplification against the non-signaling eavesdropper is known and the standard ones fail the security requirements [HRW13, AFTS12, SW16]. The fully non-signaling assumption can be achieved in the so-called parallel measurement model in which measurements on $2N$ subsystems are performed simultaneously.

Addressing the problem of upper bounds, it is enough to consider that the device shared between the honest parties consists of N independent and identically distributed (iid) copies of a non-signaling device $P(AB|XY)$. The total system of the honest parties and the eavesdropper is then a tripartite probability distribution $Q(ABE|XYZ)$, that has $P^{\otimes N}(AB|XY)$ as its marginal after discarding (tracing out) Eve's subsystem. As we assume the worst-case scenario, we allow the eavesdropper to have access to all statistical ensembles of the device of the honest parties. This power is achieved with Eve having access to the non-signaling complete extension of the device $P^{\otimes N}(AB|XY)$ denoted $\mathcal{E}(P^{\otimes N})(ABE|XYZ)$ that is a non-signaling analog of quantum purification [WDS⁺18] (for more details see Sec. 2.6). Moreover, unlike the honest parties, the malicious Eve can perform a generalized measurement M_z^G on her subsystem, allowing for additional local randomness on her input.

In the QDD and QDI scenarios, some upper bounds are based on the entanglement measure called “squashed entanglement” [Chr02, CEH⁺07, YHH⁺09, Wil16, TGW14b, TGW14a, CFH21, AFL21, FBJL⁺21]. A welcome feature measure of it is that it is an additive function of quantum states, as opposed to the relative entropy of entanglement [HHHO09, HHHO05, AEJ⁺01, PBL⁺18, PAB⁺20], which allows for avoiding the need for its regularization. Although the nonlocality analog of relative entropy has been constructed [vDGG05, GHH⁺14], no analog of the squashed entanglement has been studied until recently [KWW20]. In our approach, which is guided by the analogies between quantum entanglement and nonlocality, we introduce (independently of Ref. [KWW20]) a novel nonlocality measure “squashed nonlocality”. It is formulated differently but equivalent to the measure implicitly introduced in Ref. [AMP06]. We base our formulation on the non-signaling complete extension that is an analog of quantum purification employed in the definition of the squashed entanglement. In this way, our findings contribute to more than just the development of the NSDI scenario but additionally to the fundamental topic of nonlocality.

As mentioned, our approach bases on the analogies between different cryptographic paradigms. Firstly, we introduce a novel security criterion for the NSDI key distillation protocols Λ . The criterion is based on the non-signaling complete extension and the non-signaling norm (cf. Ref. [CT09]) that plays a role of an operational distance measure between non-signaling devices. The non-signaling norm reads

$$(2.106) \quad \|P - P'\|_{\text{NS}} := \sup_{g \in \mathcal{G}} \frac{1}{2} \|g(P) - g(P')\|_1,$$

where \mathcal{G} is a certain (broad) subset (specified in Ref. [WDH22]) of linear operations \mathcal{L} that map devices into probability distributions. In this way, a sequence of MDLOPC operations $\Lambda = \{\Lambda_N\}$ (the MDLOPC key distillation protocol) performed by the honest parties on N iid copies of a device $P(AB|XY)$ is said to be secure if the following condition holds (see Sec. 2.1.2.1)

$$(2.107) \quad \left\| P_{\text{out}} - P_{\text{ideal}}^{(d_N)} \right\|_{\text{NS}} \leq \varepsilon_N \xrightarrow{N \rightarrow \infty} 0,$$

$$(2.108) \quad P_{\text{ideal}}^{(d_N)}(s_A, s_B, q, e|z) := \frac{\delta_{s_A, s_B}}{|S_A|} \sum_{s'_A, s'_B} P_{\text{out}}(s'_A, s'_B, q, e|z),$$

where $P_{\text{out}} = \Lambda_N \left(\mathcal{E} \left(P^{\otimes N} \right) \right)$, is the non-signaling complete extension of the N iid copies of the $P(AB|XY)$ distribution, and $P_{\text{ideal}}^{(d_N)}$ is the secure (perfect) state in which Eve is uncorrelated with the honest parties. In Eq. (2.108) s_A and s_B stand for the output of the honest parties, i.e., the secret-key, q represents communication, e and z are input and output of Eve respectively, finally d_N stands for adequate dimension. Security based on the non-signaling norm constitutes an analogy to the QDD scenario, where security is based on the trace distance between quantum states. The non-signaling norm in Eq. (2.107) measures the distance between the so-called classical-device (c-d) states shared at the end of a MDLOPC protocol. In c-d states the classical

part is shared by Alice and Bob while Eve still holds a device. Furthermore, by showing an explicit form of the non-signaling norm in the case of \mathbf{c} -d states (see Proposition 2 in Ref. [WDH22]) we prove in Theorem 3 of Ref. [WDH22] that the security criterion adapted by us (see Eq. (2.107)) is equivalent to the two notions of security already present in the literature [MRC⁺14, Mas09, MPA11] and [HRW10, HRW13, HR10, Hän10], respectively. The explicit form of the non-signaling norm reads

$$(2.109) \quad \left\| P_{A,E|Z}^1 - P_{A,E|Z}^2 \right\|_{\text{NS}} = \frac{1}{2} \sum_a \sup_{\mathcal{M}_z^F} \sum_e \left| \mathcal{M}_z^F \left(P_{A,E|Z}^1 \right) (a, e) - \mathcal{M}_z^F \left(P_{A,E|Z}^2 \right) (a, e) \right|,$$

where $a \in A$ corresponds to multi-variable of outputs of the \mathbf{c} part in the \mathbf{c} -d distribution, and \mathcal{M}_z^F is a direct measurement performed on input Z . By showing the equivalence between the security criteria based on the non-signaling norm and the one in Refs. [HRW10, HRW13, HR10, Hän10] (based on the so-called distinguisher) we argue that our notion of security is restricted composable [BOM04, BOHL⁺05, Can01], i.e., it is composable provided the device is not reused (at a risk of the leakage of key generated from the first use, during second use of the device [BCK12]). Finally, the security definition based on the non-signaling complete extension guarantees that the memory of the eavesdropper is finite and minimal without compromising the eavesdropping power of Eve [WDS⁺18].

Our main goal is to provide upper bounds on the secret-key rate in the NSDI scenario. To achieve our target, we introduce the so-called squashing procedure. The squashing procedure establishes an analogy between SKA and NSDI scenarios in the following manner. Suppose $M(A : B|E)$ is a real-valued, non-negative function in the domain of tripartite probability distributions $P(ABE)$, that serves as an upper bound on the secret-key rate in the SKA scenario [Mau93, MW99, MW97, RW03, RSW03], i.e., $M(A : B|E) \geq S(A : B|E)$. Moreover, if $M(A : B|E)$ is monotonic with respect to LOPC, then we call it a secrecy monotone. Considering the above, we construct quantifiers of secret correlations in the NSDI model. Namely, the squashing procedure “lifts up” secrecy quantifiers $M(A : B|E)$ in the SKA scenario to give non-signaling secrecy quantifiers of the NSDI scenario $\widehat{M}(A : B|E)$ via mapping tripartite non-signaling devices $R(ABE|XYZ)$ into probability distributions

$$(2.110) \quad \widehat{M}(A : B|E)_{R(ABE|XYZ)} := \max_{x,y} \min_z M(A : B|E)_{(\mathcal{M}_{x,y}^F \otimes \mathcal{M}_z^G)R(ABE|XYZ)},$$

$$(2.111) \quad (\mathcal{M}_{x,y}^F \otimes \mathcal{M}_z^G)R(ABE|XYZ) = \sum_z p(z|z')R(ABE|X = x, Y = y, Z = z),$$

where by $\max_{x,y}$ we mean maximization over all possible direct measurements $\mathcal{M}_{x,y}^F$, and by \min_z we mean minimization over all possible general measurement \mathcal{M}_z^G , where optimization over probability distribution $p(Z|Z')$ is implicit. Additionally, if $R(ABE|XYZ)$ is the non-signaling complete extension, i.e., $R(ABE|XYZ) \equiv \mathcal{E}(P)(ABE|XYZ)$, we call $\widehat{M}(A : B|E)_{\mathcal{E}(P)(ABE|XYZ)}$ a non-signaling squashed secrecy quantifier. Furthermore, suppose $\widehat{M}(A : B|E)_{R(ABE|XYZ)}$ is a secrecy (MDLOPC) monotone. In that case, we call it a non-signaling secrecy quantifier

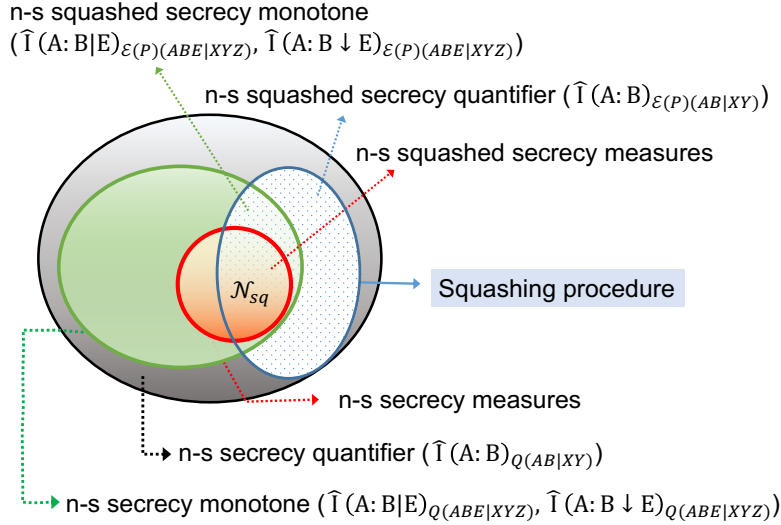


Figure 2.11: The relative hierarchy of non-signaling secrecy quantifiers in Ref. [WDH22]. The grey region represents all non-signaling secrecy quantifiers constructed from secrecy quantifiers known from the SKA security paradigm. The green region represents non-signaling secrecy monotones (MDLOPC monotones). The red circle, with the squashed nonlocality as a distinguished example, corresponds to the nonlocality measures, i.e., MDLOPC monotones that are non-negative, additive for the tensor product of devices, and zero for the local devices. Finally, the bluish region corresponds to the non-signaling secrecy quantifiers incorporating the non-signaling complete extension $\mathcal{E}(P)$ of bipartite device P in their construction, whereas Q stands for any tripartite extension of P . Between brackets, we give examples of non-signaling secrecy quantifiers belonging to each set.

monotone, and if additionally, $R(ABE|XYZ)$ is the non-signaling complete extension, we call consistently $\widehat{M}(A : B||E)_{\mathcal{E}(P)(ABE|XYZ)}$ non-signaling squashed secrecy monotone. The relation between “lifted quantifiers” is depicted in Fig. 2.11, where the inclusion between the sets (red, green, gray) is proved to be strict [WDH22].

Having the squashing procedure defined, we present our main result concerning the upper bounds on the secret-key rate (Theorem 1 in Ref. [WDH22]). Namely, we show that the rate of the secret-key $K_{DI}^{(iid)}(P)$ distilled with MDLOPC operations in the non-signaling device-independent iid scenario (see Sec. 2.1.2.1 for details) is upper bounded with any non-signaling squashed secrecy quantifier constructed with the squashing procedure

$$(2.112) \quad \forall_P K_{DI}^{(iid)}(P) \leq \widehat{M}(A : B||E)_{\mathcal{E}(P)}.$$

Here, $P(AB|XY)$ is a single copy of a bipartite non-signaling device, of which the honest parties share multiple iid copies. The above result, together with the presence of the squashing procedure, establishes a direct connection between SKA and NSDI cryptographic scenarios. In this way, we not only establish a link between two major security paradigms but also open a possibility for a systematic study of tighter upper bounds on the secret-key rate in the NSDI

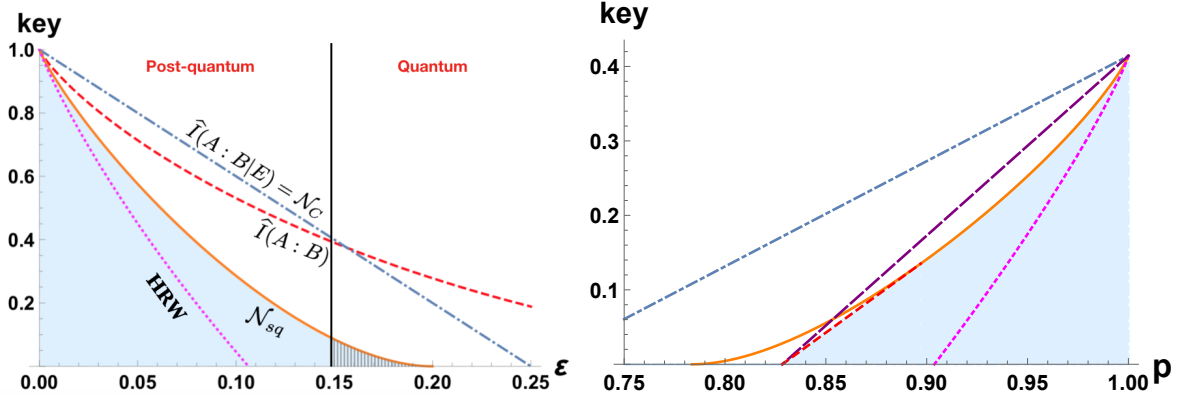


Figure 2.12: Plot of the upper and lower bounds on the secret-key rate $K_{DI}^{(iid)}(P)$ in the NSDI scenario in Ref. [WDH22]. The plot on the left corresponds to the $(2,2,2,2)$ scenario in which the devices are probabilistic mixtures of PR and anti-PR box, i.e., $P_{\text{iso}}(\varepsilon) = (1 - \varepsilon)\text{PR} + \varepsilon\overline{\text{PR}}$. The dashed red line corresponds to the non-signaling squashed mutual information. The straight blue line represents non-signaling squashed conditional mutual information (equal to the nonlocality cost [WDH22, EPR92, BCSS11]). The solid orange line represents an upper bound on the squashed nonlocality obtained with the convexification technique. The dotted magenta curve (HRW) corresponds to the lower bound obtained from Hänggi, Renner, and Wolf protocol [HRW10]. The “post-quantum” region corresponds to devices that exhibit supra-quantum correlations. The plot on the right corresponds to the $(3,2,2,2)$ scenario in which the devices are given by the $P_{\text{AMP}}(p)$ distribution (see Refs. [WDH22, AMP06] for details). The solid orange line corresponds to an upper bound on the squashed nonlocality obtained with the convexification technique. The blue dash-dotted line corresponds to the non-signaling squashed conditional mutual information. The dotted magenta curve is the lower bound by Acín, Massar, and Pironio (AMP) in Ref. [AMP06]. The purple big-dashed curve is the AMP upper bound on intrinsic information of the eavesdropping strategy in Ref. [AMP06]. The red dashed line is a segment of the lower convex hull of the orange and purple curves that illustrates the convexification technique. The shaded bluish region under orange and red curves represents the tightest obtained upper bound.

scenario (cf. Ref. [AMP06]). The performance and numerical feasibility of the constructed upper bounds are thanks to the presence of the non-signaling complete extension in the definition of the non-signaling quantifiers. The non-signaling complete extension allows for explicit representation of the eavesdropper system for which the memory is finite (cf. Ref. [WDS⁺18]).

The proof of general upper bound on the secret-key rate $K_{DI}^{(iid)}(P)$ in Eq. (2.112) (presented in Theorem 1 in Ref. [WDH22]) was obtained thanks to an intermediate technical result concerning SKA security paradigm. Namely, we show (Theorem 2 in Ref. [WDH22]) that the following definition of the secret-key rate $S(A : B||E)$ (see Sec. 2.1.2.1) is equivalent to the definition present in the literature [CK78, Mau93, MW00, CEH⁺07]

$$(2.113) \quad S(A : B||E) = \sup_{\mathcal{P}} \limsup_{N \rightarrow \infty} \frac{\log \dim_{\mathbb{A}} \left(\mathcal{P}_N \left(P^{\otimes N} (ABE) \right) \right)}{N},$$

with a security condition

$$(2.114) \quad \left\| \mathcal{P}_N \left(P^{\otimes N} (ABE) \right) - P_N^{\text{ideal}} \right\|_1 \leq \delta_N \xrightarrow{N \rightarrow \infty} 0,$$

$$(2.115) \quad P_N^{\text{ideal}}(S_A S_B C E^N) := \left(\frac{\delta_{s_A, s_B}}{|S_A|} \right) \otimes \sum_{s_A, s_B} P_N^{\text{out}}(S_A = s_A, S_B = s_B, C E^N),$$

$$(2.116) \quad P_N^{\text{out}}(S_A, S_B, C E^N) := \mathcal{P}_N \left(P^{\otimes N} (ABE) \right).$$

Here, $\mathcal{P} = \cup_{N=1}^{\infty} \{\mathcal{P}_N\}$ is a LOPC cryptographic protocol, acting on N iid copies of the classical probability distribution $P(ABE)$, and P_N^{ideal} is a distribution representing ideally secure key. Moreover, C is a random variable representing public communication in the protocol. The above technical result is interesting on its own because it rephrases the definition of the secret-key rate $S(A : B || E)$ in the SKA model into the form that is characteristic for the quantum scenarios. In this way, as a byproduct, we contribute to the SKA cryptographic paradigm.

Among all non-signaling secrecy quantifiers, we pay special attention to the non-signaling squashed intrinsic information $\mathcal{N}_{\text{sq}}(P)$, that we call the squashed nonlocality (cf. Ref. [KWW20] for parallel approach)

$$(2.117) \quad \mathcal{N}_{\text{sq}}(P) := \max_{x,y} \min_z I(A : B \downarrow E)_{(\mathcal{M}_{x,y}^F \otimes \mathcal{M}_z^G) \mathcal{E}(P)(ABE|XYZ)}.$$

We prove that the squashed nonlocality, besides being an upper bound on the secret-key rate $K_{DI}^{(iid)}(P)$, exhibits many important properties. Namely, the squashed nonlocality is (see Proposition 1 in Ref. [WDH22])

1. **Non-negative real-valued function**, equal zero for local devices,
2. **Monotonic** with respect to MDLOPC operations,
3. **Convex** with respect to mixtures of non-signaling devices,
4. **Superadditive** over joint non-signaling devices,
5. **Additive** for product of devices,
6. **Subextensive**,
7. **Non-faithful** (proof by inspection).

In consequence, we prove that the squashed nonlocality is a (novel) measure of nonlocal correlation, what makes it a nonlocality analog of the squashed entanglement being an entanglement measure. Importantly, by inspection, we prove that the squashed nonlocality is non-faithful, i.e., there exist bipartite nonlocal devices for which the squashed nonlocality is zero (see Fig. 2.12). We illustrate the non-faithfulness in the (2,2,2,2) setting (two binary inputs, two binary outputs)

for the isotropic boxes $P_{\text{iso}}(\varepsilon)$, i.e., probabilistic mixtures of the PR box and anti-PR box ($\overline{\text{PR}}$). The non-faithfulness implies that nonlocality does not imply security for the MDLOPC protocols (cf. Ref. [AGM06] where a different protocol is considered).

The intrinsic information function and, therefore, also the squashed nonlocality implicitly contains post-processing of the outputs of the eavesdropper. Therefore, in order to generate examples of the upper bounds, we perform an extensive numerical investigation that aims to find possibly optimal post-processing channels. We consider devices with statistics dependent on some set of parameters (see Fig. 2.12). However, the optimal channel can be different for each set of values of the parameters. Furthermore, if one finds a post-processing channel that diminishes the upper bound curve (with respect to the identity channel), proving that it is optimal is a formidable task. This situation motivated us to develop a convexification technique that has already proved its usefulness in situations other than the NSDI scenario [KHD22, HWD22]. The convexification technique states that (see Section I of Appendix in Ref. [WDH22]) the lower convex hull of an arbitrary number of plots, each being an upper bound on the squashed nonlocality is an upper bound on $K_{DI}^{(iid)}(P)$ itself. This result is possible because the squashed nonlocality is convex with respect to mixtures of devices. In our investigation, the upper bounds on the squashed nonlocality are functions (based on the squashed nonlocality) with typically suboptimal post-processing channels. Using the convexification technique, we obtain the tightest to our knowledge upper bound (cf. Ref. [AMP06]) on the NSDI secret-key in the (3,2,2,2) scenario (see Fig. 2.11). We also compare our upper bound with the lower bound obtained from the protocol of Hänggi, Renner, and Wolf (HRW protocol) [HRW10] in order to see the proximity of the curves, and therefore the performance of the upper bound. Furthermore, the convexification technique can be easily generalized and developed for other convex non-signaling squashed secrecy quantifiers, or different scenarios [KHD22, HWD22].

To conclude, in the described paper [WDH22], we have developed a plethora of analogies between different security paradigms. The notion of the complete extension (see Sec. 2.6 for more details) allowed us to develop a number of analogies between the NSDI, SKA and quantum scenarios. In particular, the non-signaling complete extension allowed us to introduce a family of the non-signaling squashed secrecy quantifiers being upper bounds on the MDLOPC secret-key rate in the NSDI scenario. One of the non-signaling squashed secrecy quantifiers, called the squashed nonlocality, is proved to be a novel and nonfaithful measure of nonlocality (see Ref. [KWW20] for parallel approach). The convexification technique developed in Ref. [WDH22] allowed us to find the tightest upper bound in the (3,2,2,2) NSDI scenario. In this way, we contribute not only to the development of the NSDI or SKA cryptographic paradigms but also the fundamental topic of nonlocality.

2.6 Complete Extension: the Non-Signalling Analog of Quantum Purification [E]

In the recently accepted for publication manuscript [WDS⁺18], we study a concept of the complete extension postulate (CEP) in the framework of generalized probabilistic theories (GPTs) [Bar07, Plá21, Mü121, Lam18]. The CEP is meant to be a relaxation of the purification postulate (PP) [CDP11], i.e., one of the postulates in terms of which quantum theory can be reformulated. The PP is motivated by quantum purification. Considering relaxations of postulates, in terms of which quantum theory can be formulated, is a research direction motivated by the chase for beyond-quantum theory with stronger explanatory power for physical phenomena than quantum theory. On the contrary to the PP, the CEP does not require the existence, within a theory, of pure extensions for all systems and their states. Instead, the CEP requires the existence of extension with the property of GENERATION, i.e., extensions from which any other extension of the extended system can be prepared using operations available in theory. We start by discussing the limitations of the purification postulate in terms of the no-go theorem for hyperdecoherence [LS18], the contradiction between the existence of fundamental superselection rules and PP [SSC21, Nak20, WvdW22], and quantum gravity in the context of theories with indefinite casual structure [Har05, Har07, OCB12, CDPV13, AFNB17] and discrete theories. In particular, we prove a no-go theorem in which we claim that the purification postulate fails in any discrete convex theory, i.e., in a convex theory in which the number of pure states is countable. We then formally introduce the complete extension postulate (CEP) and show that GENERATION property enables access to all probabilistic ensembles of the extended system via measurements on extending system (ACCESS), i.e., GENERATION implies ACCESS. As we discuss, classical probability theory and superselected quantum theory both satisfy CEP. We then show that in any GPT that satisfies the CEP and the no-restriction hypothesis, the bit-commitment cryptographic task is impossible, like in quantum theory. On the other hand, consecutively, we demonstrate that the proof for the no-go theorem for hyperdecoherence of beyond-quantum theory to quantum theory that holds under the PP does not hold under the CEP. Subsequently, we case study the theory of non-signaling behaviors and try to bypass the nonexistence of purifications therein by constructing extensions satisfying ACCESS and GENERATION properties, i.e., the non-signaling complete extensions (NSEAs). In fact, we show that in the theory of non-signaling behaviors, the ACCESS and GENERATION properties are equivalent. The NSEA of a generic behavior is not a pure state of the theory; still, some NSEAs are pure. As we show, NSEA of a maximally mixed binary input binary output behavior is the PR box [PR94]. The above observation can be viewed as an alternative derivation of the PR box, as it does not refer to any type of Bell inequality. Next, we derive an upper bound on the dimension of the NSEA as a function of the dimensionality of the extended system. Importantly, we show that the dimension of NSEA is always finite. We observe that NSEA, contrary to

quantum purification, does not exhibit a mirror property. Namely, in a generic case, the NSCE, ρ'_{AB} of the reduced state σ_B of the extending system B that is in NSCE, ρ_{AB} is not the original NSCE of state ρ_A of the system A . Finally, we conduct a numerical investigation in which we find NSEAs of particular contextual behaviors and behaviors lying on isotropic line (see Sec. 2.1.6.2).

Whilst the quantum theory is the best-validated description of physical reality, it lacks a single agreed-on interpretation [Per02], and it cannot explain some observable phenomena such as gravity. For this reason, a significant research effort has been put into finding a beyond-quantum theory with stronger explanatory power [PR94, Ber17, Ber18, Har01, Życ08, Smo06]. The sought theory must rather be some modification of the quantum theory. However, standard axioms of quantum theory are not independent of each other [Sto32, Gle75, MGM19], and they can not be individually modified. Therefore, recently, an interest in reformulating quantum theory in terms of mutually independent postulates occurred. This idea led to quite a number of reformulations of quantum theory in terms of information-theoretic postulates [DB11, Har01, CDP10, Har11, CBH03, Goy08, MM11, BMU14, Wil18, Höh17, BR17, SSC21, Tul20, vdW19, Nak20] within the framework of the so-called generalized probabilistic theories (GPTS) [Plá21, Mül21, Lam18], [Bar07] and references therein. One such reformulation of quantum theory introduced the purification postulate (PP) [CDP11], which gives an important insight as it distinguishes classical probability theory from quantum theory. Substantially, the purification postulate is a generalization of the notion of purification within the quantum theory to arbitrary GPTs. The PP already proved its usefulness in proving many results in the fields of pertaining to computation [LS16b, LS16a, BLS18], cryptography [CDP10, SS18], thermodynamics [CS15, CS17], and interference [BLSS17]. On the other hand, the PP should be modified, for several reasons, if we want to find the beyond-quantum theory to be a more fundamental description of nature than the quantum theory. Firstly, in Ref. [LS18], there is a no-go theory which states that if the mentioned beyond-quantum theory is causal and reduces to quantum theory via decoherence-like mechanism (hyperdecoherence), then it does not satisfy the PP. Secondly, the PP is contradictory with the existence of superselection rules; therefore, if one believes in those, the PP should be abandoned. Next, in Ref. [AFNB17], it was shown that a purification-like postulate fails to hold for all quantum process matrices. Therefore, there are doubts if the PP is suitable for theories that permit indefinite causal structure [Har05, Har07, OCB12, CDPV13], for example, quantum gravity. Finally, in Refs. [BHZ05, Mül09, Pal20], ideas motivated by the pursuit of quantum gravity suggest that on the fundamental level, the quantum state space becomes discrete. We anticipate here a bit, as according to our findings, the PP can not hold in any discrete theory. All the arguments listed above suggest asking how the PP can be weakened or replaced by another postulate yielding a theory with more explanatory power. In Ref. [WDS⁺18], we propose the complete extension postulate (CEP) as a replacement for PP and study its implications in various contexts.

We formulate the complete extension in the language of GPTs by employing the notions of extensions, ensembles, and properties of ACCESS and GENERATION. We restrict ourselves to the convex theories, i.e., theories in which convex mixtures of states are states as well, as they seem to be natural candidates for the beyond-quantum fundamental theory. Let s be a vector in the convex set of states (state space) Ω_A of a GPT system A . An ensemble for s , is then a set of pairs $\{(p_i, s_i)\}_{i \in I}$ such that $s_i \in \Omega_A$ and $\{p_i\}$ define a probability distribution ($p_i \in \mathbb{R}$, $p_i \geq 0$, and $\sum_i p_i = 1$), satisfying the convex combination

$$(2.118) \quad s = \sum_{i \in I} p_i s_i.$$

The set of all possible ensembles for a state s is denoted $\mathbf{Ens}[s]$. Next, an extension of a state $s \in \Omega_A$ of system A is a state $\sigma \in \Omega_{A \otimes E}$ of a bipartite system $A \otimes E$ for which

$$(2.119) \quad s = [\mathbb{1}_A \otimes u_E](\sigma),$$

where u_E is the unit effect on system E , and we denote the set of extensions σ , of a state s , as $\mathbf{Ext}[s]$. Importantly, in general, there will be extensions constructed on many different extending systems E . If purification exists, it is simply some $\sigma \in \mathbf{Ext}[s]$ which is pure, i.e., $\mathbf{Purif}[s] \subseteq \mathbf{Ext}[s]$. We distinguish a specific class of extensions $\mathbf{Ext}_{class}[s]$ in which the extending system is taken to be classical Δ_d for some $d \in \mathbb{N}$ (see Sec. 2.1.6.1 for details). As we show (see Proposition 9 in Ref. [WDS⁺18]) that there is an isomorphism between $\mathbf{Ext}_{class}[s]$ and $\mathbf{Ens}[s]$, i.e., $\mathbf{Ext}_{class}[s] \cong \mathbf{Ens}[s]$. Third, an extension $\sigma^* \in \mathbf{Ext}[s]$ with extending system E^* is said to be generating (satisfy GENERATION) if and only if for every $\sigma \in \mathbf{Ext}[s]$ with arbitrary extending system E there exists a transformation $T_\sigma : E^* \rightarrow E$ such that:

$$(2.120) \quad \mathbb{1}_A \otimes T_\sigma(\sigma^*) = \sigma.$$

Furthermore, an extension $\sigma^* \in \mathbf{Ext}[s]$ with extending system E^* is an extension with access (satisfies ACCESS property) if and only if, for any ensemble $\sigma \in \mathbf{Ext}_{class}[s]$ with extending system Δ_I we can find a measurement $M_\sigma : E^* \rightarrow \Delta_I$ such that:

$$(2.121) \quad \mathbb{1}_A \otimes M_\sigma(\sigma^*) = \sigma.$$

Having the basic notions defined, in Proposition 10 and 11 of Ref. [WDS⁺18], we prove that if an extension σ has the property of GENERATION, then it necessarily has the property of ACCESS, but not vice versa. We discuss the above nonequivalence in terms of quantum theory and GPT, which has the same state space as the quantum theory but restricted dynamics. Finally, we formulate the complete extension postulate (CEP).

A GPT \mathcal{G} satisfies the Complete Extension Postulate (CEP) iff: for all systems A and all states $s \in \Omega_A$, there exists an extension $\sigma^* \in \mathbf{Ext}[s]$ which is generating, that is, which has the GENERATION property.

The complete extension postulate (CEP) can be compared with the purification postulate (PP) [CDP10].

A GPT \mathcal{G} satisfies the Purification Postulate if and only if for all systems A and all states $\omega_A \in \Omega_A$, there exists purifications ($\mathbf{Purif}[\omega_A]$) which are essentially unique.

In the above definition, essential uniqueness means that all purifications constructed on the same extending system can be related via reversible transformation on the purifying system. As we discuss, the essential uniqueness property of the PP is intuitively very closely related to the GENERATION property of the CEP. However, an important difference is that essential uniqueness is a property for purifications of the same extending system, while GENERATION is a property for extensions on arbitrary extending systems. Derivation of GENERATION from essential uniqueness requires one more assumption and sets the relation between the CEP and the PP. As we prove in Proposition 13 of Ref. [WDS+18] for the set of GPTs in which the product of pure states is pure, the purification postulate implies the complete extension postulate. Consequently, quantum theory satisfies CEP. Namely, purification $|\psi\rangle$ of state ρ is also its complete extension, as it allows for GENERATION of arbitrary other extensions as well as ACCESS of arbitrary ensembles. Note that the converse statement is not true, i.e., let ω be arbitrary mixed quantum state, then $|\psi\rangle \otimes \omega$ is the complete extension of ρ , but not its purification. Moreover, as we show, also classical probability theory satisfies CEP. Indeed, the complete extension of a probability distribution of some random variable X is described by its copy. Specifically, an extension of a probability distribution $p(x)$ is any probability distribution $q(x, y)$, such that $\sum_q q(x, y) = p(x)$. In this way, $p_{ce}(x, x') := p(x)\delta_{x, x'}$ is a complete extension of $p(x)$, and any other extension can be obtained by applying a stochastic map to x' . The above results highlight the differences between the PP and the CEP, as the PP does not hold in classical probability theory. In this way, the CEP seems to be a more natural postulate to consider than the PP, as it is based on an intuitive property of GENERATION and holds both in quantum and classical theory.

As our first main result, we prove that there can not exist purification of arbitrary states in any non-signaling, convex discrete theory [PW13, CHHH17, JGBB11] of finite dimension. This no-go result was known previously only for the classical theory and the theory of non-signaling behaviors [CDP10, CDP11]. This result is not directly connected to the CEP. However, as mentioned before, the no-go theorem described below serves as an argument justifying the need to replace the PP. More precisely, a GPT \mathcal{G} is said to be discrete, if and only if, for each system

$T \in \text{Syst}[\mathcal{G}]$ there is a discrete number of pure states, that is, where each state space, Ω_T , is a polytope. The proof of Theorem 7 of Ref. [WDS⁺18], which states that in any discrete theory, there are no (non-trivial) systems that have purifications for all states, and hence, theories with purifications can not be discrete, is based on a simple observation. Namely, in Lemma 6, we observe that in any convex (non-trivial) theory, the cardinality of the set of its states is at least of power of continuum \mathfrak{c} . On the other hand, the number f of vertices in any discrete theory is finite, and $f < \aleph_0 < \mathfrak{c}$ according to a set-theoretic fact [Kur61]. So there are simply not enough vertices, in theory, to purify all the mixed states of any system. The direct consequence of Theorem 7 is Proposition 8 [WDS⁺18], which states that in any theory (with a countable number of system types), there must be at least one system with a continuum of pure states.

The task of bit-commitment, and its generalization to integer-commitment [May97, LC98, SS18], is an important two-party cryptographic primitive that can be used to build other cryptographic protocols such as coin-flipping [Blu81] and zero-knowledge proofs [GMR85]. The integer-commitment protocol for two parties, Alice and Bob, goes as follows. The task consists of two phases i) the commit phase, in which Alice chooses from the uniform random distribution an integer $j \in \{1, \dots, n\}$ and "commits" it to Bob by passing him a "token". Here, the "token" refers to a (sub)system prepared accordingly to the chosen integer. There are two of these tokens. At this stage, it should be impossible for Alice to change the value of j , as well as Bob should not be able to learn the value of j . In the reveal phase, ii) Alice communicates the integer j to Bob and sends him a second token to verify the integer. Finally, Bob should be able to verify if the integer Alice communicated is the one she drew in the first phase. The security condition for the protocol is that Alice can not change the value of the integer j after the first phase, to cheat on Bob, and Bob can not learn the value of the integer j before the second phase to cheat on Alice. The cheating probabilities p_A^* and p_B^* of Alice and Bob, respectively, quantify to which extent the two parties can deviate the protocol from the ideal one. Namely, p_A^* is the maximum probability with which Alice can reveal another integer than the one she committed to Bob. Similarly, p_B^* is the maximum probability with which Bob can learn the value of j before the reveal phase. It is known that (perfect realization of) bit-commitment is impossible both in classical and quantum theory [May97, LC98]. On the other hand, it was shown in Ref. [BDLT08] that any GPT which is non-classical but does not have entanglement allows for bit-commitment. It is, therefore, interesting to determine which properties of theories make the integer-commitment task impossible. In Ref. [CDP10, Corollary 45], the impossibility of bit-commitment was proved under a set of postulates, including the purification postulate (PP). Furthermore, in Ref. [SS18] an analytical lower bound on the product of cheating probabilities $p_A^* p_B^*$, for the integer-commitment task, has been found, again relying on the PP

$$(2.122) \quad p_A^* \cdot p_B^* \geq \frac{\alpha}{n} > \frac{1}{2n},$$

where the perfect protocol would be the one for which $p_B^* = p_A^* = 0$. As we show in Theorem 7 of Ref. [WDS⁺18], exactly the same lower bound holds if one replaces the PP with the CEP

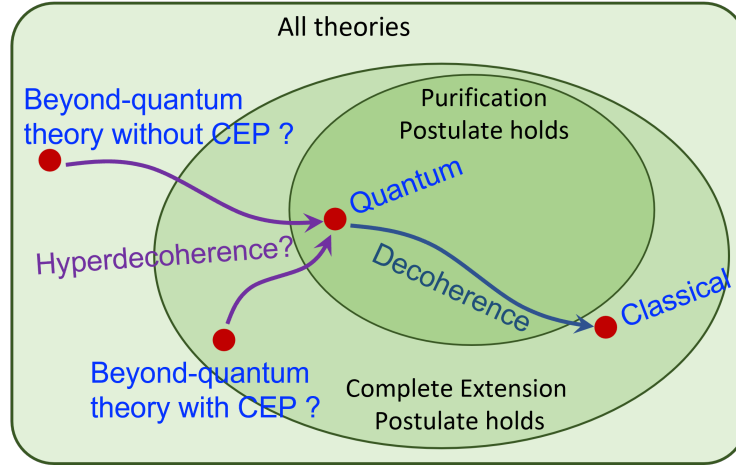


Figure 2.13: Illustration of the relationship between the set of all theories, theories which satisfy the CEP, and those which satisfy The PP. Hyperdecoherence mechanism (purple arrows) relates hypothetical beyond quantum theories and quantum theory in analogy to the decoherence mechanism (blue arrow), which relates quantum and classical theory).

and assumes the no-restriction hypothesis [CDP10]. The proof of Theorem 7 in Ref. [WDS+18] goes along the lines of the proof in Ref. [SS18], with the same cheating strategies of Alice and Bob, up to the modification in which essential uniqueness is replaced with the GENERATION property. The above demonstrates that the impossibility of the integer-commitment is not restricted to the theories that satisfy the PP, as the CEP suffices. On the other hand, it shows that the CEP is not an empty postulate as the lower bound in Eq. (2.122) holds. Additionally, using the CEP, we provide a unified proof applicable to both quantum and classical theory.

Decoherence is a mechanism that shows how quantum states, measurements, and transformations become effectively classical [CST18, SC17]. Note that the diagonal density operators are isomorphic to classical probability distributions. In Ref. [LS18], the notion of decoherence was generalized to the concept of hyperdecoherence, a term coined in [Życ08]. In analogy to decoherence, hyperdecoherence (see Fig. 2.13) is supposed to be a mechanism that would make underlying beyond-quantum theory effectively quantum and explain why we do not observe beyond-quantum effects in all our experimental tests (except the theory of gravity). The basic idea to describe hyperdecoherence is that for every system A of beyond-quantum theory, there exists a decoherence process

$$(2.123) \quad \mathbf{H}_A : A \rightarrow A,$$

which causes that system A to behave effectively as a quantum system. We assume that these hyperdecoherence processes satisfy the following properties that assure consistency of the mechanism resulting in valid physical theory

1. They are unit-effect preserving,

$$(2.124) \quad u_A \circ \mathbf{H}_A = u_A,$$

in analogy to the trace-preservation condition for quantum decoherence processes.

2. They are idempotent,

$$(2.125) \quad \mathbf{H}_A \circ \mathbf{H}_A = \mathbf{H}_A,$$

what reflects the fact that once the beyond-quantum features are lost, the processes of hyperdecoherence do nothing.

3. They must be chosen compositionally, i.e., they are such that

$$(2.126) \quad \mathbf{H}_{A \otimes B} = \mathbf{H}_A \otimes \mathbf{H}_B.$$

In other words, if two systems hyperdecohere independently, then the joint system will behave effectively as a quantum one as well.

The hyperdecoherence of a beyond-quantum theory is then modeled by replacing the unit process $\mathbb{1}_A$ with hyperdecoherence processes \mathbf{H}_A . The intuition behind the described model is that the hyperdecoherence happens at time scales much shorter than those we can probe experimentally and the systems hyperdecohere before we actually do something. In Ref. [LS18], it was shown that any beyond-quantum theory that hyperdecoheres to the quantum theory, in the way described here, must violate either the causality principle or the PP. While abandoning the PP seems more plausible, the question arises whether the CEP can be satisfied by a beyond-quantum theory that hyperdecoheres to quantum theory. We examine the proof of the main theorem in Ref. [LS18] and conclude that one can not derive, using the identical proving technique, the same no-go result by replacing PP with CEP. Therefore, it remains possible that there exists a beyond-quantum theory that hyperdecoheres to quantum theory, satisfies CEP, and does not violate the causality principle.

As a case study, for the application of the CEP, we choose the theory of non-signaling behaviors [PR94, Bar07]. The states within this theory are multipartite conditional probability distributions, e.g. $P_{\mathbf{A}|\mathbf{X}}$ for $\mathbf{A} = A_1, \dots, A_N$ and $\mathbf{X} = X_1, \dots, X_N$, that satisfy the no-signaling constraints [Bar07] (see Sec. 2.1.6.2 for more details). The theory of non-signaling behaviors allows for stronger correlations between subsystems than quantum theory. For example, it reaches the algebraic maximum of the Clauser-Horne-Shimony-Holt (CHSH) inequality [PR94], and it violates the Information Casuality Principle [PPK⁺09]. Furthermore, the theory of non-signaling behaviors is an example of a discrete theory. Hence, it does not satisfy the PP. In our case study, we aim to bypass the lack of existence of purifications, for a generic state, with the counterpart of purifications originating from the CEP, i.e., by constructing non-signaling

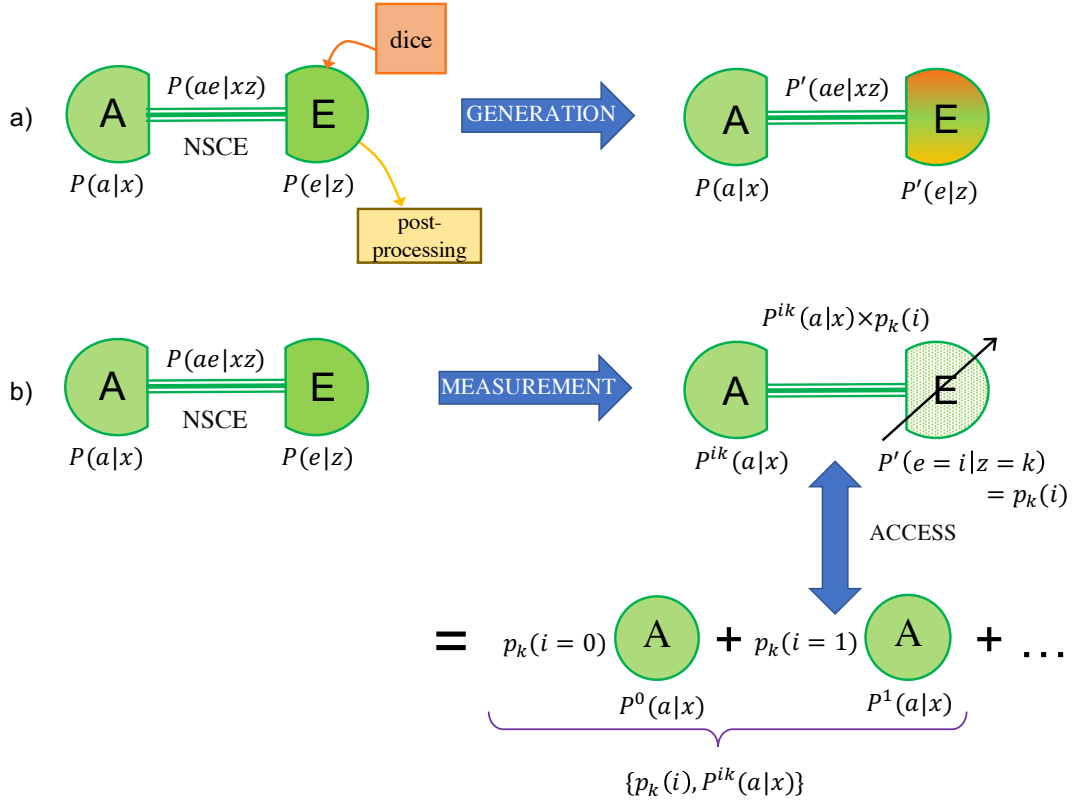


Figure 2.14: Illustration of ACCESS and GENERATION properties for the theory of non-signaling behaviors in Ref. [WDS⁺18]. a) Using NSCE, one can generate any other extension by means of randomizing the input and post-processing of output. b) The extending system E of the NSCE has access to any ensemble $\{p_k(i), P^{ik}(a|x)\}$ of the system A , which is generated upon measurement $z = k$ performed on the system E .

complete extensions (NSCEs). More precisely, for a given behavior P_A , we want to construct its non-signaling extension P_{AE} , such that it satisfies ACCESS and GENERATION properties, with an analogy to quantum purification which satisfies these. Indeed, we achieve our goal and show that NSCEs can actually be constructed, and therefore, the theory of non-signaling behaviors satisfies the CEP. We do so by first defining the overcomplete non-signaling extension with access (ONSEA) that gives access to all pure members ensembles (PMEs) of the extended system via a choice of the measurement setting (input) on the extending system. In fact, ONSEA gives access to all ensembles of extended system, also mixed ones, as those are convex mixtures of pure members ensembles (see Theorem 21 of Ref. [WDS⁺18]). Next, in Theorem 16 of Ref. [WDS⁺18], we show that ACCESS to PME is equivalent to ACCESS to minimal ensembles, i.e., these ensembles that contain a set of vertices such that any of its proper subsets does not suffice to construct an ensemble of the extended system. The equivalence comes again due to dynamics available in the theory for the extending system. This equivalence motivates us to define non-signaling extension with access (NSEA), namely

Given a behavior $P_A : P_A(a|x)$, we say that a behavior $P_{AE} : P_{AE}(ae|xz)$ is its non-signalling extension with access extended to system E if for any input choice $z = k$ an outcome $e = i$ occurred in the extending system, there holds

$$(2.127) \quad P_{AE}(a, e = i|x, z = k) = P_E^{i,k}(a|x)p(e = i|z = k)$$

such, that, for each k , the ensemble $\left\{ \left(p(e = i|z = k), P_E^{i,k}(a|x) \right) \right\}_i$ is a minimal ensemble of the behavior P_A . Moreover, corresponding to each minimal ensemble of P_A , there is exactly one input $z = k$, in part of the extending system which generates it.

In Proposition 19 of Ref. [WDS⁺18], we show that for each behavior P_A , its NSEA $\mathcal{E}(P)_{AE}$ is unique up to local relabeling (of inputs and outputs) on the extending system E , what makes an analogy to essential uniqueness. In virtue of Corollary 22 of Ref. [WDS⁺18] that follows from Corollary 20 and Theorem 21 of Ref. [WDS⁺18], NSEA gives access to all ensembles of the extended system, and therefore, NSEA satisfies ACCESS. Next, in Theorem 23 of Ref. [WDS⁺18], we show that in the theory of non-signaling behaviors, the properties of ACCESS and GENERATION are equivalent. Finally, in Corollary 23 of Ref. [WDS⁺18], we conclude that NSEA is NSCE as it satisfies ACCESS and GENERATION, and therefore, the theory of non-signaling behaviors satisfies the CEP (see Fig. 2.14). Finally, later, in Proposition 27 of Ref. [WDS⁺18], we show that NSEA (or NSCE equivalently), of arbitrary behavior P , has the lowest dimension amongst all non-signaling behaviors of P having the property of ACCESS.

We further give an explicit example of how NSCEs can be constructed in the theory of non-signaling behaviors. One of the behaviors for which we show the construction of NSCE is the maximally mixed behavior with a single binary input and single binary output given by

$$(2.128) \quad P_A^m(a|x) = \begin{array}{c|cc} & x & \\ \hline a & & \\ \hline & 0 & 1 \\ \hline & 0 & 1/2 & 1/2 \\ \hline & 1 & 1/2 & 1/2 \end{array}$$

where a is the output and x is the input (measurement setting) of the behavior on system A . As we show, the behavior given in Eq. (2.128) above provides an interesting example of NSCE. We start our construction by determining vertices of the polytope, i.e., state space P_A^m belongs to,

$$(2.129) \quad P_E^0 = \begin{array}{c|cc} & x & \\ \hline a & & \\ \hline & 0 & 1 \\ \hline & 0 & 1 & 1 \\ \hline & 1 & 0 & 0 \end{array}, \quad P_E^1 = \begin{array}{c|cc} & x & \\ \hline a & & \\ \hline & 0 & 1 \\ \hline & 0 & 0 & 0 \\ \hline & 1 & 1 & 1 \end{array}, \quad P_E^2 = \begin{array}{c|cc} & x & \\ \hline a & & \\ \hline & 0 & 1 \\ \hline & 0 & 1 & 0 \\ \hline & 1 & 0 & 1 \end{array}, \quad P_E^3 = \begin{array}{c|cc} & x & \\ \hline a & & \\ \hline & 0 & 1 \\ \hline & 0 & 0 & 1 \\ \hline & 1 & 1 & 0 \end{array},$$

where the subscript E corresponds to the fact that behaviors above are "extremal" points of the polytope. We now employ the fact that any point inside a convex polytope can be expressed

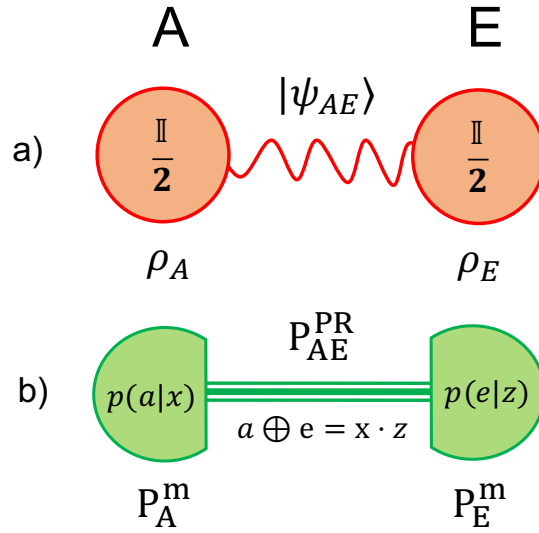


Figure 2.15: Illustrations of purifications of system A to an extended system E in Ref. [WDS⁺18]. a) The purification of maximally mixed state of qubit $\frac{1}{2}$ is a Bell's state which is maximally entangled. b) NSCE of a maximally mixed behavior P_A^m (see Eq. (2.128)) is the Popescu-Rohrlich box which is also pure.

as a convex combination of its vertices. In this way, we find all minimal ensembles of the P_A^m behavior

$$(2.130) \quad \mathcal{M}_0(P_A^m) = \{(1/2, P_E^0); (1/2, P_E^1)\}$$

$$(2.131) \quad \mathcal{M}_1(P_A^m) = \{(1/2, P_E^2); (1/2, P_E^3)\}.$$

Now, using the definition of NSCE (or equivalently NSEA) above, these two minimal ensembles are obtained on the extended system for two different input choices on the extending system. We associate now the choice of minimal ensembles $\mathcal{M}_0, \mathcal{M}_1$ in Eqs. (2.130), (2.131), with input choices on the extending system $z = 0, z = 1$, respectively. In this way, we obtain the NSCE of the behavior P_A^m

$$(2.132) \quad P_{AE}^{\text{PR}}(ae|xz) = \begin{array}{c|cc|cc|cc} & x & 0 & 1 & & & & \\ & a & & & & & & \\ z & e & & & & & & \\ \hline 0 & 0 & \frac{1}{2} & 0 & \frac{1}{2} & 0 & & \\ & 1 & 0 & \frac{1}{2} & 0 & \frac{1}{2} & & \\ \hline 1 & 0 & \frac{1}{2} & 0 & 0 & \frac{1}{2} & & \\ & 1 & 0 & \frac{1}{2} & \frac{1}{2} & 0 & & \end{array}.$$

We recognize the NSCE of P_A^m to be the famous Popescu-Rohrlich (PR) box [PR94] (see Sec. 2.1.6.2 for more details). The PR box is an extreme point in its polytope. In this point, we

recognize the analogy to the purification of a maximally mixed state of a qubit, i.e., $\rho = \frac{1}{2}\mathbb{1}$, which purification is maximally entangled Bell's state (up to local isometry) [HHHH09] (see Fig. 2.15). As we conclude in Corollary 33 of Ref. [WDS⁺18], the PR box is a purification of maximally mixed behavior with a single binary input and a single binary output P_A^m . We also discuss that observing that PR box is the purification of P_A^m can be seen as a derivation of PR box without referring to any notion of CHSH or the so-called Bell inequality, in contrast to Ref. [PR94]. In essence, the independence of our derivation relies on the fact that we assume only a) the structure of single-partite systems and b) that the NSCE of any single-partite system is a valid state of the theory. However, we do not presuppose: i) that all non-signaling behaviors all valid states of the theory, ii) a specific composition rule, iii) the no-restriction hypothesis [CDP10]. Furthermore, in section B.1 of the appendix of Ref. [PR94], we show explicit construction of NSCE for certain single-partite behavior, which is not pure. In section B.3 of the appendix of Ref. [PR94], we construct NSCE if three-cycle contextual behavior [AB11, AQB⁺13] (aka Specker's triangle [Spe90, LSW11]). Eventually, in section B.4 of the appendix of Ref. [PR94], we determine all minimal ensembles of bipartite behaviors lying on the isotropic line (see Sec. 2.1.6.2) that can be readily used to construct NSCEs, and study their properties.

As our next main result, we study the dimension of the non-signaling complete extension (NSCE). By the dimension of NSCE or any other non-signaling behavior, we understand the dimension of the polytope (state space) it belongs to. Our purpose is to derive an upper bound on the dimension of the NSCE of n -partite behavior belonging to polytope \mathcal{B} , with m_i inputs for parties, and v_{ij} outputs respectively. We denote by $\tilde{\mathcal{B}}$ the polytope that contains NSCE of arbitrary behavior in \mathcal{B} . In Theorem 25 of Ref. [WDS⁺18], we show the following upper bound on the dimension of polytope $\tilde{\mathcal{B}}$

$$(2.133) \quad \dim \tilde{\mathcal{B}} < (\dim \mathcal{B} + 1) \times \left(\left(\binom{2t - \lfloor t/2 \rfloor - \dim \mathcal{B}}{\lfloor t/2 \rfloor} + \binom{3t - \lfloor t/2 \rfloor - (\dim \mathcal{B} + 1)}{t - \lfloor t/2 \rfloor - 1} \right) \dim \mathcal{B} + 1 \right),$$

where:

$$(2.134) \quad \dim \mathcal{B} = \prod_{i=1}^n \left(\sum_{j=1}^{m_i} (v_{ij} - 1) + 1 \right) - 1, \quad t = \prod_{i=1}^n \sum_{j=1}^{m_i} v_{ij}.$$

In the proof, we base on the formula in Eq. (2.134) [Pir05] (Theorem 1 therein) and some facts from convex geometry. To find an upper bound on $\dim \tilde{\mathcal{B}}$, we have to determine upper bounds on the number of inputs m_{n+1} and outputs $v_{n+1,j}$ of the extending party. We firstly determine the maximal number of outputs, for each input generating minimal ensemble, of the extending party via Carathéodory theorem [Zie95] to be $\dim \mathcal{B} + 1$, i.e., $v_{n+1,j} \leq \dim \mathcal{B} + 1$. On the other hand, the number of inputs of the extending party is equivalent to the number of minimal ensembles. Assuming V to be the number of vertices in polytope \mathcal{B} , and using Carathéodory theorem again, we obtain $m_{n+1} \leq \binom{V}{\dim \mathcal{B} + 1}$. Finally, the number of vertices V can be upper

bounded using McMullen’s Upper Bound Theorem [McM70, Zie95, ELS06] and properties of non-signaling polytopes

$$(2.135) \quad V \leq \binom{2t - \lfloor t/2 \rfloor - \dim \mathcal{B}}{\lfloor t/2 \rfloor} + \binom{3t - \lfloor t/2 \rfloor - (\dim \mathcal{B} + 1)}{t - \lfloor t/2 \rfloor - 1}.$$

Combining the above upper bounds yields the formula in Eq. (2.133). An immediate corollary from Theorem 25 (Corollary 26 in Ref. [WDS⁺18]) is that the dimension of the NSCE is always finite. Investigating the dimension of the NSCE is not only interesting on its own. Namely, access to all possible ensembles of a non-signaling system has been considered operationally the worst-case extension the extending party might possess in the context of device-independent cryptography against a non-signaling adversary [BHK05, Mas06, HRW10, Hän10] (see also Ref. [WDH22] and Sec. 2.5), and also in the context of private randomness [CR12, GMT⁺13, MGP15, BRG⁺16, RBH⁺16]. The adversary having access to the extending system of NSCE possesses, therefore, ultimate operational power at the lowest memory cost required for having access to all ensembles of the extended system.

In summary, in Ref. [WDS⁺18], we introduce a new concept of the complete extension postulate (CEP) as a relaxation of the purification postulate (PP). We show that the CEP postulate is satisfied in classical theory, quantum theory, the theory of non-signaling behaviors, and, moreover, in any theory in which the product of pure states is pure. We have shown that the CEP does not allow for the integer-commitment cryptographic task, as well as that it does not exclude, in contrast to the purification postulate, a beyond-quantum theory that hyperdecoheres to quantum theory. In this way, the CEP sets a demarcation line between results that require the PP to hold and those for which the CEP is enough. Our case study in the theory of non-signaling behaviors presents an alternative derivation of the Popescu-Rohrlich box as the non-signaling complete extension (NSCE) of maximally mixed behavior with a single binary input and single binary output. We study the dimension of the NSCE state space. We showed that the dimension of NSCE is always finite and pointed out important implications for cryptography against the non-signaling adversary.

Outlook

The results described in the articles incorporated in this dissertation establish fundamental limitations of the major quantum and supra-quantum secret key distribution paradigms that are about to be the main building blocks of the future quantum-secure Internet. The vast landscape of adamant limitations on the secret key distribution tasks drawn in this dissertation is, however, a single element of the quantum revolution that is about to happen in the not-too-distant future. Still, determining the upper bounds is of no lesser importance than constructing new cryptographic protocols and tasks. This concern is because any newly created or experimentally applied secret key distribution protocol has to confront the limitations of the secret key distillation rates. In this context, the upper bounds described in this dissertation set the aforementioned type of limitations on the several selected secret key distribution scenarios of crucial importance. Importantly, any violation of the upper bounds on the secret key rate instantly implies that the produced cryptographic key can not be secure and points out theoretical shortcomings in the application of the proposed protocol or technological problems with the devices used for its implementation. Furthermore, if one finds a correct and feasible protocol for a specific cryptographic task that achieves the corresponding upper bound, then the search for a better protocol can be completed. This situation can only happen if we have the upper bounds on the considered scenario or the upper bounds we know they are tight enough. The findings described in this dissertation open a pathway for further investigations in the meaningful field of upper bounds on the secret key rates. The techniques developed here allow both, at least in some cases, for developing tighter upper bounds and extensions to other cryptographic scenarios. Moreover, apart from the application-directed perspective on the upper bounds that establish the core of this dissertation, the novel measures of nonlocality and multipartite entanglement presented here in the context of upper bounds on the secret key rate, have their reflection in the foundations of the quantum theory. Determining the properties

of the proposed measures, along with the development of their frameworks, constitute tasks of no lesser importance than finding their applications in the description of the cryptographic scenarios. In the forthcoming paragraphs, we sketch possible further progress that can be done in the direction indicated in this dissertation.

Regarding the results of Ref. [SWRH20] described in Sec. 2.2. The performance of the countermeasure to the rerouting attack proposed therein is benchmarked with two quantities, i.e., the gap of the scheme and the percentage of the memory overhead. The definitions of the gap of the scheme and the memory overhead include the rate of the device-dependent secret key. It is, therefore, a crucial question if the solution proposed as a countermeasure to the rerouting attack can be extended to the case of device-independent security of the secret key while preserving the reasonable gap and memory overhead of the scheme. The change of the security paradigm in this place is, however, not meaningless regarding the performance of the scheme. Private states and PPT states approximating them are known to provide a low rate of device-independent key, significantly lower than in the device-dependent key [CFH21]. This drawback can dramatically diminish the gap of the scheme and significantly increase the memory overhead resulting in extra low performance of the countermeasure. Determination of the set of states yielding relatively good performance of the secure hybrid network scheme is, therefore interesting direction for further investigation.

The upper bounds on the secret key rates derived in Ref. [DBWH21] and described here in Sec. 2.3 are given by various types of relative entropy functions [Ume62, MLDS⁺13, Dat09b, Dat09a, MLDS⁺13, WWY14, BD10, WR12]. The framework developed in Ref. [DBWH21] allows for describing a plethora of different scenarios in a unified way. On the other hand, the squashed entanglement [Chr02, CEH⁺07] and the c-squashed entanglement [YHH⁺09] are known to be entanglement measures that upper bound the bipartite and multipartite (conference) secret key rates respectively. It is, therefore, an interesting open question if one can develop a similar unified framework based on a multipartite generalization of the squashed entanglement rather than the one based on relative entropies originating from the generalized divergence. Furthermore, as stated in Ref. [DBWH21], the identification of new information processing tasks and determination of bounds on the achievable rates for the classical and quantum communication protocols over a multiplex channel (see, for example, Refs. [Sha61, GK11, BHTW10, WDW17, Das18, LALS20, TR19, DW19]) is undoubtedly an important future direction of research.

Speaking of the possible future developments of results in Ref. [HWD22], described in Sec. 2.4, the situation is a mirror reflection to the described in the previous paragraph in the case of Ref. [DBWH21]. In the Ref. [HWD22], the derived upper bounds on device-independent conference key rate that are given in terms of the c-squashed entanglement based on the squashed and cc-squashed entanglement functions [Chr02, CEH⁺07, YHH⁺09, AFL21, KHD22]. It is, therefore, tempting to ask if tighter upper bounds can be derived in terms of the relative

entropy functions [Ume62, MLDS⁺13, Dat09b, Dat09a, MLDS⁺13, WWY14, BD10, WR12] and in particular in terms of the reduced relative entropy of entanglement in analogy to results in Ref. [KHD22]. Furthermore, as remarked in Ref. [HWD22] our results hold in the static scenario of quantum states. A possible further investigation includes a generalization of the results in Ref. [MLK⁺16] to the dynamic case of quantum channels by employing and developing the formalism developed in Ref. [HWD22]. Moreover, determining whether there exists a strict gap between $K_{DI,dev}$ and $K_{DI,par}$ is an interesting question on its own. Finally, it would be interesting to determine whether Theorem 8 in Ref. [HWD22] still holds if we consider the TOBL class of devices instead of the locally quantum class.

In Ref. [WDH22], we developed a method for obtaining upper bounds on the device-independent secret key secure against a non-signaling adversary (NSDI scenario) by lifting the upper bounds on the secret key rate in the secret key agreement scenario (SKA). The upper bound we concentrate on the most in Ref. [WDH22] is the new measure of nonlocality, i.e., the so-called squashed nonlocality (non-signaling squashed intrinsic information) \mathcal{N}_{sq} that is based on the intrinsic information function [MW99, MW97]. However, it is known that reduced intrinsic information [RSW03, RW03] is an example of an upper bound on SKA key rate that is, for some probability distributions, strictly lower than intrinsic information [RW03]. Therefore, the first step in obtaining tighter upper bounds on the NSDI key rate is to construct non-signaling squashed reduced intrinsic information and investigate its properties. Furthermore, it is an important question to answer whether the squashed nonlocality [WDH22] and intrinsic non-locality developed in Ref. [KWW20] are in fact the same functions. The negative answer to the aforementioned issue would yield many new questions about the relations between the squashed nonlocality and intrinsic non-locality. The generalization of the method developed in Ref. [WDH22] to the multipartite case constitutes a new primary research direction (cf. Ref. [PKBW23]). Eventually, the study of the properties of the squashed nonlocality and its generalization is interesting on its own. For instance, the study that proves the asymptotic continuity property of the squashed nonlocality is currently in an advanced stage of development.

Finally, the investigation conducted in Ref. [WDS⁺18], and described here in Sec. 2.6, considers the beyond quantum theory analog of the purification postulate (PP) that holds in quantum theory, i.e., the so-called complete extension postulate (CEP). Our hitherto research raises numerous further questions not only about the consequences of the complete extension postulate (CEP) but also about its best shape. In particular, an interesting further research direction is highlighted when one considers generalized probabilistic theories (GPTs) in which GENERATION and ACCESS properties are non-equivalent. In the present shape, the complete extension postulate is defined to satisfy the GENERATION property by definition. As proved in Ref. [WDS⁺18], GENERATION implies ACCESS, but the reverse statement is not true in the general case. Therefore, replacing the GENERATION property with the ACCESS property in the definition of the complete extension postulate can supposedly bring a richer and more

interesting structure of the complete extensions. The existence of a GPT in which the (complete) extensions allow access to all of the statistical ensembles of the extended system but do not allow to generate an arbitrary extension (assuming no-restriction hypothesis [CDP10] holds) is an interesting question about the structure of such a theory. Furthermore, the next step to develop the framework of CEP would be to consider the possibility of studying its dynamical counterpart, i.e., a GPT analog of the Stinespring dilation theorem. Finally, there are still some open questions regarding the non-signaling complete extension (NSCE). One of them is the lack of the mirror property that holds for quantum purification but does not hold for NSCE (see Ref. [WDS⁺18] for more details).

In summary, the articles included in this dissertation [SWRH20, DBWH21, HWD22, WDS⁺18] provide many fundamental and application-directed results in the fields of quantum and supra-quantum cryptography. However, as it usually is, the insightful investigations conducted therein yield even many more interesting questions apart from those already answered and open numerous new research directions. Fortunately, the approaches developed in Refs. [SWRH20, DBWH21, HWD22, WDS⁺18] have structures flexible enough to immediately notice the possible generalizations and further applications of their formalism. We look forward to the new results in the research areas considered in this dissertation, as well as those connected to them.

Bibliography

- [AB11] Samson Abramsky and Adam Brandenburger.
The sheaf-theoretic structure of non-locality and contextuality.
New Journal of Physics, 13(11):113036, November 2011.
- [ABB⁺06] A. Acín, J. Bae, E. Bagan, M. Baig, Ll. Masanes, and R. Muñoz-Tapia.
Secrecy properties of quantum channels.
Physical Review A, 73(1), January 2006.
- [ABG⁺07] Antonio Acín, Nicolas Brunner, Nicolas Gisin, Serge Massar, Stefano Pironio, and Valerio Scarani.
Device-independent security of quantum cryptography against collective attacks.
Physical Review Letters, 98(23), June 2007.
- [AEJ⁺01] K. Audenaert, J. Eisert, E. Jané, M. B. Plenio, S. Virmani, and B. De Moor.
Asymptotic relative entropy of entanglement.
Physical Review Letters, 87(21), November 2001.
- [AFDF⁺18] Rotem Arnon-Friedman, Frédéric Dupuis, Omar Fawzi, Renato Renner, and Thomas Vidick.
Practical device-independent quantum cryptography via entropy accumulation.
Nature Communications, 9(1), January 2018.
- [AFL21] Rotem Arnon-Friedman and Felix Leditzky.
Upper bounds on device-independent quantum key distribution rates and a revised Peres conjecture.
IEEE Transactions on Information Theory, 67(10):6606–6618, October 2021.
- [AFNB17] Mateus Araújo, Adrien Feix, Miguel Navascués, and Časlav Brukner.
A purification postulate for quantum mechanics with indefinite causal order.
Quantum, 1:10, April 2017.
- [AFOV08] Luigi Amico, Rosario Fazio, Andreas Osterloh, and Vlatko Vedral.
Entanglement in many-body systems.
Reviews of Modern Physics, 80(2):517–576, May 2008.

BIBLIOGRAPHY

- [AFRV19] Rotem Arnon-Friedman, Renato Renner, and Thomas Vidick.
Simple and tight device-independent security proofs.
SIAM Journal on Computing, 48(1):181–225, January 2019.
- [AFTS12] Rotem Arnon-Friedman and Amnon Ta-Shma.
Limits of privacy amplification against nonsignaling memory attacks.
Physical Review A, 86(6), December 2012.
- [AGM06] Antonio Acín, Nicolas Gisin, and Lluís Masanes.
From Bell’s theorem to secure quantum key distribution.
Physical Review Letters, 97(12), September 2006.
- [AH09a] R. Augusiak and P. Horodecki.
W-like bound entangled states and secure key distillation.
EPL (Europhysics Letters), 85(5):50001, March 2009.
- [AH09b] Remigiusz Augusiak and Paweł Horodecki.
Multipartite secret key distillation and bound entanglement.
Physical Review A, 80(4), October 2009.
- [Ahl06] R. Ahlswede.
On Concepts of Performance Parameters for Channels, pages 639–663.
Springer Berlin Heidelberg, Berlin, Heidelberg, 2006.
- [AML16] Koji Azuma, Akihiro Mizutani, and Hoi-Kwong Lo.
Fundamental rate-loss trade-off for the quantum internet.
Nature Communications, 7(1), November 2016.
- [AMP06] Antonio Acín, Serge Massar, and Stefano Pironio.
Efficient quantum key distribution secure against no-signalling eavesdroppers.
New Journal of Physics, 8(8):126–126, August 2006.
- [AQB⁺13] Mateus Araújo, Marco Túlio Quintino, Costantino Budroni, Marcelo Terra Cunha,
and Adán Cabello.
All noncontextuality inequalities for the n-cycle scenario.
Physical Review A, 88(2), August 2013.
- [Ard92] M. Ardehali.
Bell inequalities with a magnitude of violation that grows exponentially with the
number of particles.
Physical Review A, 46(9):5375–5378, November 1992.
- [ATL15] Koji Azuma, Kiyoshi Tamaki, and Hoi-Kwong Lo.

- All-photonic quantum repeaters.
Nature Communications, 6(1), April 2015.
- [BA17] Stefan Bäuml and Koji Azuma.
Fundamental limitation on quantum broadcast networks.
Quantum Science and Technology, 2(2):024004, May 2017.
- [Bar07] Jonathan Barrett.
Information processing in generalized probabilistic theories.
Physical Review A, 75(3), March 2007.
- [Bau21] Craig P. Bauer.
Secret History.
Chapman and Hall/CRC, February 2021.
- [BB84] Charles H. Bennett and Gilles Brassard.
Quantum cryptography: Public key distribution and coin tossing.
In *International Conference on Computer System and Signal Processing, IEEE, 1984*, pages 175–179, 1984.
- [BBP⁺96] Charles H. Bennett, Gilles Brassard, Sandu Popescu, Benjamin Schumacher, John A. Smolin, and William K. Wootters.
Purification of noisy entanglement and faithful teleportation via noisy channels.
Physical Review Letters, 76(5):722–725, January 1996.
- [BCHW15] Stefan Bäuml, Matthias Christandl, Karol Horodecki, and Andreas Winter.
Limitations on quantum key repeaters.
Nature Communications, 6(1), April 2015.
- [BCK12] Jonathan Barrett, Roger Colbeck, and Adrian Kent.
Unconditionally secure device-independent quantum key distribution with only two devices.
Physical Review A, 86(6), December 2012.
- [BCP⁺14] Nicolas Brunner, Daniel Cavalcanti, Stefano Pironio, Valerio Scarani, and Stephanie Wehner.
Bell nonlocality.
Reviews of Modern Physics, 86(2):419–478, April 2014.
- [BCSS11] Nicolas Brunner, Daniel Cavalcanti, Alejo Salles, and Paul Skrzypczyk.
Bound nonlocality and activation.
Physical Review Letters, 106(2), January 2011.

- [BD10] Francesco Buscemi and Nilanjana Datta.
The quantum capacity of channels with arbitrarily correlated noise.
IEEE Transactions on Information Theory, 56(3):1447–1460, March 2010.
- [BDCZ98] H.-J. Briegel, W. Dür, J. I. Cirac, and P. Zoller.
Quantum repeaters: The role of imperfect local operations in quantum communication.
Physical Review Letters, 81(26):5932–5935, December 1998.
- [BDLT08] Howard Barnum, Oscar C.O. Dahlsten, Matthew Leifer, and Ben Toner.
Nonclassicality without entanglement enables bit commitment.
In *2008 IEEE Information Theory Workshop*. IEEE, May 2008.
- [BDSW96] Charles H. Bennett, David P. DiVincenzo, John A. Smolin, and William K. Wootters.
Mixed-state entanglement and quantum error correction.
Physical Review A, 54(5):3824–3851, November 1996.
- [BDW18] Stefan Bäuml, Siddhartha Das, and Mark M. Wilde.
Fundamental limits on the capacities of bipartite quantum interactions.
Physical Review Letters, 121(25), December 2018.
- [Ben92] Charles H. Bennett.
Quantum cryptography using any two nonorthogonal states.
Physical Review Letters, 68(21):3121–3124, May 1992.
- [Ber17] Wolfgang Bertram.
An essay on the completion of quantum theory. i: General setting, 2017.
arXiv:1711.08643.
- [Ber18] Wolfgang Bertram.
An essay on the completion of quantum theory. ii: Unitary time evolution, 2018.
arXiv:1807.04650.
- [BHH⁺14] Piotr Badziąg, Karol Horodecki, Michał Horodecki, Justin Jenkinson, and Stanisław J. Szarek.
Bound entangled states with extremal properties.
Physical Review A, 90(1), July 2014.
- [BHK05] Jonathan Barrett, Lucien Hardy, and Adrian Kent.
No signaling and quantum key distribution.
Physical Review Letters, 95(1), June 2005.

- [BHTW10] Kamil Brádler, Patrick Hayden, Dave Touchette, and Mark M. Wilde.
Trade-off capacities of the quantum Hadamard channels.
Physical Review A, 81(6), June 2010.
- [BHZ05] Roman V. Buniy, Stephen D.H. Hsu, and A. Zee.
Is Hilbert space discrete?
Physics Letters B, 630(1-2):68–72, December 2005.
- [BK93] A V Belinskiĭ and D N Klyshko.
Interference of light and Bell's theorem.
Physics-Uspekhi, 36(8):653–693, August 1993.
- [BLM⁺05] Jonathan Barrett, Noah Linden, Serge Massar, Stefano Pironio, Sandu Popescu,
and David Roberts.
Nonlocal correlations as an information-theoretic resource.
Physical Review A, 71(2), February 2005.
- [BLS18] Howard Barnum, Ciarán M. Lee, and John H. Selby.
Oracles and query lower bounds in generalised probabilistic theories.
Foundations of Physics, 48(8):954–981, July 2018.
- [BLSS17] Howard Barnum, Ciarán Lee, Carlo Scandolo, and John Selby.
Ruling out higher-order interference from purity principles.
Entropy, 19(6):253, June 2017.
- [Blu81] Manuel Blum.
Coin flipping by telephone.
In *Advances in Cryptology: A Report on CRYPTO 81, IEEE Workshop on Communications Security*, pages 11–15, 1981.
- [BMU14] Howard Barnum, Markus P Müller, and Cozmin Ududec.
Higher-order interference and single-system postulates characterizing quantum theory.
New Journal of Physics, 16(12):123029, December 2014.
- [BOHL⁺05] Michael Ben-Or, Michał Horodecki, Debbie W. Leung, Dominic Mayers, and
Jonathan Oppenheim.
The universal composable security of quantum key distribution.
In *Theory of Cryptography*, pages 386–406. Springer Berlin Heidelberg, 2005.
- [BOM04] Michael Ben-Or and Dominic Mayers.
General security definition and compossibility for quantum & classical protocols,
2004.

- arXiv:quant-ph/0409062.
- [BP12] Samuel L. Braunstein and Stefano Pirandola.
Side-channel-free quantum key distribution.
Physical Review Letters, 108(13), March 2012.
- [BPR⁺00] Charles H. Bennett, Sandu Popescu, Daniel Rohrlich, John A. Smolin, and
Ashish V. Thapliyal.
Exact and asymptotic measures of multipartite pure-state entanglement.
Physical Review A, 63(1), December 2000.
- [BR17] Agung Budiyono and Daniel Rohrlich.
Quantum mechanics as classical statistical mechanics with an ontic extension and
an epistemic restriction.
Nature Communications, 8(1), November 2017.
- [BRA⁺19] C. E. Bradley, J. Randall, M. H. Aboeih, R. C. Berrevoets, M. J. Degen, M. A.
Bakker, M. Markham, D. J. Twitchen, and T. H. Taminiau.
A ten-qubit solid-state spin register with quantum memory up to one minute.
Physical Review X, 9(3), September 2019.
- [BRG⁺16] Fernando G. S. L. Brandão, Ravishankar Ramanathan, Andrzej Grudka, Karol
Horodecki, Michał Horodecki, Paweł Horodecki, Tomasz Szarek, and Hanna
Wojewódka.
Realistic noise-tolerant randomness amplification using finite number of devices.
Nature Communications, 7(1), April 2016.
- [BRPB13] Georg T. Becker, Francesco Regazzoni, Christof Paar, and Wayne P. Burleson.
Stealthy dopant-level hardware trojans.
In *Cryptographic Hardware and Embedded Systems - CHES 2013*, pages 197–214.
Springer Berlin Heidelberg, 2013.
- [Cab00] Adan Cabello.
Multiparty key distribution and secret sharing based on entanglement swapping,
2000.
arXiv:quant-ph/0009025.
- [CAL19] Marcos Curty, Koji Azuma, and Hoi-Kwong Lo.
Simple security proof of twin-field type quantum key distribution protocol.
npj Quantum Information, 5(1), July 2019.
- [Can01] R. Canetti.
Universally composable security: a new paradigm for cryptographic protocols.

- In *Proceedings 42nd IEEE Symposium on Foundations of Computer Science*, pages 136–145. IEEE, 2001.
- [CBH03] Rob Clifton, Jeffrey Bub, and Hans Halvorson.
Characterizing quantum theory in terms of information-theoretic constraints.
Foundations of Physics, 33(11):1561–1591, 2003.
- [CCL11] Wei Cui, Eric Chitambar, and Hoi-Kwong Lo.
Randomly distilling W-class states into general configurations of two-party entanglement.
Physical Review A, 84(5), November 2011.
- [CDP10] Giulio Chiribella, Giacomo Mauro D’Ariano, and Paolo Perinotti.
Probabilistic theories with purification.
Physical Review A, 81(6), June 2010.
- [CDP11] Giulio Chiribella, Giacomo Mauro D’Ariano, and Paolo Perinotti.
Informational derivation of quantum theory.
Physical Review A, 84(1), July 2011.
- [CDPV13] Giulio Chiribella, Giacomo Mauro D’Ariano, Paolo Perinotti, and Benoit Valiron.
Quantum computations without definite causal structure.
Physical Review A, 88(2), August 2013.
- [CEH⁺07] Matthias Christandl, Artur Ekert, Michał Horodecki, Paweł Horodecki, Jonathan Oppenheim, and Renato Renner.
Unifying classical and quantum key distillation.
In Salil P. Vadhan, editor, *Theory of Cryptography*, pages 456–478, Berlin, Heidelberg, 2007. Springer Berlin Heidelberg.
- [CF17] Matthias Christandl and Roberto Ferrara.
Private states, quantum data hiding, and the swapping of perfect secrecy.
Physical Review Letters, 119(22), November 2017.
- [CFH21] Matthias Christandl, Roberto Ferrara, and Karol Horodecki.
Upper bounds on device-independent quantum key distribution.
Physical Review Letters, 126(16), April 2021.
- [CHHH17] L. Czekaj, M. Horodecki, P. Horodecki, and R. Horodecki.
Information content of systems as a physical principle.
Physical Review A, 95(2), February 2017.
- [CHL10] Wei Cui, Wolfram Helwig, and Hoi-Kwong Lo.

- Bounds on probability of transformations between multipartite pure states.
Physical Review A, 81(1), January 2010.
- [Chr02] M. Christandl.
The quantum analog to intrinsic information.
Diploma Thesis, Institute for Theoretical Computer Science, ETH Zurich, 2002.
- [Cir80] B. S. Cirel'son.
Quantum generalizations of Bell's inequality.
Letters in Mathematical Physics, 4(2):93–100, March 1980.
- [CK78] I. Csiszar and J. Korner.
Broadcast channels with confidential messages.
IEEE Transactions on Information Theory, 24(3):339–348, May 1978.
- [CL05] Kai Chen and Hoi-Kwong Lo.
Conference key agreement and quantum sharing of classical secrets with noisy GHZ states.
In *Proceedings. International Symposium on Information Theory, 2005. ISIT 2005*.
IEEE, 2005.
- [CLL⁺09] Teng-Yun Chen, Hao Liang, Yang Liu, Wen-Qi Cai, Lei Ju, Wei-Yue Liu, Jian Wang, Hao Yin, Kai Chen, Zeng-Bing Chen, Cheng-Zhi Peng, and Jian-Wei Pan.
Field test of a practical secure communication network with decoy-state quantum cryptography.
Optics Express, 17(8):6540, April 2009.
- [CMG22] Giacomo Carrara, Gláucia Murta, and Federico Grasselli.
Overcoming fundamental bounds on quantum conference key agreement, 2022.
- [CMH17] Matthias Christandl and Alexander Müller-Hermes.
Relative entropy bounds on quantum, private and repeater capacities.
Communications in Mathematical Physics, 353(2):821–852, May 2017.
- [CMS02] N. J. Cerf, S. Massar, and S. Schneider.
Multipartite classical and quantum secrecy monotones.
Physical Review A, 66(4), October 2002.
- [Coe14] Bob Coecke.
Terminality implies non-signalling.
Electronic Proceedings in Theoretical Computer Science, 172:27–35, dec 2014.
arXiv:1405.3681.

- [CR12] Roger Colbeck and Renato Renner.
Free randomness can be amplified.
Nature Physics, 8(6):450–453, May 2012.
- [CS15] Giulio Chiribella and Carlo Maria Scandolo.
Entanglement and thermodynamics in general probabilistic theories.
New Journal of Physics, 17(10):103027, October 2015.
- [CS17] Giulio Chiribella and Carlo Maria Scandolo.
Microcanonical thermodynamics in general physical theories.
New Journal of Physics, 19(12):123043, December 2017.
- [CST18] Bob Coecke, John Selby, and Sean Tull.
Two roads to classicality.
Electronic Proceedings in Theoretical Computer Science, 266:104–118, February 2018.
- [CT09] Matthias Christandl and Ben Toner.
Finite de Finetti theorem for conditional probability distributions describing physical theories.
Journal of Mathematical Physics, 50(4):042104, April 2009.
- [CW04] Matthias Christandl and Andreas Winter.
“Squashed entanglement”: An additive entanglement measure.
Journal of Mathematical Physics, 45(3):829–840, March 2004.
- [CYW⁺19] Chaohan Cui, Zhen-Qiang Yin, Rong Wang, Wei Chen, Shuang Wang, Guang-Can Guo, and Zheng-Fu Han.
Twin-field quantum key distribution without phase postselection.
Physical Review Applied, 11(3), March 2019.
- [CZC⁺21] Yu-Ao Chen, Qiang Zhang, Teng-Yun Chen, Wen-Qi Cai, Sheng-Kai Liao, Jun Zhang, Kai Chen, Juan Yin, Ji-Gang Ren, Zhu Chen, Sheng-Long Han, Qing Yu, Ken Liang, Fei Zhou, Xiao Yuan, Mei-Sheng Zhao, Tian-Yin Wang, Xiao Jiang, Liang Zhang, Wei-Yue Liu, Yang Li, Qi Shen, Yuan Cao, Chao-Yang Lu, Rong Shu, Jian-Yu Wang, Li Li, Nai-Le Liu, Feihu Xu, Xiang-Bin Wang, Cheng-Zhi Peng, and Jian-Wei Pan.
An integrated space-to-ground quantum communication network over 4, 600 kilometres.
Nature, 589(7841):214–219, January 2021.
- [Das18] Siddhartha Das.

- Bipartite Quantum Interactions: Entangling and Information Processing Abilities.*
PhD thesis, Louisiana State University, October 2018.
arXiv:1901.05895.
- [Dat09a] Nilanjana Datta.
Max-relative entropy of entanglement, alias Log robustness.
International Journal of Quantum Information, 07(02):475–491, March 2009.
- [Dat09b] Nilanjana Datta.
Min- and max-relative entropies and a new entanglement monotone.
IEEE Transactions on Information Theory, 55(6):2816–2826, June 2009.
- [DB11] Borivoje Dakić and Časlav Brukner.
Quantum Theory and Beyond: Is Entanglement Special?, pages 365–392.
Cambridge University Press, 2011.
In “Deep Beauty: Understanding the Quantum World through Mathematical Innovation”, Editor: Hans Halvorson.
- [DBCZ99] W. Dür, H.-J. Briegel, J. I. Cirac, and P. Zoller.
Quantum repeaters based on entanglement purification.
Physical Review A, 59(1):169–181, January 1999.
- [DBW20] Siddhartha Das, Stefan Bäuml, and Mark M. Wilde.
Entanglement and secret-key-agreement capacities of bipartite quantum interactions and read-only memory devices.
Physical Review A, 101(1), January 2020.
- [DBWH21] Siddhartha Das, Stefan Bäuml, Marek Winzewski, and Karol Horodecki.
Universal limitations on quantum key distribution over a network.
Physical Review X, 11(4), October 2021.
- [Dev05] I. Devetak.
The private classical capacity and quantum capacity of a quantum channel.
IEEE Transactions on Information Theory, 51(1):44–55, January 2005.
- [DKD18] Siddhartha Das, Sumeet Khatri, and Jonathan P. Dowling.
Robust quantum network architectures and topologies for entanglement distribution.
Physical Review A, 97(1), January 2018.
- [DKDD⁺11] K. Dobek, M. Karpiński, R. Demkowicz-Dobrzański, K. Banaszek, and P. Horodecki.
Experimental extraction of secure correlations from a noisy private state.

-
- Physical Review Letters*, 106(3), January 2011.
- [DM03] Jonathan P. Dowling and Gerard J. Milburn.
Quantum technology: the second quantum revolution.
Philosophical Transactions of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences, 361(1809):1655–1674, June 2003.
- [Doo18] John F. Dooley.
History of Cryptography and Cryptanalysis.
Springer International Publishing, 2018.
- [DSW18] Noah Davis, Maksim E. Shirokov, and Mark M. Wilde.
Energy-constrained two-way assisted private and quantum capacities of quantum channels.
Physical Review A, 97(6), June 2018.
- [DVC00] W. Dür, G. Vidal, and J. I. Cirac.
Three qubits can be entangled in two inequivalent ways.
Physical Review A, 62(6), November 2000.
- [DW04] I. Devetak and A. Winter.
Relating quantum privacy and quantum coherence: An operational approach.
Physical Review Letters, 93(8), August 2004.
- [DW05] Igor Devetak and Andreas Winter.
Distillation of secret key and entanglement from quantum states.
Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences, 461(2053):207–235, January 2005.
- [DW19] Siddhartha Das and Mark M. Wilde.
Quantum rebound capacity.
Physical Review A, 100(3), September 2019.
- [Eke91] Artur K. Ekert.
Quantum cryptography based on Bell’s theorem.
Physical Review Letters, 67(6):661–663, August 1991.
- [ELS06] Khaled Elbassioni, Zvi Lotker, and Raimund Seidel.
Upper bound on the number of vertices of polyhedra with 0,1-constraint matrices.
Information Processing Letters, 100(2):69–71, 2006.
- [EPR35] Albert Einstein, Boris Podolsky, and Nathan Rosen.
Can quantum-mechanical description of physical reality be considered complete?
Physical Review, 47(10):777–780, May 1935.

- [EPR92] Avshalom C. Elitzur, Sandu Popescu, and Daniel Rohrlich.
Quantum nonlocality for each pair in an ensemble.
Physics Letters A, 162(1):25–28, 1992.
- [FBJL⁺21] Máté Farkas, Maria Balanzó-Juandó, Karol Łukanowski, Jan Kołodyński, and Antonio Acín.
Bell nonlocality is not sufficient for the security of standard device-independent quantum key distribution protocols.
Physical Review Letters, 127(5), July 2021.
- [FL07] Ben Fortescue and Hoi-Kwong Lo.
Random bipartite entanglement from W and W -like states.
Physical Review Letters, 98(26), June 2007.
- [FL08] Ben Fortescue and Hoi-Kwong Lo.
Random-party entanglement distillation in multiparty states.
Physical Review A, 78(1), July 2008.
- [FvdG99] C.A. Fuchs and J. van de Graaf.
Cryptographic distinguishability measures for quantum-mechanical states.
IEEE Transactions on Information Theory, 45(4):1216–1227, 1999.
- [FYCC15] Yao Fu, Hua-Lei Yin, Teng-Yun Chen, and Zeng-Bing Chen.
Long-distance measurement-device-independent multiparty quantum communication.
Physical Review Letters, 114(9), March 2015.
- [GBP97] M. Grassl, Th. Beth, and T. Pellizzari.
Codes for the quantum erasure channel.
Physical Review A, 56(1):33–38, July 1997.
- [GEW16] K Goodenough, D Elkouss, and S Wehner.
Assessing the performance of quantum repeaters for all phase-insensitive gaussian bosonic channels.
New Journal of Physics, 18(6):063005, June 2016.
- [GHH⁺14] A. Grudka, K. Horodecki, M. Horodecki, P. Horodecki, R. Horodecki, P. Joshi, W. Kłobus, and A. Wójcik.
Quantifying contextuality.
Physical Review Letters, 112(12), March 2014.
- [GHZ89] Daniel M. Greenberger, Michael A. Horne, and Anton Zeilinger.
Going beyond Bell’s theorem.

- In *Bell's Theorem, Quantum Theory and Conceptions of the Universe*, pages 69–72. Springer Netherlands, 1989.
- [GK11] Abbas El Gamal and Young-Han Kim.
Network Information Theory.
Cambridge University Press, December 2011.
- [GKB19] Federico Grasselli, Hermann Kampermann, and Dagmar Bruß.
Conference key agreement with single-photon interference.
New Journal of Physics, 21(12):123002, December 2019.
- [Gle75] Andrew M. Gleason.
Measures on the closed subspaces of a Hilbert space.
In *The Logico-Algebraic Approach to Quantum Mechanics*, pages 123–133. Springer Netherlands, 1975.
- [GMR85] S Goldwasser, S Micali, and C Rackoff.
The knowledge complexity of interactive proof-systems.
In *Proceedings of the seventeenth annual ACM symposium on Theory of computing - STOC '85*. ACM Press, 1985.
- [GMT⁺13] Rodrigo Gallego, Lluís Masanes, Gonzalo De La Torre, Chirag Dhara, Leandro Aolita, and Antonio Acín.
Full randomness from arbitrarily deterministic events.
Nature Communications, 4(1), October 2013.
- [Goy08] Philip Goyal.
Information-geometric reconstruction of quantum theory.
Physical Review A, 78(5), November 2008.
- [GRTZ02] Nicolas Gisin, Grégoire Ribordy, Wolfgang Tittel, and Hugo Zbinden.
Quantum cryptography.
Reviews of Modern Physics, 74(1):145–195, March 2002.
- [GS18] Stefano Gogioso and Carlo Maria Scandolo.
Categorical probabilistic theories.
Electronic Proceedings in Theoretical Computer Science, 266:367–385, feb 2018.
arXiv:1701.08075.
- [GWAN12] Rodrigo Gallego, Lars Erik Würflinger, Antonio Acín, and Miguel Navascués.
Operational framework for nonlocality.
Physical Review Letters, 109(7), August 2012.

- [HA06] Paweł Horodecki and Remigiusz Augusiak.
Quantum states representing perfectly secure bits are always distillable.
Physical Review A, 74(1), July 2006.
- [Hän10] E. Hänggi.
Device-independent quantum key distribution.
PhD thesis, ETH Zürich, December 2010.
arXiv:1012.3878.
- [Har28] R. V. L. Hartley.
Transmission of information.
Bell System Technical Journal, 7(3):535–563, July 1928.
- [Har01] Lucien Hardy.
Quantum theory from five reasonable axioms, 2001.
arXiv:quant-ph/0101012.
- [Har05] Lucien Hardy.
Probability theories with dynamic causal structure: A new framework for quantum gravity, 2005.
arXiv:gr-qc/0509120.
- [Har07] Lucien Hardy.
Towards quantum gravity: a framework for probabilistic theories with non-fixed causal structure.
Journal of Physics A: Mathematical and Theoretical, 40(12):3081–3099, March 2007.
- [Har11] Lucien Hardy.
Reformulating and reconstructing quantum theory, 2011.
arXiv:.1104.2066.
- [HBD⁺15] B. Hensen, H. Bernien, A. E. Dréau, A. Reiserer, N. Kalb, M. S. Blok, J. Ruitenber, R. F. L. Vermeulen, R. N. Schouten, C. Abellán, W. Amaya, V. Pruneri, M. W. Mitchell, M. Markham, D. J. Twitchen, D. Elkouss, S. Wehner, T. H. Taminiau, and R. Hanson.
Loophole-free Bell inequality violation using electron spins separated by 1.3 kilometres.
Nature, 526(7575):682–686, October 2015.
- [HĆRS18] K. Horodecki, P. Őwikliński, A. Rutkowski, and M Studziński.
On distilling secure key from reducible private states and (non)existence of entangled key-undistillable states.

-
- New Journal of Physics*, 20(8):083021, August 2018.
- [HGJ⁺15] Karol Horodecki, Andrzej Grudka, Pankaj Joshi, Waldemar Kłobus, and Justyna Łodyga.
Axiomatic approach to contextuality and nonlocality.
Physical Review A, 92(3), September 2015.
- [HHH98] Michał Horodecki, Paweł Horodecki, and Ryszard Horodecki.
Mixed-state entanglement and distillation: Is there a “bound” entanglement in Nature?
Physical Review Letters, 80(24):5239–5242, June 1998.
- [HHH99] Michał Horodecki, Paweł Horodecki, and Ryszard Horodecki.
General teleportation channel, singlet fraction, and quasidistillation.
Physical Review A, 60(3):1888–1898, September 1999.
- [HHHH09] Ryszard Horodecki, Paweł Horodecki, Michał Horodecki, and Karol Horodecki.
Quantum entanglement.
Reviews of Modern Physics, 81(2):865–942, June 2009.
- [HHHO05] Karol Horodecki, Michał Horodecki, Paweł Horodecki, and Jonathan Oppenheim.
Secure key from bound entanglement.
Physical Review Letters, 94(16), April 2005.
- [HHHO09] Karol Horodecki, Michał Horodecki, Paweł Horodecki, and Jonathan Oppenheim.
General paradigm for distilling classical key from quantum states.
IEEE Transactions on Information Theory, 55(4):1898–1929, April 2009.
- [HŁPHH08] Karol Horodecki, Łukasz Pankowski, Michał Horodecki, and Paweł Horodecki.
Low-dimensional bound entanglement with one-way distillable cryptographic key.
IEEE Transactions on Information Theory, 54(6):2621–2625, June 2008.
- [Höh17] Philipp Höhn.
Quantum theory from rules on information acquisition.
Entropy, 19(3):98, March 2017.
- [Hor09] Karol Horodecki.
General paradigm for distilling classical key from quantum states - on quantum entanglement and security.
PhD thesis, University of Warsaw, 2009.
- [HR10] Esther Hänggi and Renato Renner.
Device-independent quantum key distribution with commuting measurements, 2010.

arXiv:1009.1833.

- [HRW10] Esther Hänggi, Renato Renner, and Stefan Wolf.
Quantum cryptography based solely on Bell's theorem.
In *Proceedings of Advances in Cryptology - EUROCRYPT, 2010*, pages 216–234, 2010.
- [HRW13] Esther Hänggi, Renato Renner, and Stefan Wolf.
The impossibility of non-signaling privacy amplification.
Theoretical Computer Science, 486:27–42, May 2013.
- [HSD15] Dipankar Home, Debashis Saha, and Siddhartha Das.
Multipartite Bell-type inequality by generalizing Wigner's argument.
Physical Review A, 91(1), January 2015.
- [HWD22] Karol Horodecki, Marek Winzewski, and Siddhartha Das.
Fundamental limitations on the device-independent quantum conference key agreement.
Physical Review A, 105(2), February 2022.
Corrected in erratum [[HWD23](#)].
- [HWD23] Karol Horodecki, Marek Winzewski, and Siddhartha Das.
Erratum: Fundamental limitations on the device-independent quantum conference key agreement [Phys. Rev. A 105, 022604 (2022)].
Phys. Rev. A, 107:029902, Feb 2023.
- [JGBB11] Peter Janotta, Christian Gogolin, Jonathan Barrett, and Nicolas Brunner.
Limits on nonlocal correlations from the structure of the local state space.
New Journal of Physics, 13(6):063024, June 2011.
- [Kau20] Eneet Kaur.
Limitations on Protecting Information Against Quantum Adversaries.
PhD thesis, LSU Doctoral Dissertations, 2020.
- [KGR05] B. Kraus, N. Gisin, and R. Renner.
Lower and upper bounds on the secret-key rate for quantum key distribution protocols using one-way classical communication.
Physical Review Letters, 95(8), August 2005.
- [KHC17] Aleks Kissinger, Matty Hoban, and Bob Coecke.
Equivalence of relativistic causal structure and process terminality.
arXiv preprint arXiv:1708.04118, 2017.

- [KHD22] Eneet Kaur, Karol Horodecki, and Siddhartha Das.
Upper bounds on device-independent quantum key distribution rates in static and dynamic scenarios.
Physical Review Applied, 18(5), November 2022.
- [Kim08] H. J. Kimble.
The quantum internet.
Nature, 453(7198):1023–1030, June 2008.
- [KT10] S. Kintaş and S. Turgut.
Transformations of W-type entangled states.
Journal of Mathematical Physics, 51(9):092202, September 2010.
- [Kur61] K. Kuratowski.
Introduction To Set Theory & Topology, volume 101 of *International series of monographs in pure and applied mathematics*.
PWN, Warsaw, 1961.
- [KWW20] Eneet Kaur, Mark M Wilde, and Andreas Winter.
Fundamental limits on key rates in device-independent quantum key distribution.
New Journal of Physics, 22(2):023039, February 2020.
- [LALS20] Felix Leditzky, Mohammad A. Alhejji, Joshua Levin, and Graeme Smith.
Playing games with multiple access channels.
Nature Communications, 11(1), March 2020.
- [Lam18] Ludovico Lami.
Non-classical correlations in quantum mechanics and beyond, 2018.
arXiv:1803.02902.
- [LC98] Hoi-Kwong Lo and H.F. Chau.
Why quantum bit commitment and ideal quantum coin tossing are impossible.
Physica D: Nonlinear Phenomena, 120(1-2):177–187, September 1998.
- [LCQ12] Hoi-Kwong Lo, Marcos Curty, and Bing Qi.
Measurement-device-independent quantum key distribution.
Physical Review Letters, 108(13), March 2012.
- [LL18] Jie Lin and Norbert Lütkenhaus.
Simple security analysis of phase-matching measurement-device-independent quantum key distribution.
Physical Review A, 98(4), October 2018.

- [LP17] Riccardo Laurenza and Stefano Pirandola.
General bounds for sender-receiver capacities in multipoint quantum communications.
Physical Review A, 96(3), September 2017.
- [LS16a] Ciarán M Lee and John H Selby.
Deriving Grover's lower bound from simple physical principles.
New Journal of Physics, 18(9):093047, September 2016.
- [LS16b] Ciarán M Lee and John H Selby.
Generalised phase kick-back: the structure of computational algorithms from physical principles.
New Journal of Physics, 18(3):033023, March 2016.
- [LS18] Ciarán M. Lee and John H. Selby.
A no-go theorem for theories that decohere to quantum mechanics.
Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences, 474(2214):20170732, June 2018.
- [LSW11] Yeong-Cherng Liang, Robert W. Spekkens, and Howard M. Wiseman.
Specker's parable of the overprotective seer: A road to contextuality, nonlocality and complementarity.
Physics Reports, 506(1-2):1–39, September 2011.
Corrected in erratum [\[LSW17\]](#).
- [LSW17] Yeong-Cherng Liang, Robert W. Spekkens, and Howard M. Wiseman.
Erratum to “specker's parable of the over-protective seer: A road to contextuality, nonlocality and complementarity” [phys. rep. 506 (2011) 1–39].
Physics Reports, 666:110–111, January 2017.
- [Luo22] Ming-Xing Luo.
Fully device-independent model on quantum networks.
Physical Review Research, 4(1), March 2022.
- [LWW⁺19] Hui Liu, Wenyuan Wang, Kejin Wei, Xiao-Tian Fang, Li Li, Nai-Le Liu, Hao Liang, Si-Jie Zhang, Weijun Zhang, Hao Li, Lixing You, Zhen Wang, Hoi-Kwong Lo, Teng-Yun Chen, Feihu Xu, and Jian-Wei Pan.
Experimental demonstration of high-rate measurement-device-independent quantum key distribution over asymmetric channels.
Physical Review Letters, 122(16), April 2019.
- [LYDS18] M. Lucamarini, Z. L. Yuan, J. F. Dynes, and A. J. Shields.

- Overcoming the rate–distance limit of quantum key distribution without quantum repeaters.
Nature, 557(7705):400–403, May 2018.
- [LZZ⁺22] Wen-Zhao Liu, Yu-Zhe Zhang, Yi-Zheng Zhen, Ming-Han Li, Yang Liu, Jingyun Fan, Feihu Xu, Qiang Zhang, and Jian-Wei Pan.
Toward a photonic demonstration of device-independent quantum key distribution.
Physical Review Letters, 129(5), July 2022.
- [Mak09] Vadim Makarov.
Controlling passively quenched single photon detectors by bright light.
New Journal of Physics, 11(6):065003, June 2009.
- [Mas06] Lluís Masanes.
Asymptotic violation of Bell inequalities and distillability.
Physical Review Letters, 97(5), August 2006.
- [Mas09] Lluís Masanes.
Universally composable privacy amplification from causality constraints.
Physical Review Letters, 102(14), April 2009.
- [MATN15] William J. Munro, Koji Azuma, Kiyoshi Tamaki, and Kae Nemoto.
Inside quantum repeaters.
IEEE Journal of Selected Topics in Quantum Electronics, 21(3):78–90, May 2015.
- [Mau93] U.M. Maurer.
Secret key agreement by public discussion from common information.
IEEE Transactions on Information Theory, 39(3):733–742, May 1993.
- [May97] Dominic Mayers.
Unconditionally secure quantum bit commitment is impossible.
Physical Review Letters, 78(17):3414–3417, April 1997.
- [McM70] P. McMullen.
The maximum numbers of faces of a convex polytope.
Mathematika, 17(2):179–184, 1970.
- [Mer90] N. David Mermin.
Extreme quantum entanglement in a superposition of macroscopically distinct states.
Physical Review Letters, 65(15):1838–1840, October 1990.
- [MGKB20] Gláucia Murta, Federico Grasselli, Hermann Kampermann, and Dagmar Bruß.

- Quantum conference key agreement: A review.
Advanced Quantum Technologies, 3(11):2000025, September 2020.
- [MGM19] Lluís Masanes, Thomas D. Galley, and Markus P. Müller.
The measurement postulates of quantum mechanics are operationally redundant.
Nature Communications, 10(1), March 2019.
- [MGP15] Piotr Mironowicz, Rodrigo Gallego, and Marcin Pawłowski.
Robust amplification of santha-vazirani sources with three devices.
Physical Review A, 91(3), March 2015.
- [Mil82] F. Miller.
Telegraphic Code to Insure Privacy and Secrecy in the Transmission of Telegrams.
C.M. Cornwell, 1882.
- [MLDS⁺13] Martin Müller-Lennert, Frédéric Dupuis, Oleg Szehr, Serge Fehr, and Marco Tomamichel.
On quantum Rényi entropies: A new generalization and some properties.
Journal of Mathematical Physics, 54(12):122203, December 2013.
- [MLK⁺16] Sreraman Muralidharan, Linshu Li, Jungsang Kim, Norbert Lütkenhaus, Mikhail D. Lukin, and Liang Jiang.
Optimal architectures for long distance quantum communication.
Scientific Reports, 6(1), February 2016.
- [MM11] Lluís Masanes and Markus P Müller.
A derivation of quantum theory from physical requirements.
New Journal of Physics, 13(6):063001, June 2011.
- [MPA11] Lluís Masanes, Stefano Pironio, and Antonio Acín.
Secure device-independent quantum key distribution with causally independent measurement devices.
Nature Communications, 2(1), March 2011.
- [MPR⁺19] M. Minder, M. Pittaluga, G. L. Roberts, M. Lucamarini, J. F. Dynes, Z. L. Yuan, and A. J. Shields.
Experimental quantum key distribution beyond the repeaterless secret key capacity.
Nature Photonics, 13(5):334–338, March 2019.
- [MRC⁺14] Lluís Masanes, Renato Renner, Matthias Christandl, Andreas Winter, and Jonathan Barrett.
Full security of quantum key distribution from no-signaling constraints.
IEEE Transactions on Information Theory, 60(8):4973–4986, August 2014.

- [Mül09] Markus Müller.
Does probability become fuzzy in small regions of spacetime?
Physics Letters B, 673(2):166–167, March 2009.
- [Mül21] Markus Müller.
Probabilistic theories and reconstructions of quantum theory.
SciPost Physics Lecture Notes, March 2021.
- [MW97] U. Maurer and S. Wolf.
The intrinsic conditional mutual information and perfect secrecy.
In *Proceedings of IEEE International Symposium on Information Theory*. IEEE,
29 June – 4 July 1997.
- [MW99] U.M. Maurer and S. Wolf.
Unconditionally secure key agreement and the intrinsic conditional information.
IEEE Transactions on Information Theory, 45(2):499–514, March 1999.
- [MW00] Ueli Maurer and Stefan Wolf.
Information-theoretic key agreement: From weak to strong secrecy for free.
In *Advances in Cryptology — EUROCRYPT 2000*, pages 351–368. Springer Berlin
Heidelberg, 2000.
- [MY04] Dominic Mayers and Andrew Yao.
Self testing quantum apparatus.
Quantum Info. Comput., 4(4):273286, July 2004.
- [MZZ18] Xiongfeng Ma, Pei Zeng, and Hongyi Zhou.
Phase-matching quantum key distribution.
Physical Review X, 8(3), August 2018.
- [Nak20] Kenji Nakahira.
Derivation of quantum theory with superselection rules.
Physical Review A, 101(2), February 2020.
- [NC00] M. A. Nielsen and I. L. Chuang.
Quantum Computation and Quantum Information.
Cambridge University Press, Cambridge, 2000.
- [NDN⁺22] D. P. Nadlinger, P. Drmota, B. C. Nichol, G. Araneda, D. Main, R. Srinivas, D. M.
Lucas, C. J. Ballance, K. Ivanov, E. Y.-Z. Tan, P. Sekatski, R. L. Urbanke,
R. Renner, N. Sangouard, and J.-D. Bancal.
Experimental quantum key distribution certified by Bell's theorem.
Nature, 607(7920):682–686, July 2022.

BIBLIOGRAPHY

- [Nyq24] H. Nyquist.
Certain factors affecting telegraph speed.
Bell System Technical Journal, 3(2):324–346, April 1924.
- [OCB12] Ognyan Oreshkov, Fabio Costa, and Časlav Brukner.
Quantum correlations with no causal order.
Nature Communications, 3(1), October 2012.
- [OLLP19] Carlo Ottaviani, Cosmo Lupo, Riccardo Laurenza, and Stefano Pirandola.
Modular network for high-rate quantum conferencing.
Communications Physics, 2(1), September 2019.
- [PAB⁺09] Stefano Pironio, Antonio Acín, Nicolas Brunner, Nicolas Gisin, Serge Massar, and Valerio Scarani.
Device-independent quantum key distribution secure against collective attacks.
New Journal of Physics, 11(4):045021, April 2009.
- [PAB⁺20] S. Pirandola, U. L. Andersen, L. Banchi, M. Berta, D. Bunandar, R. Colbeck, D. Englund, T. Gehring, C. Lupo, C. Ottaviani, J. L. Pereira, M. Razavi, J. Shamsul Shaari, M. Tomamichel, V. C. Usenko, G. Vallone, P. Villoresi, and P. Wallden.
Advances in quantum cryptography.
Advances in Optics and Photonics, 12(4):1012, December 2020.
- [Pal20] T. N. Palmer.
Discretization of the Bloch sphere, fractal invariant sets and Bell’s theorem.
Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences, 476(2236):20190350, April 2020.
- [PBL⁺18] Stefano Pirandola, Samuel L Braunstein, Riccardo Laurenza, Carlo Ottaviani, Thomas P W Cope, Gaetana Spedalieri, and Leonardo Banchi.
Theory of channel simulation and bounds for private communication.
Quantum Science and Technology, 3(3):035009, May 2018.
- [PDK⁺19] S. J. Pauka, K. Das, R. Kalra, A. Moini, Y. Yang, M. Trainer, A. Bousquet, C. Cantaloube, N. Dick, G. C. Gardner, M. J. Manfra, and D. J. Reilly.
A cryogenic interface for controlling many qubits, 2019.
arXiv:1912.01299.
- [PdV22] Carlos Palazuelos and Julio I. de Vicente.
Genuine multipartite entanglement of quantum states in the multiple-copy scenario.
Quantum, 6:735, June 2022.

- [Per02] Asher Peres.
Karl Popper and the Copenhagen interpretation.
Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics, 33(1):23–34, March 2002.
- [PGPBL09] Stefano Pirandola, Raul García-Patrón, Samuel L. Braunstein, and Seth Lloyd.
Direct and reverse secret-key capacities of a quantum channel.
Physical Review Letters, 102(5), February 2009.
- [PH10] Łukasz Pankowski and Michał Horodecki.
Low-dimensional quite noisy bound entanglement with a cryptographic key.
Journal of Physics A: Mathematical and Theoretical, 44(3):035301, December 2010.
- [Pir05] Stefano Pironio.
Lifting Bell inequalities.
Journal of Mathematical Physics, 46(6):062112, June 2005.
- [Pir19] Stefano Pirandola.
End-to-end capacities of a quantum communication network.
Communications Physics, 2(1), May 2019.
- [PKBW23] Aby Philip, Eneet Kaur, Peter Bierhorst, and Mark M. Wilde.
Multipartite intrinsic non-locality and device-independent conference key agreement.
Quantum, 7:898, January 2023.
- [Plá21] Martin Plávala.
General probabilistic theories: An introduction, 2021.
arXiv:2103.07469.
- [PLG⁺19] Ignatius William Primaatmaja, Emilien Lavie, Koon Tong Goh, Chao Wang, and Charles Ci Wen Lim.
Versatile security analysis of measurement-device-independent quantum key distribution.
Physical Review A, 99(6), June 2019.
- [PLOB17] Stefano Pirandola, Riccardo Laurenza, Carlo Ottaviani, and Leonardo Banchi.
Fundamental limits of repeaterless quantum communications.
Nature Communications, 8(1), April 2017.
- [POS⁺15] Stefano Pirandola, Carlo Ottaviani, Gaetana Spedalieri, Christian Weedbrook, Samuel L. Braunstein, Seth Lloyd, Tobias Gehring, Christian S. Jacobsen, and Ulrik L. Andersen.

- High-rate measurement-device-independent quantum cryptography.
Nature Photonics, 9(6):397–402, May 2015.
- [PPK⁺09] Marcin Pawłowski, Tomasz Paterek, Dagomir Kaszlikowski, Valerio Scarani, Andreas Winter, and Marek Żukowski.
Information causality as a physical principle.
Nature, 461(7267):1101–1104, October 2009.
- [PR94] Sandu Popescu and Daniel Rohrlich.
Quantum nonlocality as an axiom.
Foundations of Physics, 24(3):379–385, March 1994.
- [Pre18] John Preskill.
Quantum computing in the NISQ era and beyond.
Quantum, 2:79, August 2018.
- [PW13] Corsin Pfister and Stephanie Wehner.
An information-theoretic principle implies that any discrete physical theory is classical.
Nature Communications, 4(1), May 2013.
- [Ras85] Peter Rastall.
Locality, Bell's theorem, and quantum mechanics.
Foundations of Physics, 15(9):963–972, September 1985.
- [RBH⁺16] Ravishankar Ramanathan, Fernando G. S. L. Brandão, Karol Horodecki, Michał Horodecki, Paweł Horodecki, and Hanna Wojewódka.
Randomness amplification under minimal fundamental assumptions on the devices.
Physical Review Letters, 117(23), November 2016.
- [Ren05] Renato Renner.
Security of Quantum Key Distribution.
PhD thesis, ETH Zürich, December 2005.
arXiv:quant-ph/0512258.
- [RKB⁺18] Luca Rigovacca, Go Kato, Stefan Bäuml, M S Kim, W J Munro, and Koji Azuma.
Versatile relative entropy bounds for quantum networks.
New Journal of Physics, 20(1):013033, January 2018.
- [RMW18] Jérémy Ribeiro, Gláucia Murta, and Stephanie Wehner.
Fully device-independent conference key agreement.
Physical Review A, 97(2), February 2018.

- [RP10] Tim C. Ralph and Geoff J. Pryde.
Optical quantum computation.
In *Progress in Optics*, pages 209–269. Elsevier, 2010.
- [RSW03] R. Renner, J. Skripsky, and S. Wolf.
A new measure for conditional mutual information and its properties.
In *IEEE International Symposium on Information Theory, 2003. Proceedings.*, pages 259–259, 2003.
- [RW03] Renato Renner and Stefan Wolf.
New bounds in secret-key agreement: The gap between formation and secrecy extraction.
In Eli Biham, editor, *Advances in Cryptology — EUROCRYPT 2003*, pages 562–577, Berlin, Heidelberg, 2003. Springer Berlin Heidelberg.
- [SC17] John Selby and Bob Coecke.
Leaks: Quantum, classical, intermediate and more.
Entropy, 19(4):174, April 2017.
- [Sch35] E. Schrödinger.
Die gegenwärtige Situation in der Quantenmechanik.
Naturwissenschaften, 23:807–812, 1935.
- [SdVSK17] C. Spee, J. I. de Vicente, D. Sauerwein, and B. Kraus.
Entangled pure state transformations via local operations assisted by finitely many rounds of classical communication.
Physical Review Letters, 118(4), January 2017.
- [SG01] Valerio Scarani and Nicolas Gisin.
Quantum key distribution between N partners: Optimal eavesdropping and Bell’s inequalities.
Physical Review A, 65(1), December 2001.
- [SGB⁺06] Valerio Scarani, Nicolas Gisin, Nicolas Brunner, Lluís Masanes, Sergi Pino, and Antonio Acín.
Secrecy extraction from no-signaling correlations.
Physical Review A, 74(4), October 2006.
- [Sha61] Claude E. Shannon.
Two-way communication channels.
In *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability, Volume 1: Contributions to the Theory of Statistics*, pages 611–644, Berkeley, California, 1961. University of California Press.

BIBLIOGRAPHY

- [Sha01] C. E. Shannon.
A mathematical theory of communication.
SIGMOBILE Mob. Comput. Commun. Rev., 5(1):355, January 2001.
- [Sho94] Peter W. Shor.
Algorithms for quantum computation: Discrete logarithms and factoring.
In *Proceedings of the 35th Annual Symposium on Foundations of Computer Science*, pages 124–134, Los Alamitos, California, 1994. IEEE Computer Society Press.
- [SME20] A. Streltsov, C. Meignant, and J. Eisert.
Rates of multipartite entanglement transformations.
Physical Review Letters, 125(8), August 2020.
- [Smo06] Lee Smolin.
Could quantum mechanics be an approximation to another theory?, 2006.
arXiv:quant-ph/0609109.
- [SNS⁺21] Takahiko Satoh, Shota Nagayama, Shigeya Suzuki, Takaaki Matsuo, Michal Hajdušek, and Rodney Van Meter.
Attacking the quantum Internet.
IEEE Transactions on Quantum Engineering, 2:1–17, 2021.
- [Spe90] Ernst Specker.
Die Logik nicht gleichzeitig entscheidbarer Aussagen.
In *Ernst Specker Selecta*, pages 175–182. Springer, 1990.
- [SS02] Michael Seevinck and George Svetlichny.
Bell-type inequalities for partial separability in N-particle systems and quantum mechanical violations.
Physical Review Letters, 89(6), July 2002.
- [SS18] Jamie Sikora and John Selby.
Simple proof of the impossibility of bit commitment in generalized probabilistic theories using cone programming.
Physical Review A, 97(4), April 2018.
- [SSC21] John H. Selby, Carlo Maria Scandolo, and Bob Coecke.
Reconstructing quantum theory from diagrammatic postulates.
Quantum, 5:445, April 2021.
- [Sto32] M. H. Stone.
On one-parameter unitary groups in Hilbert space.
The Annals of Mathematics, 33(3):643, July 1932.

- [Sve87] George Svetlichny.
Distinguishing three-body from two-body nonseparability by a Bell-type inequality.
Physical Review D, 35(10):3066–3069, May 1987.
- [SVW05] John A. Smolin, Frank Verstraete, and Andreas Winter.
Entanglement of assistance and multipartite state distillation.
Physical Review A, 72(5), November 2005.
- [SW16] Benno Salwey and Stefan Wolf.
Stronger attacks on causality-based key agreement.
In *2016 IEEE International Symposium on Information Theory (ISIT)*. IEEE,
July 2016.
- [SWRH20] Omer Sakarya, Marek Winczewski, Adam Rutkowski, and Karol Horodecki.
Hybrid quantum network design against unauthorized secret-key generation, and
its memory cost.
Physical Review Research, 2(4), October 2020.
- [TGW14a] Masahiro Takeoka, Saikat Guha, and Mark M. Wilde.
Fundamental rate-loss tradeoff for optical quantum key distribution.
Nature Communications, 5(1), October 2014.
- [TGW14b] Masahiro Takeoka, Saikat Guha, and Mark M. Wilde.
The squashed entanglement of a quantum channel.
IEEE Transactions on Information Theory, 60(8):4987–4998, August 2014.
- [TLWL18] Kiyoshi Tamaki, Hoi-Kwong Lo, Wenyuan Wang, and Marco Lucamarini.
Information theoretic security of quantum key distribution overcoming the re-
peaterless secret key capacity bound, 2018.
arXiv:1805.05511.
- [TR19] Si-Hui Tan and Peter P. Rohde.
The resurgence of the linear optics quantum interferometer — recent advances &
applications.
Reviews in Physics, 4:100030, November 2019.
- [TSW16] Masahiro Takeoka, Kaushik P. Seshadreesan, and Mark M. Wilde.
Unconstrained distillation capacities of a pure-loss bosonic broadcast channel.
In *2016 IEEE International Symposium on Information Theory (ISIT)*. IEEE,
July 2016.
- [Tul20] Sean Tull.
A categorical reconstruction of quantum theory.

- Logical Methods in Computer Science ; Volume 16*, pages Issue 1 ; 1860–5974, 2020.
- [Ume62] Hisaharu Umegaki.
Conditional expectation in an operator algebra. IV. entropy and information.
Kodai Mathematical Journal, 14(2), January 1962.
- [VBD⁺15] Giuseppe Vallone, Davide Bacco, Daniele Dequal, Simone Gaiarin, Vincenza Luceri, Giuseppe Bianco, and Paolo Villoresi.
Experimental satellite quantum communications.
Physical Review Letters, 115(4), July 2015.
- [VC15] Péter Vrana and Matthias Christandl.
Asymptotic entanglement transformation between W and GHZ states.
Journal of Mathematical Physics, 56(2):022204, 2015.
- [VC19] Peter Vrana and Matthias Christandl.
Distillation of Greenberger–Horne–Zeilinger states by combinatorial methods.
IEEE Transactions on Information Theory, 65(9):5945–5958, September 2019.
- [vDGG05] W. van Dam, R.D. Gill, and P.D. Grunwald.
The statistical strength of nonlocality proofs.
IEEE Transactions on Information Theory, 51(8):2812–2835, August 2005.
- [VDM02] Frank Verstraete, Jeroen Dehaene, and Bart De Moor.
On the geometry of entangled states.
Journal of Modern Optics, 49(8):1277–1287, July 2002.
- [vdW19] John van de Wetering.
An effect-theoretic reconstruction of quantum theory.
Compositionality, 1:1, December 2019.
- [VP98] V. Vedral and M. B. Plenio.
Entanglement measures and purification procedures.
Physical Review A, 57(3):1619–1633, March 1998.
- [VPRK97] V. Vedral, M. B. Plenio, M. A. Rippin, and P. L. Knight.
Quantifying entanglement.
Physical Review Letters, 78(12):2275–2279, March 1997.
- [VV14] Umesh Vazirani and Thomas Vidick.
Fully device-independent quantum key distribution.
Physical Review Letters, 113(14), September 2014.

- [WDH22] Marek Winczewski, Tamoghna Das, and Karol Horodecki.
Limitations on a device-independent key secure against a nonsignaling adversary via squashed nonlocality.
Physical Review A, 106(5), November 2022.
- [WDS⁺18] Marek Winczewski, Tamoghna Das, John H. Selby, Karol Horodecki, Paweł Horodecki, Łukasz Pankowski, Marco Piani, and Ravishankar Ramanathan.
Complete extension: the non-signaling analog of quantum purification.
Accepted for publication in Quantum 27.01.2023, 2018.
arXiv:1810.02222.
- [WDW17] Qingle Wang, Siddhartha Das, and Mark M. Wilde.
Hadamard quantum broadcast channels.
Quantum Information Processing, 16(10), August 2017.
- [WEH18] Stephanie Wehner, David Elkouss, and Ronald Hanson.
Quantum internet: A vision for the road ahead.
Science, 362(6412), October 2018.
- [Wie83] Stephen Wiesner.
Conjugate coding.
ACM SIGACT News, 15(1):78–88, January 1983.
- [Wil96] R.J. Wilson.
Introduction to Graph Theory.
Longman, 1996.
- [Wil16] Mark M. Wilde.
Squashed entanglement and approximate private states.
Quantum Information Processing, 15(11):4563–4580, September 2016.
- [Wil18] Alexander Wilce.
A royal road to quantum theory (or thereabouts).
Entropy, 20(4):227, March 2018.
- [Wil21] Mark M. Wilde.
Second law of entanglement dynamics for the non-asymptotic regime.
In *2021 IEEE Information Theory Workshop (ITW)*. IEEE, oct 2021.
- [WR12] Ligong Wang and Renato Renner.
One-shot classical-quantum capacity and hypothesis testing.
Physical Review Letters, 108(20), May 2012.

- [WTB17] Mark M. Wilde, Marco Tomamichel, and Mario Berta.
Converse bounds for private communication over quantum channels.
IEEE Transactions on Information Theory, 63(3):1792–1817, March 2017.
- [WvdW22] Bas Westerbaan and John van de Wetering.
A computer scientist’s reconstruction of quantum theory*.
Journal of Physics A: Mathematical and Theoretical, 55(38):384002, August 2022.
- [WW01] R. F. Werner and M. M. Wolf.
All-multipartite Bell-correlation inequalities for two dichotomic observables per site.
Physical Review A, 64(3), August 2001.
- [WWY14] Mark M. Wilde, Andreas Winter, and Dong Yang.
Strong converse for the classical capacity of entanglement-breaking and hadamard channels via a sandwiched Rényi relative entropy.
Communications in Mathematical Physics, 331(2):593–622, July 2014.
- [WZ82] W. K. Wootters and W. H. Zurek.
A single quantum cannot be cloned.
Nature, 299(5886):802–803, October 1982.
- [WZG⁺16] Yadong Wu, Jian Zhou, Xinbao Gong, Ying Guo, Zhi-Ming Zhang, and Guangqiang He.
Continuous-variable measurement-device-independent multipartite quantum communication.
Physical Review A, 93(2), February 2016.
- [XMZ⁺20] Feihu Xu, Xiongfeng Ma, Qiang Zhang, Hoi-Kwong Lo, and Jian-Wei Pan.
Secure quantum key distribution with realistic devices.
Reviews of Modern Physics, 92(2), May 2020.
- [YCZ⁺12] Sixia Yu, Qing Chen, Chengjie Zhang, C. H. Lai, and C. H. Oh.
All entangled pure states violate a single Bell’s inequality.
Physical Review Letters, 109(12), September 2012.
- [YHH⁺09] Dong Yang, Karol Horodecki, Michał Horodecki, Paweł Horodecki, Jonathan Oppenheim, and Wei Song.
Squashed entanglement for multipartite states and entanglement measures based on the mixed convex roof.
IEEE Transactions on Information Theory, 55(7):3375–3387, July 2009.

- [ŻB02] Marek Żukowski and Časlav Brukner.
Bell’s theorem for general N-qubit states.
Physical Review Letters, 88(21), May 2002.
- [Zie95] G. M. Ziegler.
Lectures on Polytopes, volume GTM 152.
Springer Verlag, New York, 1995.
- [ZPD⁺18] M. Zwerger, A. Pirker, V. Dunjko, H. J. Briegel, and W. Dür.
Long-range big quantum-data transmission.
Physical Review Letters, 120(3), January 2018.
- [ZvLR⁺22] Wei Zhang, Tim van Leent, Kai Redeker, Robert Garthoff, René Schwonnek, Florian Fertig, Sebastian Eppelt, Wenjamin Rosenfeld, Valerio Scarani, Charles C.-W. Lim, and Harald Weinfurter.
A device-independent quantum key distribution system for distant users.
Nature, 607(7920):687–691, jul 2022.
arXiv:2110.00575.
- [ZXC⁺18] Qiang Zhang, Feihu Xu, Yu-Ao Chen, Cheng-Zhi Peng, and Jian-Wei Pan.
Large scale quantum key distribution: challenges and solutions [invited].
Optics Express, 26(18):24260, August 2018.
- [Życ08] Karol Życzkowski.
Quartic quantum theory: an extension of the standard quantum mechanics.
Journal of Physics A: Mathematical and Theoretical, 41(35):355302, July 2008.
- [ŻZHE93] M. Żukowski, A. Zeilinger, M. A. Horne, and A. K. Ekert.
“Event-ready-detectors” Bell experiment via entanglement swapping.
Physical Review Letters, 71(26):4287–4290, December 1993.

Hybrid quantum network design against unauthorized secret-key generation, and its memory cost

Omer Sakarya ¹, Marek Winczewski ^{2,3}, Adam Rutkowski ⁴ and Karol Horodecki ^{1,3}

¹*Institute of Informatics, National Quantum Information Centre, Faculty of Mathematics, Physics and Informatics, University of Gdańsk, Wita Stwosza 57, 80-308 Gdańsk, Poland*

²*Institute of Theoretical Physics and Astrophysics, National Quantum Information Centre, Faculty of Mathematics, Physics and Informatics, University of Gdańsk, Wita Stwosza 57, 80-308 Gdańsk, Poland*

³*International Centre for Theory of Quantum Technologies, University of Gdańsk, Wita Stwosza 63, 80-308 Gdańsk, Poland*

⁴*Institute of Theoretical Physics and Astrophysics, University of Gdańsk, 80-308 Gdańsk, Poland*



(Received 26 March 2020; accepted 30 July 2020; published 5 October 2020)

A significant number of servers that constitute the Internet are to provide private data via private communication channels to mutually anonymous registered users. Such are the servers of banks, hospitals that provide cloud storage and many others. Replacing communication channels by maximally entangled states is a promising idea for the quantum-secured Internet (QI). While it is an important idea for large distances secure communication, for the case of the mentioned class of servers pure entanglement based solution is not only unnecessary but also opens a threat. A crack stimulating a node to generate secure connections via entanglement swapping between two hackers can cause uncontrolled consumption of resources. Turning into positive a recently proven no-go result by S. Bäuml *et al.* [*Nat. Commun.* **6**, 6908 (2015)], we propose a natural countermeasure against this threat. The solution bases on connections between hub-nodes and end-users realized with states that contain secure key but do not allow for swapping of this key. We then focus on the study of the quantum memory cost of such a scheme and prove a fundamental lower bound on its memory overhead. In particular, we show that to avoid the possibility of entanglement swapping, it is necessary to store at least twice as much memory than it is the case in standard quantum-repeater-based network design. For schemes employing either states with positive partial transposition that approximates certain private states or private states hardly distinguishable from their attacked versions, we derive much tighter lower bounds on required memory. Our considerations yield upper bounds on a two-way repeater rate for states with positive partial transposition (PPT), which approximates strictly irreducible private states. As a byproduct, we provide a lower bound on the trace distance between PPT and private states, shown previously only for private bits.

DOI: [10.1103/PhysRevResearch.2.043022](https://doi.org/10.1103/PhysRevResearch.2.043022)

I. INTRODUCTION

The domain of quantum information processing, which shows how the rules of quantum mechanics can meet the needs of information society [1,2], has reached its maturity in recent years. We are about to enter the NISQ era of quantum computing with the noisy intermediate scale quantum (NISQ) devices ahead of us [3]. In parallel, a huge effort has been done towards building the quantum Internet (QI) [4–6], which is predicted to be built within several years [7]. It is viewed as a network of NISQ devices with their memory and the central processing unit (CPU), which exchange *qubits* rather than classical bits between each other.

The main welcome feature of the Quantum Internet in comparison with the traditional Internet is its, speaking of theory, the *inherent security of sent signals*. The first-generation

QI [5] bases on the quantum correlations called *entanglement* and its advantageous property of *transitivity*. In theory, a two otherwise disconnected nodes can obtain mutual unconditionally secure connection if only they share maximally entangled state (singlet) with a common node, via the *entanglement swapping* protocol [8,9]. Due to the high attenuation of quantum signals in optical fiber and impossibility of their amplification by cloning [10], the number of intermediate nodes which perform entanglement swapping (*quantum repeaters* [4]), needs to be large, and function in high coordination. Let us recall here that the quantum repeaters protect sent qubits against eavesdropping because entanglement swapping uses, in fact, quantum teleportation [9]. Indeed, quantum teleportation protocol allows for a transfer of data without any intermediate point in space-time, where it could be attacked.

While the QI is about to come, a number of serious attacks on the traditional Internet which is working already for about a half a century is being more and more often reported. This happens in accordance with a growing interest in network cybersecurity. One of the simplest attacks on the network is the hijacking of a node, via a *malware*—a malicious piece of software which changes its functioning at a wish of a

Published by the American Physical Society under the terms of the [Creative Commons Attribution 4.0 International](https://creativecommons.org/licenses/by/4.0/) license. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.

hacker. Possible attacks on future quantum Internet has been recently considered [11,12]: a piece of software infects the CPU of a quantum device of the node of quantum repeater, leading, e.g., to local change of topology of the network. While proposals for overcoming the implications of such an attack are developed, we focus on a solution which to some extent, prevents it due to laws of physics.

Hybrid quantum network. As it is common in quantum information theory, a no-go (impossibility) in processing of quantum data can be exploited as its potential: quantum no-cloning led to the seminal ideas of quantum money and quantum cryptography protocols [1,2] while impossibility of prediction of measurement outcomes (attributing the so-called *hidden variable model*) led further to the device independent quantum security [13,14]. Our countermeasure to hijacking is also based on a recently found no-go, which can be stated as follows. *There exist quantum states which allow for point-to-point security of classical data against quantum adversary, and in spite of this fact can not be effectively used in quantum key repeaters* [15].

The above result shows that quantum security is not always *transitive*: for certain states (call them *nonrepeatable secure states* ρ), conversely to entanglement swapping, when A has secure link (possessing ρ) with B and B with C , there is no possibility for B to help A and C , via a three-partite local quantum operations and classical communication (3-LOCC), to share a secure link, protected against B as well. Certain bound entangled states [16] (from which no pure entanglement can be distilled by local operations and communication [17]) and highly noisy private states [18], has been recently shown, to fit the scheme in case of arbitrary 3-way and one-way classical communication (from the node B to A and C) respectively [15,19].

In this manuscript, we propose a general idea of *physical protection against malware* by presenting a flip side of the presented limitation on quantum repeaters. It amounts to deliberate use of the quantum states which disallow for repeating of secure key, in order to protect against any unauthorized network user who wants to perform it for his own purposes.

In the language of computer science, we propose an architecture and a model of the physical layer of the quantum network to exclude the possibility that its local topology is changed via attacking the network at the application layer.

We show that specially designed *hybrid quantum network*, i.e., based on both *repeaters* and special *relay stations* called here *hubs*, is more robust against special kind of attacks than original repeaters. We put forward a particular example of an attack and study properties of its countermeasure. To show the idea, we focus on a subnetwork of the hybrid quantum network, whose graph is a star, i.e., with a *central hub-node* and several (Δ) connected *end-nodes* (see Fig. 1).

Our approach suits the scenario in which (1) the hub-node can be connected by a quantum repeater with other hub nodes; (2) only classical data need to be sent between the hub node and the end-nodes; (3) the distance between the hub-node and the end-nodes is maximally of metropolitan scale (up to the distance available for repeaterless quantum networks [20,21]); (4) only disconnected hub-nodes and their two adjacent end-nodes are attacked at a time; and (5) the attack is honest but curious: only functioning of the classical processor

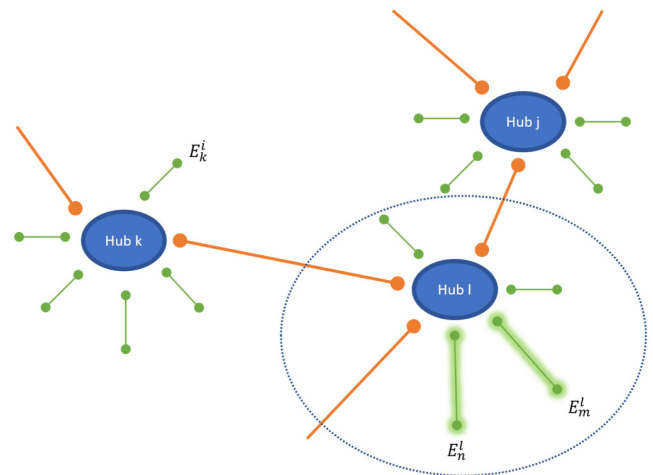


FIG. 1. Structure of the proposed hybrid network: thin green lines connect end-users with hubs; in these, only classical data can be transferred. Thick orange lines connect hubs being routing nodes of the network; these connections allow for passing a quantum state. Shaded lines connect two end-users that communicate securely with the hub node only classical data. Selected region denotes single hub of our interest.

is changed by malware, while classical data at the node remain unread.

The hybrid network is shown in Fig. 3. The above example fits the real use case, as in the network of the traditional Internet. Indeed in the Internet, there is quite a number of nodes representing servers that deliver certain utilities in the form of classical data, access to which is charged, and limited to a group of registered users. Moreover, the task of these nodes is not to connect the users, that are usually anonymous but to provide them an access to data via private link. Servers for online banking, access to medical data of a laboratory, online shops, and last but not least, providers of the data clouds form far from a complete list of examples of the latter. In some of these cases, the users are local so that the assumptions about the distance between end-nodes is satisfied. We focus on a star-shaped network with central hub connected to end-users. In this case, the data are generated in classical form. The distance between the users is usually not too big. It is also, needless to say, that security is vital since it is important for end-users, users of the hub, administrators or the owner of the server. We also focus on the case when two dishonest users of the network hijack a single node. Their task is to obtain a free secure connection. In other words the attack is a theft of processor time and power aimed at generation of secret-key. The main feature of our solution is that the *topology of the network* is naturally, physically protected against modification.

As it is usual, any good comes at a price. In the above case, the price will come out in the number of qubits needed to be stored (or processed) in quantum memory of a node. In the NISQ era, it is of prime importance to find how much of quantum memory is necessary in order to realize a given quantum network architecture. This is because, as of now, there is no technology to store coherently qubits for a long time. We therefore study lower bounds on the memory

cost related to the hybrid quantum network and also show that this architecture can be realized with relatively modest quantum memory requirements. In principle, to form the links of a quantum network, one can take any state containing the private key. Useful from a cryptographic point of view are in that respect the so-called private states [18,22] or their approximate versions with positive partial transposition. The private states have directly accessible key via measurement on their subsystems called key parts. However, especially those having low repeatable key [19], have also the shielding systems (shields). A shield protects the key but costs quantum memory. As we will show, these states are not the only ones that have memory cost.

We therefore first provide lower bounds on the memory cost of our secure network scheme, which is related to the *density of the secure key in quantum states* - a natural quantity that was implicitly used in Refs. [15,22,23]. To our knowledge, this quantity has not been explicitly studied on its own so far. We introduce a *memory overhead* as a measure of the cost. For a *scheme* S_ρ (that assures security of the hub node), its overhead is defined as

$$V(S_\rho) := M(\rho)(1 - \mathcal{D}(\rho)), \quad (1)$$

where \mathcal{D} is the density of the key, i.e., ratio of the key to the dimension of the state, and $M(\rho)$ is the total memory of the scheme. This intuitive quantity is 0 for maximally entangled states, as their whole memory has a form of the key. However, in general case of mixed quantum states, $V(S_\rho)$ is strictly larger than zero.

We then represent each link in the network by the same state ρ and study its usefulness in the context of hacking. The efficiency of a given scheme we quantify by the difference between the key that can be repeated R and the initial key of the link K_D . We then say that a scheme is (θ, η) -good, when $K_D \geq \eta$ but $R \leq \theta$, along with assumption $\eta > \theta$. This means that the link provides security and because it is not realized by pure state, one can not abuse the link to connect with someone else in the network.

We prove the general lower bound showing that for any state serving as reasonable secure network scheme at least half of the memory qubits (approximately) shall not be used for key distillation, i.e., $V(S_\rho) \geq \frac{1}{2}M(\rho)$. Different, however asymptotically equivalent bound we obtain for the private states [18,22]. For these specific states, we prove that the shield must be at least the size of the key part to assure the security of the scheme. We do so by finding explicit formula for the coherent information of a private state [24,25].

Aiming at set of states for which there are known examples that assure an $\approx (0, 1)$ -good scheme, we consider states that have positive partial transposition (PPT states), and approximate some private states. More precisely, we provide lower bounds for the memory cost of our secure network schemes (hubs) employing PPT states approximating strictly irreducible private states [26]. As a related problem being of independent interest, we provide an upper bound on two-way repeater rate for PPT states. These states (i) approximate strictly irreducible pdits for any dimension of the key part d_k (ii) satisfy structural constraints on its behavior under partial transposition map. For the considered class of states, the overhead approaches 1 in the limit of large dimensions.

However, the speed of this convergence is rather modest. We conclude from the formulas, that e.g., for a scheme with 80% gap, i.e., where $\theta - \eta \geq \frac{8}{10}$, it suffices to spend eight qubits on shield for one qubit in the key part. States realizing such schemes are known [22].

As a byproduct, we prove a lower bound for the trace norm distance between private states and PPT states approximating them. So far, only $d_k = 2$ case was known, which we also tighten. Finally, let us stress that, to our knowledge, the hybrid architecture of a quantum network proposed here is the first application of states with low distillable entanglement (or even bound entangled [16]) in practice.

The paper is organized as follows. In the next section II, we specify and describe an example of the proposed secure-network scheme. In Sec. III, we introduce the memory overhead of the scheme and the density of key. In subsequent Sec. IV, we provide lower bound on overhead for irreducible private states and also a general lower. In Sec. V, we quantify the scheme that uses private states hardly distinguishable from their attacked versions, whereas in Sec. VI, we concentrate on bounds for certain PPT states. Section VII is left for discussion.

II. STAR-SHAPED NETWORK: THE CASE STUDY OF THE ATTACK AND COUNTERMEASURE

In this section, we describe in detail the scenario for which, given quantum Internet happens to be realized in a form suggested nowadays, an attack via malware could be done. We then describe countermeasure invoking recent results on limitations on quantum key repeaters

A. Attack on the star-shaped, pure entanglement based quantum network.

We focus the following specific example of the above-explained scenario. The hub shares secure links with many end-users E_i with $i \in \{1, \dots, n\}$, in particular with Adam and Eve [see Fig. 2(a)]. The natural topology of the network of secure links is the *star* one (see Fig. 1), so that each end-user is connected with the hub. The hub network node is assumed to be a unit with classical and quantum computer inside. The crucial observation is that if the links are quantum, and based on pure entanglement, they allow via entanglement swapping for the change of topology of the network. Indeed, it can change from star to a disconnected graph of at least two components: star without some nodes and a pair of end-users having a secure connection between them and sharing no more the connection with the hub.

For the above reason, setting up a star network based on pure entanglement, the hub opens a possibility to provide security to pairs of end-users [see Fig. 2(b)]. On the other hand, states allowing quantum communication seem to be an overkill in the case where the node exchanges with subnodes inherently classical information like in the mentioned list of examples of online services. Such an additional side-effect possibility should be under control of the hub that owns the subnetwork. A solution would be to designate a person who sells the connections. If it is not the case, there is a possibility of two dangers: firstly, the administrators of the hub can sell

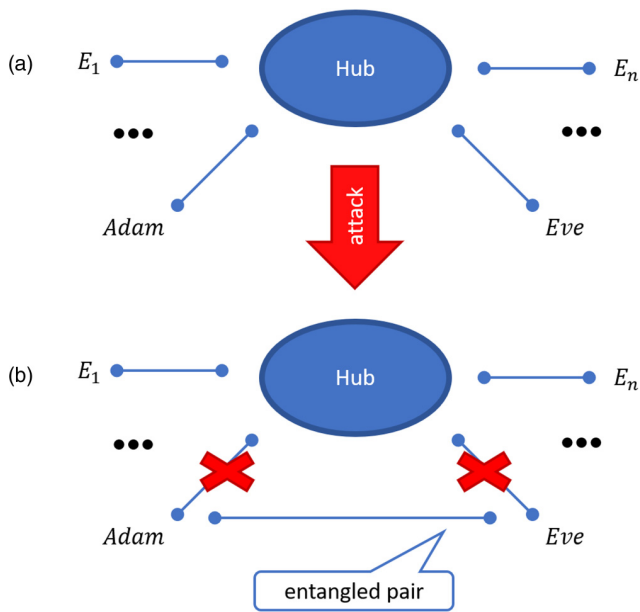


FIG. 2. The main idea of an attack: (a) the hub shares entangled pairs with end-users E_i , in particular, with Adam and Eve. Adam and Eve can attack the hub via quantum malware which performs for them entanglement swapping (b) Adam and Eve share an entangled pair after successful attack.

the secure links by themselves and earn illegally without notice of the hub’s owner. Secondly, a more dangerous threat is possible: two end-users Adam and Eve can hack the system installing a Trojan quantum software, which serves them as a source of cheap security. Even more importantly, in this way energy consumed for performing quantum operations would be stolen, again, without notice of the hub. Let us note that the same holds if the hub is one of a number of repeater stations [4], and the links are improved via entanglement distillation [9].

We will distinguish here two kinds of attacks: a *general one* where the hacked node can perform any three-party classical communication with two other nodes and *one-way attack* where only the central node can communicate classical information to the hackers.

B. Countermeasure via noisy entangled states

In what follows, we observe, that using appropriate noisy entangled states solves the mentioned problem in cases of general (two-way) and one-way attack (see Fig. 3).

Recently a fundamental result has been shown in this context, indicating that for some states (having at least one separable key attacked state) the rate R of repeated secure key is strongly related to the so-called *distillable entanglement* [9,17,19] by the following result:

$$R^{H_1H_2 \rightarrow A:E}(\gamma_{AH_1}, \gamma_{H_2E}) \leq E_D^{H_1H_2 \rightarrow A:E}(\gamma_{AH_1} \otimes \gamma_{H_2E}), \quad (2)$$

where \rightarrow stands for the classical communication restricted to one-way from the intermediate node $H \equiv H_1H_2$ to nodes A and E , and γ_{AH_i} denotes a private state [18]—a state possessing ideal security directly accessible via measuring its subsystem called *key part*.

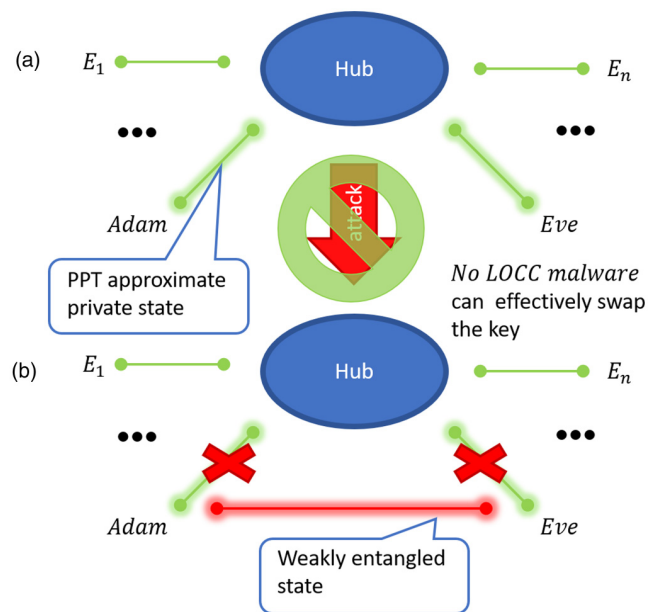


FIG. 3. The main idea of the countermeasure: (a) hub shares bound entangled states (green lines) with end-users E_i (each having at least 1 bit of key), in particular, with Adam and Eve (shaded lines). No LOCC malware can efficiently swap the key. (b) Adam and Eve can not share a state (shaded red line) with non-negligible amount of key (compare [15]).

Notation 1. Private state with d_k dimensional key part and d_s dimensional shield part per one party, shared between A or E and H is denoted γ_{d_k, d_s} .

We present the following countermeasure: instead of having a star network with the end-users, which is pure entanglement based, the hub can set a star-shaped network of point-to-point links based on bound entangled states which are approximate private states (see Fig. 3). Let us note that this is legitimate when the hub needs to encrypt only classical data. Furthermore, if end-users had a quantum connection with the hub, then we could have a case of hub’s network abuse. These bound entangled states are *weakly* transitive. This means that there *does not exist* a quantum software that can be run on the quantum computer of the hub, or even a quantum tripartite LOCC protocol between Adam, Eve, and the quantum computer of the hub, to achieve this task. The *no-go* is hence turned into a success. The hub employing the bound entanglement based quantum links keeps secure communication but needs not to control the setup. There is *no* physical map in the considered scenario, that can create a secure link with a non-negligible amount of secrecy.

Let us note that although we have mentioned here entanglement swapping, in Ref. [15], it is shown that even if the links with the hub are provided in the form of a private state $\gamma_{AH_i}^{\otimes n}$ (or an approximate private state), the rate of the output secure key for Adam and Eve is negligible as a function of dimension of the bound entangled approximate private states. By *negligible* amount, we mean the rate which goes to zero with growing dimension of the shield system of the state γ_{AH_i} . Hence the countermeasure works in the asymptotic regime up to the fact that some small rate of key can be obtained by Adam and

Eve. Note here that the key of the links is not in the form of pure entanglement. The states are chosen such that they have negligible or even zero distillable entanglement. However, such a choice of states for the links is not a constraint for the security of the network. This is because we consider only hubs that send classical data, hence their encryption does not need to involve pure entanglement. Moreover, the states considered in our solution, have large enough distillable key.

III. MEMORY OVERHEAD OF THE COUNTERMEASURE

We now focus on the *quantum memory cost* of implementation of the proposed countermeasure. We recall first the definition of the key repeater rate. Let us stress here that according to our approach, the lower it is, the better for the security of the node.

We further focus on the scheme represented by a private state with d_k dimensional key-part and d_s dimensional shield. This state reads a form [18]:

Definition 1. Private quantum state

$$\gamma_{d_k, d_s} := \sum_{i, j=0}^{d_k-1} \frac{1}{d_k} |ii\rangle\langle jj| \otimes X_{ij}, \quad (3)$$

where $X_{ij} = U_i \sigma U_j^\dagger$ for some state σ of $\mathcal{C}^{d_k} \otimes \mathcal{C}^{d_s}$ and U_i are some unitary transformations.

Notation 2. We follow the notation in which

$$\|X\|_1 = \text{Tr} \sqrt{XX^\dagger}. \quad (4)$$

Additionally we skip the subscript, as it doesn't lead to any ambiguity.

Remark 1. Through the rest of the paper, we assume that each considered quantum state ρ acts on $\mathcal{H}_H \otimes \mathcal{H}_N$ being tensor product of subspaces associated with the hub and a node, and $\dim \mathcal{H}_H = \dim \mathcal{H}_N < \infty$. What is more, both subspaces are assumed to be partitioned into key and shield parts (of dimensions d_k and d_s , respectively) in the same way at both sides, unless stated differently.

Notation 3. Here we adapt shortened notation in which $X_{ij} \equiv X_{ii, jj}$. In calculations, we mainly incorporate full notation. Additionally for $i \neq j$, we define $X_{ij, ij} \equiv 0$, as they do not enter to definition of a private state.

Note that X_{ii} are, in fact, subnormalized states, obtained on the shield system upon observing key $|i\rangle$ on the key part. We call them *conditional states*. According to definition, $K_D(\gamma_{d_k, d_s}) \geq \log_2 d_k$, while in case of equality, a private state is called *irreducible*: its whole secure content is available from the key part via direct measurement. In the case in which X_{ii} are additionally separable, we call these states strictly irreducible private states. In fact, it is conjectured that all irreducible private states are of the form of strictly irreducible ones [26], it is so if there do not exist entangled but key-undistillable states.

Definition 2. The distillable key rate with respect to arbitrary LOCC operations is defined as

$$K_D(\rho) := \inf_{\epsilon > 0} \limsup_{n \rightarrow \infty}$$

$$\sup_{\Lambda_n^{\text{LOCC}}, \gamma_{d_k, d_s}} \left\{ \frac{\log_2 d_k}{n} : \Lambda_n^{\text{LOCC}}(\rho^{\otimes n}) \approx_\epsilon \gamma_{d_k, d_s} \right\}, \quad (5)$$

where ρ is a bipartite state shared by the parties. Λ is a LOCC protocol with two-way classical communication.

Definition 3. The quantum key repeater rate with respect to arbitrary LOCC operations among A , E , and H is defined as

$$R^{A \leftrightarrow H \leftrightarrow E}(\rho, \rho') := \lim_{\epsilon \rightarrow 0} \lim_{n \rightarrow \infty} \sup_{\Lambda_n^{\text{LOCC}}, \gamma_{d_k, d_s}} \left\{ \frac{\log_2 d_k}{n} : \text{Tr}_H \Lambda_n^{\text{LOCC}}((\rho \otimes \rho')^{\otimes n}) \approx_\epsilon \gamma_{d_k, d_s} \right\}, \quad (6)$$

where Adam and Hub share state ρ while Hub and Eve share ρ' . $\Lambda := \{\Lambda_n^{\text{LOCC}}\}$ are tripartite LOCC protocols with two-way classical communication. In the case in which communication between central node and A , E systems is restricted to one-way from H to A and E , we denote this rate with $R^{H \rightarrow A:E}$.

Notation 4. For the repeater rate in the case in which $\rho = \rho'$, we introduce simplified notation $R^{\rightarrow(\leftrightarrow)}(\rho)$.

An ultimate goal would be to provide a nonrepeatable key with the smallest possible memory cost, being a precious resource in NISQ era of quantum computing. Our solution to the problem is represented by a bipartite quantum state ρ shared between the central node H and one of the end-users (Adam), however, its specific parameters are important enough to write them out explicitly. The scheme will be represented by the following tuple:

$$S_\rho^{\rightarrow(\leftrightarrow)} := (\rho, \log_2 \dim_H(\rho), \Delta, K_D(\rho), R^{\rightarrow(\leftrightarrow)}(\rho)). \quad (7)$$

The arrow(s) in the superscript are dropped if the results hold for both cases. The state ρ_{HA} is shared between the central node H and a single end-user (Adam). Δ is the degree of the node (number of connections).

Definition 4. The scheme S_ρ is one-way (two-way) (θ, η) -good if $R^{\rightarrow(\leftrightarrow)}(\rho) \leq \theta$ and $K_D(\rho) \geq \eta$, and $\eta > \theta$.

Let us note here, that by definition of the key repeater rate, $K_D \geq R^{\leftrightarrow} \geq R^{\rightarrow}$. This means that if $R^{\leftrightarrow} = \theta$, we have $\theta \leq \eta$. However, since θ is only an upper bound on R^{\leftrightarrow} , we had to assume in the above definition desired order of parameters. Moreover we demand strict inequality $\eta > \theta$, since the scheme with $\eta = \theta$ is not "good," i.e., does not have any advantage over quantum repeaters design. Indeed, in our approach, we are interested in the largest possible gap between K_D and $R^{\rightarrow(\leftrightarrow)}$, while keeping memory overhead considerably small at the same time. We quantify this gap by its lower bound defined as a difference $\eta - \theta$, and call it *efficiency* of the scheme.

Definition 5. The *overhead* of the scheme S_ρ is the following quantity:

$$V(S_\rho) := \Delta(\log_2 \dim_H(\rho) - K_D(\rho)), \quad (8)$$

where ρ is a bipartite state shared between the hub and a end-user.

The overhead is the difference between the qubits of memory at the node: $\rho^{\otimes \Delta}$ has

$$M(\rho) := \Delta \log_2 \dim_H(\rho), \quad (9)$$

of qubits of subsystem H , and the number of bits of security which the node shares with the other part of the quantum Internet.

Definition 6. For a scheme that is (θ, η) -good, the difference $\eta - \theta \geq 0$ is the gap of the scheme.

We note here that such defined overhead bears strong connection with the other, to our knowledge not explicitly studied notion, which is that of *density of the private key*.

Definition 7. For a quantum state ρ ($\dim_H(\rho) \geq 2$) shared between the hub H and Adam A or Eve E , the density of the private key \mathcal{D} reads

$$\mathcal{D}(\rho) := \frac{K_D(\rho)}{\log_2 \dim_H(\rho)}. \quad (10)$$

In the above definition, we have included a restriction $\dim_H(\rho) \geq 2$ to exclude trivial, not relevant cases. We then have the dependence

$$V(S_\rho) = M(\rho)(1 - \mathcal{D}(\rho)). \quad (11)$$

From the above form it is clear to see that the overhead is a non-negative quantity, as the density is a quantity less than or equal to 1. In what follows, we provide several lower bounds on the overhead V of the countermeasure, that satisfies

$$0 \leq V(S_\rho) \leq M(\rho). \quad (12)$$

The first from the above inequalities follows from the fact that secure key $K_D(\rho)$, can not be larger than memory size $\log_2 \dim_H(\rho)$, and hence Δ is non-negative. Presenting results on plots in the next sections, we will concentrate on the fraction between memory overhead and total memory of the hub, i.e., the percentage of total memory that is not used for storing secret-key.

IV. LOWER BOUNDS ON THE OVERHEAD OF THE SECURE NETWORK SCHEME

Let us first focus on the class of *one-way attacks*: the attacked hub node can send data to two receiver nodes owned by malicious parties that can communicate freely. We begin with preliminary definitions and facts.

Definition 8. The coherent information of a quantum state ρ_{AB}

$$I_{\text{coh}}(A)B = S(B) - S(AB), \quad (13)$$

where $S(B)$ is the Von Neumann entropy of state $\rho_B = \text{Tr}_A(\rho_{AB})$ and $S(AB)$ is that of state the ρ_{AB} .

The key repeater rate is an upper bound on distillable entanglement in each of the two links of the star-shaped network. We therefore provide a lower bound on one-way distillable entanglement $E_D^\rightarrow(\cdot)$ of a private state, via the Devetak-Winter hashing protocol [27].

$$E_D^\rightarrow(\gamma_{d_k, d_s}) \geq \log_2 d_k + \sum_i \frac{1}{d_k} I_{\text{coh}}(A')B'_{\sigma_i}, \quad (14)$$

where I_{coh} is the coherent information [17], and σ_i are the conditional states of a private state. We have then the following observation.

Observation 1. For any private state γ_{d_k, d_s} , one-way distillable entanglement is lower bounded as follows:

$$E_D^\rightarrow(\gamma_{d_k, d_s}) \geq \log_2 d_k + \sum_i \frac{1}{d_k} I_{\text{coh}}(A')B'_{\sigma_i}, \quad (15)$$

where $\sigma_i = U_i \rho_{A'B'} U_i^\dagger$ are conditional states.

For the proof of the above observation see Appendix.

Let us note that the above bound is achievable given a choice $\forall_i \sigma_i = \frac{1}{d_s}$, i.e., for pdits with twisted-in maximally mixed state.

Since coherent information can not be smaller than $-\log_2 d$ for a d dimensional state, we have the following general result.

Corollary 1. For any private state γ_{d_k, d_s} one-way distillable entanglement is lower bounded by the following expression:

$$E_D^\rightarrow(\gamma_{d_k, d_s}) \geq \log_2 d_k - \log_2 d_s, \quad (16)$$

where d_k and d_s , are dimensions of the key part and shield part respectively.

Proof. It follows from the fact that for any state σ_i of dimension d_s^2 , there is $I_{\text{coh}}(A')B'_{\sigma_i} \geq -\log_2 d_s$. Indeed, $S(B') - S(A'B') = I(A' : B') - S(A') \geq 0 - \log_2 |A'| = -\log_2 d_s$, as the entropy is maximally $\log_2 |A'|$ while $I(A' : B') \geq 0$. ■

Following the fact that one-way distillable entanglement constitutes a lower bound on both one-way and two-way repeater rates, we conclude that in schemes incorporating private states, it is reasonable to assume $d_s \geq d_k$. This assumption is a necessary condition for having low repeater rates.

As we have discussed, we obtain the following lower bound on overhead of schemes based on irreducible private states.

Theorem 1. If an irreducible private state γ_{d_k, d_s} serves as an $(\theta, \log_2 d_k)$ -good secure network scheme $S_{\gamma_{d_k, d_s}}^\rightarrow$ with degree Δ , then its overhead satisfies a lower bound:

$$V(S_{\gamma_{d_k, d_s}}^\rightarrow) \geq \Delta \log_2(d_k d_s) \left(1 - \frac{1}{2 - \frac{\theta}{\log_2 d_k}} \right) \quad (17)$$

$$\approx_{\theta \approx 0} \frac{1}{2} M(\gamma_{d_k, d_s}). \quad (18)$$

For the proof of the above theorem see Appendix.

This theorem shows that memory used by a private state which allows only for θ of repeated key must have at least as big shield system as its key part, see Fig. 4 for exemplary lower bounds. The technique used for proving theorem 1 inspired us to find a general lower bound on the overhead of any scheme, which is presented below.

Theorem 2. Any state ρ that serves as (θ, η) -good secure network scheme, satisfies

$$V(S_\rho) \geq M(\rho) \left(\frac{1}{2} - \frac{\theta}{\log_2 d_H} \right) \approx_{\theta \approx 0} \frac{1}{2} M(\rho). \quad (19)$$

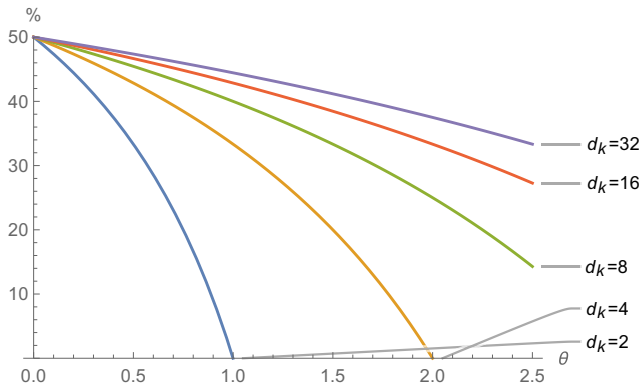


FIG. 4. Plots of lower bounds on percentage of memory overhead from theorem 1, with different values of d_k .

For the proof of the above theorem see Appendix. The above theorem is based on observation that distillable key is upper bounded by $S(A)/2$ if only coherent information is nonpositive. As we will show below on Fig. 5, this bound is the only bound on key repeater rate for certain amount of one-way distillable entanglement.

The bound shown in Fig. 5 as dotted blue line segment reads

$$R^{\rightarrow}(\rho) \leq R^{\leftrightarrow}(\rho) \leq K_D(\rho) \leq E_{\text{sq}}(\rho) \leq \frac{S(A)}{2} + \frac{E_D^{\rightarrow}(\rho)}{2}. \quad (20)$$

The inequality in Eq. (20) is a known fact, since one-way communication from the hub H to hosts A and E can not allow to repeat more key than in two-way communication setup. The second inequality comes from the fact that it is not possible to

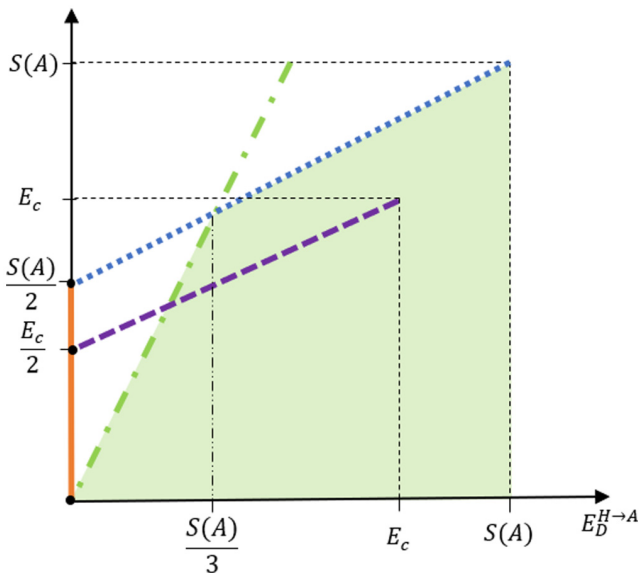


FIG. 5. Upper bounds on quantum key repeater rates. Dotted blue line: introduced here in eqn. (20), solid orange line: special case for $I_{\text{coh}} = 0$, dashed violet line: [15], and dotted-dashed green line: only for some states [19]. Shaded region corresponds to a combination of bounds.

have more of a repeatable key than a distillable key. On the other hand, it is possible that the quantum key repeater rate is smaller than the distillable key of a particular state ρ . The third inequality is true because squashed entanglement is an upper bound on distillable key [28]. Finally, the last inequality is the upper bound on R^{\rightarrow} observed in this work, which is a direct consequence from the proof of lemma 18 in Ref. [15]. Similar results on private capacity for quantum channels were obtained in Ref. [29]. As one can see the result for states which we prove here is much simpler than the analogous one proved there for channels.

The dotted-dashed green line segment is the upper bound on quantum key repeater rate derived in Ref. [19]:

$$R^{\rightarrow}(\rho) \leq 2E_D^{\rightarrow}(\rho), \quad (21)$$

which holds for special class of *block states*. Here the hub can send messages to Adam and Eve, but not receive from them. Adam and Eve can communicate in both ways freely.

The dashed violet line segment is the upper bound introduced in Ref. [15]:

$$R^{A \leftarrow H \rightarrow E}(\rho) \leq \frac{E_D^{\rightarrow}(\rho)}{2} + \frac{E_C(\rho)}{2}. \quad (22)$$

In this case, only the communication from Hub to Adam is one-way and between Hub and Eve the communication is one-way, no other data transfer is allowed.

The solid orange line segment is the upper bound for states that have $I_{\text{coh}} = 0$. These states do not have more of distillable key than $E_C/2$ or $S(A)/2$.

Even though dotted-dashed green and dashed violet bounds intersect in $E_D^{\rightarrow} = E_C/3$, they are different scenarios in which the classical communication is not in the same direction. Therefore, they are incomparable. It is the same for dotted blue and dashed violet bounds. On the other hand, the directions of classical communication for dotted-dashed green and dotted blue bounds are the same, so it is possible to compare them. The upper bound introduced in this work is more accurate than the bound derived in Ref. [19] starting from $E_D^{\rightarrow} = S(A)/3$.

V. LOWER BOUND ON OVERHEAD FOR PRIVATE STATES HARDLY DISTINGUISHABLE FROM THEIR ATTACKED VERSIONS

In this section, we derive lower bounds for the memory overhead for schemes utilizing private states hardly distinguishable from their attacked versions. We first briefly explain the approach and then formalize the presented idea.

Let us note, that to assure $\eta > 0$ in our scheme S_ρ , we need to know how much a given state of it ρ has distillable key. A good choice is then a strictly irreducible private state, as for this state, we know that $K_D(\gamma_{(d_k, d_s)}) = \log_2 d_k$, however, such $\gamma_{(d_k, d_s)}$ should not be too much distillable, as $R^{\leftrightarrow} \geq E_D(\rho)$. Thus, to also have that scheme is (θ, η) -good for small θ , we need to assure $E_D(\gamma_{(d_k, d_s)}) \leq \theta$. This can be done in various ways, including logarithmic negativity bound $E_D(\rho) \leq -\log_2 \|\rho^\Gamma\|$ [30]. From Ref. [19], it follows that R^{\rightarrow} is small since it is upper bounded by $2E_D^{\rightarrow}$. The next theorem encapsulates this approach and proves the lower bound on the memory cost of such a solution.

We first use the bound that employs measure called log-negativity [30,31].

Observation 2. For a private state such that $X_{ii} \in PPT$, and at least one from its conditional key attacked states is separable, there is the following bound on the one-way quantum key repeater rate:

$$E_D^{\rightarrow}(\gamma_{d_k, d_s} \otimes \gamma_{d_k, d_s}) \leq 2 \log_2 (1 + \|\gamma_{d_k, d_s}^{\Gamma} - \hat{\gamma}_{d_s, d_k}^{\Gamma}\|), \quad (23)$$

where $\hat{\gamma}_{d_k, d_s} = \sum_i \frac{1}{d_k} |ii\rangle\langle ii| \otimes X_{ii}$ is an irreducible private state after measurement on the key part (attacked), and Γ is an operation of partial transposition.

For the proof of the above observation see Appendix.

For technical reasons, we deal more specifically with the right-hand side of the above inequality, as encapsulated in the following observation.

Observation 3. The following identity holds:

$$\|\gamma_{d_k, d_s}^{\Gamma} - \hat{\gamma}_{d_k, d_s}^{\Gamma}\| = \sum_{i \neq j} \frac{1}{d_k} \|X_{ij}^{\Gamma}\|, \quad (24)$$

where $\hat{\gamma}_{d_k, d_s}^{\Gamma} = \sum_i \frac{1}{d_k} |ii\rangle\langle ii| \otimes X_{ii}^{\Gamma}$ is the private state after measurement on the key part and Γ is the partial transpose operation.

For the proof of the above observation see Appendix.

In the next lemma, we argue that some private states, that are hardly distinguishable from their attacked versions, have large dimension of the shield in relation to the dimension of the key part.

Lemma 1. For a special private state γ_{d_k, d_s} , which satisfies condition $X_{ii}^{\Gamma} \geq 0$, and $\|\gamma_{d_k, d_s}^{\Gamma} - \hat{\gamma}_{d_k, d_s}^{\Gamma}\| \leq \epsilon$, there is

$$d_s \geq \frac{d_k - 1}{\epsilon}. \quad (25)$$

For the proof of the above lemma see Appendix.

The above technical lemma and observation lead us to the main result of this section. It states that the overhead in case of private states that are hardly distinguishable from their attacked versions tends to 1 with the parameter of distinguishability approaching zero.

Theorem 3. A strictly irreducible private state $\gamma_{(d_k, d_s)}$ ($X_{ii} \in SEP$, $d_k \geq 2$) satisfying $\|\gamma_{(d_k, d_s)}^{\Gamma} - \hat{\gamma}_{(d_k, d_s)}^{\Gamma}\| \leq \epsilon$ and $\frac{d_k - 1}{d_s} \leq \epsilon$ serves as (θ, η) -good secure network scheme with

$$V(S_{\gamma_{(d_k, d_s)}^{\rightarrow}}) \geq M(\gamma_{(d_k, d_s)}) \left(1 - \frac{\log_2 d_k}{\log_2 d_k + \log_2 \frac{d_k - 1}{\epsilon}} \right) \quad (26)$$

$$\approx_{\epsilon \rightarrow 0} M(\gamma_{(d_k, d_s)}), \quad (27)$$

for $\theta = 2 \log_2 (1 + \epsilon) \approx_{\epsilon \ll 1} \frac{2}{\ln 2} \epsilon$ and $\eta = \log_2 d_k$.

For the proof of the above theorem see Appendix. For the performance of lower bound in different dimensions of key part see Fig. 6, for the behaviour of a gap see Fig. 7.

A. Example of the gap for low dimensional state

In general, one would like to diminish the repeater rate of the scheme as much as possible. Unfortunately, in theorem 3,

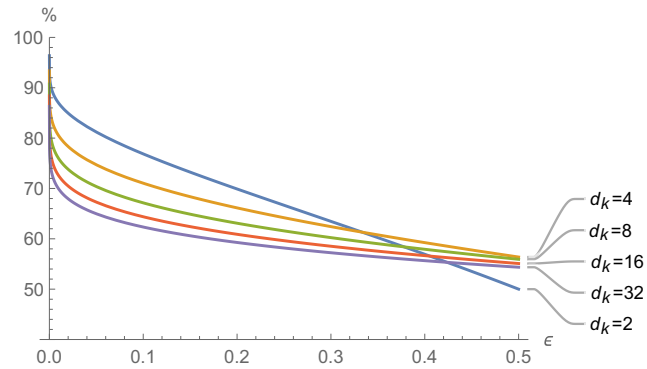


FIG. 6. Plots of lower bounds on percentage of memory overhead from theorem 3, with different values of d_k .

the parameter ϵ appears both in formula for the repeater rate and the overhead. This is the reason why one can not reduce repeater rate to zero keeping the overhead smaller than total memory cost. In this situation, one should decide on an acceptable level of repeater rate, for which the overhead is still reasonable. A small dimensional example of a pbit state which allows for such a control is known [18,32]. Block matrix representation of such a pbit is

$$\Omega_{d_s} = \frac{1}{2} \begin{bmatrix} \frac{1}{d_s^2} & 0 & 0 & \frac{F}{d_s^2} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \frac{F}{d_s^2} & 0 & 0 & \frac{1}{d_s^2} \end{bmatrix}, \quad (28)$$

where F is a matrix of swap quantum logic gate of dimension d_s^2 implying $\|\Omega_{d_s}^{\Gamma} - \hat{\Omega}_{d_s}^{\Gamma}\| = \frac{1}{d_s}$. We estimate now the size of the gap for a scheme using this state. Let us assume a scheme with minimal amount of memory by setting $\epsilon = \frac{1}{d_s}$ (see that conditions of lemma 1 and theorem 3 are satisfied). We obtain a lower bound $V(S_{\gamma_{(d_k, d_s)}^{\rightarrow}}) \geq M(\gamma_{(d_k, d_s)}) (1 - \frac{1}{1 + \log_2 d_s})$, for scheme being $(\frac{2}{\ln 2} \frac{1}{d_s}, 1)$ -good. For $d_s = 2$, it saturates also the general lower bound on overhead from theorem 2 with value of $\frac{1}{2}$, although in this case the rate of repeater R^{\rightarrow} is upper bounded with $\frac{1}{\ln 2} \approx 1.44$, what is an unsatisfying result. The first nontrivial case, in that secure network scheme

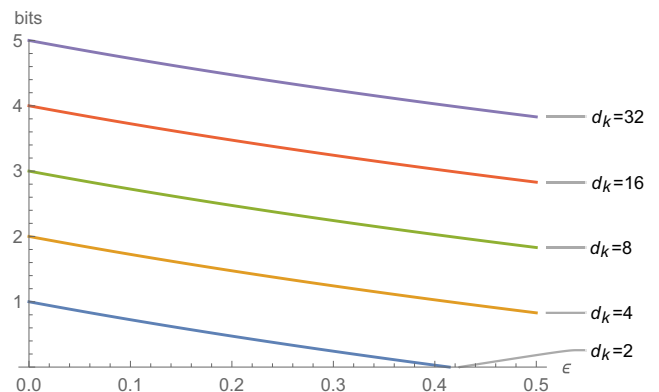


FIG. 7. Plots of lower bound on gap between η and θ for the scheme in theorem 3.

has an advantage over malicious parties, appears for $d_s = 3$, in which repeatable rate drops to $R^{\rightarrow} \leq \frac{2}{3 \ln 2} \approx 0.96$ being strictly smaller than key rate $K = 1$, what follows from its irreducibility.

VI. LOWER BOUNDS ON OVERHEAD FOR PPT STATES

As it was argued in Ref. [15] (see supplemental material note 6) the states which are PPT and approximate private bits are of rather high dimension. This fact can be found as a consequence of the following earlier statement [23] (see also Ref. [33]):

$$\forall_{\rho \in PPT} \gamma \in C^2 \otimes C^2 \otimes C^{d_s} \otimes C^{d_s} \|\rho - \gamma_{2,d_s}\| \geq \frac{1}{2(d_s + 1)}. \quad (29)$$

We conclude that a quantum PPT state close by ϵ in the trace norm to strictly irreducible private state γ_{d_k,d_s} has dimension of the shield at least $d_s \geq \frac{1-2\epsilon}{2\epsilon}$.

It is known that for two-way repeater rate to be zero, the state has to be bound entangled $[R^{\leftrightarrow}(\rho) \geq E_D(\rho)]$ [18]. Thus, in this section, we investigate the overhead using such schemes.

Notation 5. We adopt a notation in which PPT state ρ has the following form:

$$\rho := \sum_{i,j,k,l=0}^{d_k-1} |ij\rangle\langle kl| \otimes A_{ij,kl}, \quad (30)$$

where come $A_{ij,kl}$ are blocks of dimension d_s^2 .

Proposition 1. If ρ is a state with positive partial transpose, that approximates a strictly irreducible private bit $\|\rho - \gamma_{(2,d_s)}\| \leq \epsilon$ for $\epsilon \geq \frac{1}{2(d_s+1)}$, $\|A_{01,10}^\Gamma\| \leq \epsilon$, and its conditional shield states are separable, then its two-way repeater rate $R^{\leftrightarrow}(\rho)$ is upper bounded as follows:

$$R^{\leftrightarrow}(\rho) \leq 2\left(\sqrt{\epsilon} + \frac{3}{2}\epsilon\right)(1 + \log_2 d_s) + (1 + 2\sqrt{\epsilon} + 3\epsilon)h\left(\frac{2\sqrt{\epsilon} + 3\epsilon}{1 + 2\sqrt{\epsilon} + 3\epsilon}\right). \quad (31)$$

For the proof of the above proposition see Appendix.

Note that PPT states from proposition 1 above do exist. One example can be states for which $A_{01,10} = A_{00,11}^\Gamma$. For upper bounds on key repeater rate from this proposition for dimensions $d_s = 2, \dots, d_s = 32$ see Fig. 8.

Theorem 4. If a state with positive partial transpose ρ approximates strictly irreducible private bit $\|\rho - \gamma_{(2,d_s)}\| \leq \epsilon$ for $\frac{1}{2(d_s+1)} \leq \epsilon < \frac{1}{2}$, $\|A_{01,10}^\Gamma\| \leq \epsilon$, and its conditional shield states are separable, then it serves as a two-way (θ, η) -good secure network scheme S_ρ with degree Δ , and its overhead satisfies a lower bound:

$$V(S_\rho) \geq M(\rho) \left(1 - \frac{1 + (1 + \frac{\epsilon}{2})h\left(\frac{\frac{\epsilon}{2}}{1 + \frac{\epsilon}{2}}\right)}{1 + \log_2\left(\frac{1-2\epsilon}{2\epsilon}\right)} - \frac{\epsilon}{2}\right) \quad (32)$$

with $\eta = 1 - 8\epsilon - 4h(\epsilon)$ [where $h(\cdot)$ is the binary Shannon entropy] and $\theta = 2(\sqrt{\epsilon} + \frac{3}{2}\epsilon)(1 + \log_2 d_s) + (1 + 2\sqrt{\epsilon} + 3\epsilon)h\left(\frac{2\sqrt{\epsilon} + 3\epsilon}{1 + 2\sqrt{\epsilon} + 3\epsilon}\right)$.

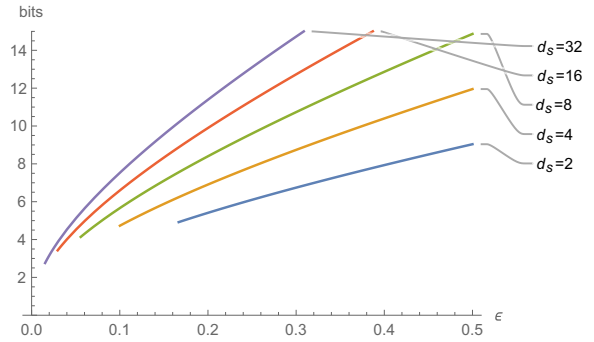


FIG. 8. Upper bounds on key repeater rate from proposition 1. Domains are constrained with $\epsilon \geq \frac{1}{2(d_s+1)}$ condition.

For the proof of the above theorem see Appendix. For the plot of lower bound of percentage of memory overhead from the above theorem see Fig. 9. The plots of lower bounds on gap between η and θ in this theorem are depicted in Fig. 10. From inequality (29), we obtain

$$\log_2 d_s \geq \log_2 \left(\frac{1 - 2\epsilon}{2\epsilon}\right). \quad (33)$$

We then see that focusing on states which have positive partial transposition and approximate private bits is quite costly: the overhead approximates the whole memory of the scheme for small ϵ . In particular, obtaining a reasonable amount of key in links ≈ 1 bits for each of Δ links implies that the whole memory cost is that of an overhead. However, an advantage of this scheme is that it is no longer limited to one-way communication. In this case, there *does not exist any three-partite LOCC protocol which can break the scheme.*

We now generalize the above result for larger dimensions of the key part than qubit, and study it in case of private state. In order to achieve this we need a number of technical observations and lemmas, which we present below.

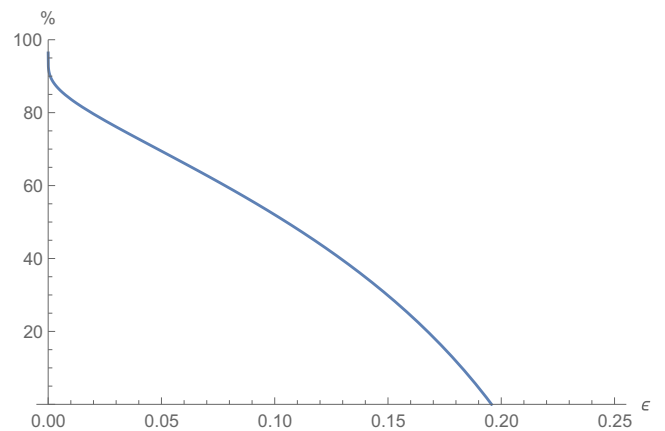


FIG. 9. Plot of lower bound on percentage of memory overhead from theorem 4.

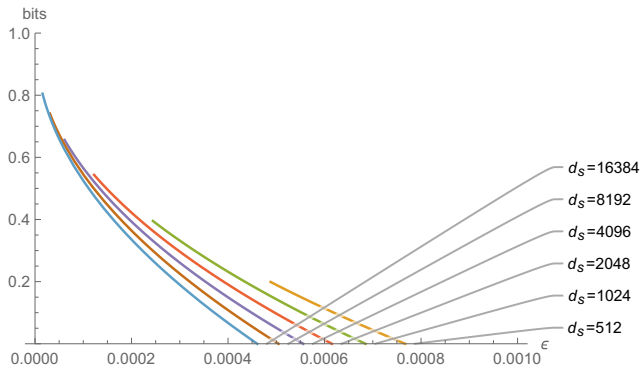


FIG. 10. Plots of lower bounds on gap between η and θ for the scheme in theorem 4. The case $d_s = 64$ is a setting with lowest dimension for obtaining positive lower bound on gap under $\epsilon \geq \frac{1}{2(d_s+1)}$ condition.

Observation 4. Denoting with $A_{ij,kl}$ matrices some of them ($A_{ii,jj}$) being unnormalized conditional states of the shield of a state $\rho = \sum_{ijkl} |ij\rangle\langle kl| \otimes A_{ij,kl}$, we prove the following relations:

$$\|\rho - \gamma\| \leq \epsilon \Rightarrow \forall_{i \neq j} \|A_{ii,jj}\| \geq \frac{1}{d_k} - \epsilon \quad (34)$$

and

$$\|\rho - \gamma\| \leq \epsilon \Rightarrow \sum_{i \neq j} \|A_{ij,ij}\| \leq \epsilon. \quad (35)$$

In the following lemma and subsequent corollary, we prove a general lower bound on the distance between private states (of any dimension of the key part) from PPT states [23].

Lemma 2. For any state $\rho \in PPT$, there is

$$\|\rho - \gamma_{d_k, d_s}\| \leq \epsilon \Rightarrow d_s \geq \left(\frac{d_k - 1}{\epsilon}\right)(1 - \epsilon d_k), \quad (36)$$

where γ is a private state with d_k^2 dimensional key part and d_s^2 dimensional shield subsystem.

Corollary 2. For any state $\rho \in PPT$ approximating private state, the following lower bound holds:

$$\|\rho - \gamma_{d_k, d_s}\| \geq \frac{d_k - 1}{d_s + d_k(d_k - 1)}. \quad (37)$$

The important properties of lower bound presented in corollary 2 are the fact that it is not trivial for values of d_k but also that it yields tighter bound for $d_k = 2$ known form [23] [see Eq. (29)]. Concluding as a byproduct, we have found a nontrivial (nonzero) lower bound on the distance between any private state and a PPT state in any dimension [23,32].

Corollary 3. For any state $\rho \in PPT$ of dimension $2d_s \otimes 2d_s$ approximating private bit there is

$$\|\rho - \gamma_{2, d_s}\| \leq \epsilon \Rightarrow \|A_{01,10}^\Gamma\| \leq \frac{\epsilon}{2}. \quad (38)$$

The upper bound on the norm in Corollary 3 is tighter than the one in [23]. This is due to modification in the proof technique. This motivates us to assume $\sum_{i \neq j} \|A_{ij,ji}^\Gamma\| \leq \epsilon$, instead of $2 \sum_{i \neq j} \|A_{ij,ji}^\Gamma\| \leq \epsilon$ what would be analogous to assumption in proposition 1.

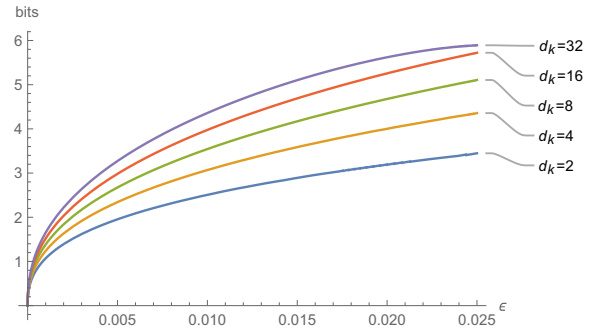


FIG. 11. Upper bounds on key repeater rate from proposition 2. We extract from condition $\frac{d_k-1}{d_s+d_k(d_k-1)} \leq \epsilon$ minimal value of $d_s = \lceil \frac{d_k-1-\epsilon d_k(d_k-1)}{\epsilon} \rceil$ yielding best value for upper bound.

Proposition 2. If ρ is a state with positive partial transpose approximates a strictly irreducible private dit (pdit) $\|\rho - \gamma_{(d_k, d_s)}\| \leq \epsilon$ for $\frac{d_k-1}{d_s+d_k(d_k-1)} \leq \epsilon$, $\sum_{i \neq j} \|A_{ij,ji}^\Gamma\| \leq \epsilon$, and conditional shield states of ρ are separable, then its two-way repeater rate $R^{\leftrightarrow}(\rho)$ is upper bounded as follows:

$$R^{\leftrightarrow}(\rho) \leq 2(\sqrt{\epsilon} + \epsilon) \log_2 \dim_H(\rho) + (1 + 2\sqrt{\epsilon} + 2\epsilon)h\left(\frac{\sqrt{\epsilon} + \epsilon}{\frac{1}{2} + \sqrt{\epsilon} + \epsilon}\right). \quad (39)$$

For the proof of the above proposition see Appendix. It is easy to notice that the upper bound in proposition 2 evaluated for pbits is tighter than the corresponding one from proposition 1. This is because with slightly different assumption on $A_{ij,ji}$ blocks. For upper bounds on the key repeater rate for exemplary dimensions of the key part provided in proposition 2, see Fig. 11.

Theorem 5. If a state with positive partial transpose ρ approximates strictly irreducible private dit (pdit) $\|\rho - \gamma_{(d_k, d_s)}\| \leq \epsilon$ for $\frac{d_k-1}{d_s+d_k(d_k-1)} \leq \epsilon < \frac{1}{d_k}$, $\sum_{i \neq j} \|A_{ij,ji}^\Gamma\| \leq \epsilon$, and its conditional shield states are separable, then it serves as a two-way (θ, η) -good secure network scheme S_ρ with degree Δ , and its overhead is lower bounded with

$$V(S_\rho) \geq M(\rho) \left(1 - \frac{\epsilon}{2} - f(d_k, \epsilon)\right), \quad (40)$$

$$f(d_k, \epsilon) := \frac{\log_2 d_k + \left(1 + \frac{\epsilon}{2}\right)h\left(\frac{\frac{\epsilon}{2}}{1 + \frac{\epsilon}{2}}\right)}{\log_2 d_k + \log_2 \left(\frac{d_k-1}{\epsilon}\right) + \log_2(1 - \epsilon d_k)}, \quad (41)$$

with $\eta = \log_2 d_k - 8\epsilon \log_2 d_k - 4h(\epsilon)$ [where $h(\cdot)$ is the binary Shannon entropy] and $\theta = 2(\sqrt{\epsilon} + \epsilon) \log_2 \dim_H(\rho) + (1 + 2\sqrt{\epsilon} + 2\epsilon)h\left(\frac{\sqrt{\epsilon} + \epsilon}{\frac{1}{2} + \sqrt{\epsilon} + \epsilon}\right)$.

For the proof of the above theorem see Appendix. For the lower bound on percentage of memory overhead from this theorem for different values of d_k see Fig. 12, while lower bounds on gap between η and θ for the presented scheme are depicted in Fig. 13.

A. On relaxation of the honest-but-curious attack assumption

In this section, we discuss in more detail a possible relaxation of the assumption 5 (about the honest-but-curious

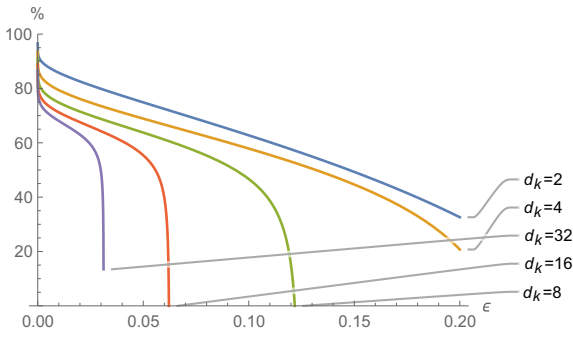


FIG. 12. Plots of lower bound on percentage of memory overhead from theorem 5, with different values of d_k .

attack). We argue that one can combine our countermeasure with the protection of the data and/or key of the node.

Indeed, there two major attacks that are in due. The first aims to learn the key of some of the links of the hub with end-users. The second concentrates on the direct learning of classical data stored in the hub’s server. One can also consider a mixed strategy aiming at learning both types of data. We note that in order to secure any type of data, one can perform the quantum one-time-pad [34], i.e., rotate each qubit randomly by one of the four Pauli operations. However, this type of encryption does not allow for manipulating the data by the honest party. It only shifts the problem of hacking to the place where the keys of the randomness are stored. A more clever solution involves the so-called *homomorphic encryption* [35,36]. This one aims at allowing to execute some quantum operation on encrypted data without its decryption. Such a solution can be therefore composed with our countermeasure. In particular the efficiency of our scheme (the gap between distillable and repeated key) after composition with the homomorphic encryption stays the same. Indeed, this encryption can be viewed as a local operation on the hub, while our scheme is secure against such operations. In turn, effectiveness against an LOCC protocol (tripartite or from the hub to the users) does not change. However, homomorphic encryption costs a non-negligible amount of

quantum memory. In the proposed solution for a number of protected qubits q gets enlarged to $a \times q$ where a is a natural number (constant of the solution). We can still measure the efficiency of such a complex solution with the introduced memory overhead. In this case, it reads

$$V(S'_\rho) = M'(\rho)(1 - D'(\rho)), \tag{42}$$

where $M'(\rho) = \Delta[a(\log_2 \dim_H(\rho))]$, $D'(\rho) = K_D(\rho)/[a(\log_2 \dim_H(\rho))]$. This implies that $V(S'_\rho) - V(S_\rho) = (a - 1) \log_2 \dim_H(\rho)$. The value of a can vary depending on a chosen protocol for homomorphic encryption. In Ref. [36], the value of a is modest, as it equals 3. We note here, that we count only the quantum memory of this solution, rather than classical, as the former is hard to be realized experimentally. In such an approach, the hub has to store classical keys needed for the homomorphic encryption in some separate memory, which is inaccessible to the hacker. However, the quantum part of encryption can be exposed to attacks against the reading of the data.

Finally, let us note that in the above countermeasure, the hacker, in spite of the impossibility of reading, can modify the data. Related countermeasure that can be composed with ours is the intrusion detection obtained recently via the so-called trap-codes [37]. In this case, the memory overhead can also report the cost of relaxation of the security of the data.

VII. CONCLUSIONS

In this manuscript, we have observed a particular attack on quantum network, and studied the quantum memory cost of its remedy—the hybrid quantum network. A common approach in designing quantum-secured Internet is to connect its nodes via pure entangled states or channels that distribute such entanglement. In this paper, we observe that this practice is not needed for a number of nodes of the Internet, and moreover, would open a threat.

As a case study of such a threat, we consider the possibility of performing entanglement swapping between the data basis of the hub and its two end-users Adam and Eve. We imagine that in future due to development of quantum technologies the link between each of them and the hub would be a quantum one. As a countermeasure, we propose to replace these links with those sharing/distributing bound entangled states which approximate private states. As for end-users, it is enough to communicate only classical information with the hub. What is more a functionality to pass a quantum state seems not only to be a redundant feature but also opens a gateway for a possible abuse.

While in the case of a maximally entangled state, one can generate 1 bit of key per 1 qubit of local memory, this is not the case for mixed entangled quantum states. We, therefore, study the memory cost of the proposed solution. We have introduced two notions: (i) that of a scheme (a choice of states shared by the node and users) and (ii) that of the memory overhead. The latter quantity reports how many qubits of the memory are not directly used up to generate key, but only assures security of its generation. We then focus on schemes that are represented by a single quantum state distributed in all the links. As the quality of the scheme, we propose the gap between the key that can be obtained from the state and the upper bound on the

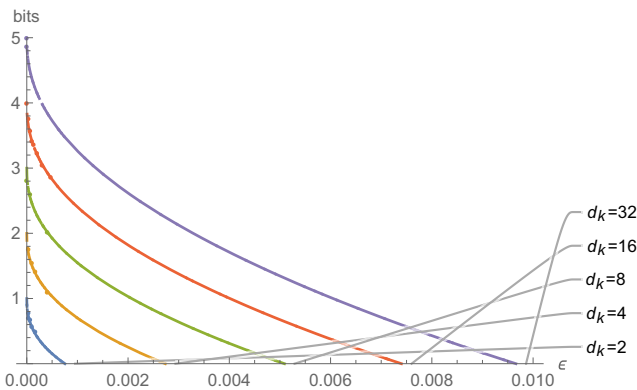


FIG. 13. Plots of lower bounds on gap between η and θ for the scheme in theorem 5. For obtaining possibly optimal value of upper bounds we attribute with d_s its minimal possible value of $d_s = \lceil \frac{d_k - 1 - \epsilon d_k (d_k - 1)}{\epsilon} \rceil$.

key that can be obtained via hacking. We called it a gap of the scheme.

We first focus on what is more or less straightforward to obtain from the well-established facts in entanglement theory based approach to quantum cryptography. This leads us to two different but asymptotically equivalent lower bounds for the memory overhead of the scheme. One is for private states, and the other for all quantum states. It implies that at least half of the memory of the scheme need to assist security of the scheme rather than can be turned to security itself.

We then consider particular bound entangled states as well as private states for which we know the construction of our proposal can be realized. These are PPT states that approximate private states and that are at the same time highly indistinguishable by LOPC operations from their attacked versions, which are separable. Although, in general, the overhead, in that case, is asymptotically 1, the convergence to 1 is modest.

The presented results allow to tune the exemplary states to the size of the gap of the scheme. As a byproduct, we have both sharpen the lower bounds on the distance between PPT and private bits and gave the first lower bound on this distance between PPT states and private dits for arbitrary dimension of the key part d_k . It would be then interesting to find the schemes based on private dits, rather than those that are based on tensor products of private bits.

Let us note here that we consider the attack to be honest-but curious. Both in the case of quantum repeater and the proposed hybrid repeater, the nodes can be hijacked, and in principle, the data can be traded via blackmail, and therefore, as we have discussed, should be kept e.g., homomorphically encrypted [35,36]. Finding the most effective scheme in terms of memory is an important open problem.

Finally, we admit, that another simple to consider solution for the considered threat, is to live with the fact of possibility of a malware and let every registered user of a node be connected with any other by quantum switch (no matter what is the type of the node) and sell e.g. utility. This, however, would need to be done at a certain price, in similarity to a utility that any smart phone can be turned into a network router within the price of the subscription. In general, one can ask for any other nontransitive property (nonhackable), that can be incorporated to provide security. That will be studied elsewhere (see in this context recent Ref. [38]).

While large effort to make QI happen is begin taken [7], it is also important to know a novel, inherently quantum threats that can come from the new quantum network design. To our knowledge this direction of research needs separate attention, as has not been studied in deep so far [11,12].

ACKNOWLEDGMENTS

K.H., O.S., and M.W. acknowledge grant Sonata Bis 5 (No. 2015/18/E/ST2/00327) from the National Science Center. K.H. and M.W. acknowledge partial support by the Foundation for Polish Science through IRAP project co-financed by EU within Smart Growth Operational Programme (Contract No. 2018/MAB/5).

APPENDIX

Proof of Observation 1.

$$E_D^{\rightarrow}(\gamma_{d_k, d_s}) \geq I_{\text{coh}}(AA')BB')_{\gamma_{d_k, d_s}} \quad (\text{A1})$$

$$= S(BB')_{\gamma_{d_k, d_s}} - S(AA'BB')_{\gamma_{d_k, d_s}} \quad (\text{A2})$$

$$= \log_2 d_k + \sum_i \frac{1}{d_k} S(\sigma_{iB'}) - S(\rho_{A'B'}) \quad (\text{A3})$$

$$= \log_2 d_k + \sum_i \frac{1}{d_k} [S(\text{Tr}_{A'} U_i \rho_{A'B'} U_i^\dagger) - S(U_i \rho_{A'B'} U_i^\dagger)] \quad (\text{A4})$$

$$= \log_2 d_k + \sum_i \frac{1}{d_k} [S(B') - S(A'B')]_{[U_i \rho_{A'B'} U_i^\dagger]} \quad (\text{A5})$$

$$= \log_2 d_k + \sum_i \frac{1}{d_k} I_{\text{coh}}(A')B')_{[U_i \rho_{A'B'} U_i^\dagger]} \quad (\text{A6})$$

$$= \log_2 d_k + \sum_i \frac{1}{d_k} I_{\text{coh}}(A')B')_{\sigma_i}, \quad (\text{A7})$$

where $\sigma_{iB'} = \text{Tr}_{A'} U_i \rho_{A'B'} U_i^\dagger$.

The first inequality is due to the fact that the one-way distillable entanglement is lower bounded by the coherent information. Then the first equality follows from direct calculation, and the fact that $S(\sum_i \frac{1}{d_k} |i\rangle\langle i| \otimes \text{Tr}_{A'} U_i \rho_{A'B'} U_i^\dagger) = \log_2 d_k + \sum_i \frac{1}{d_k} S(\text{Tr}_{A'} U_i \rho_{A'B'} U_i^\dagger)$. Equality $S(AA'BB') = S(\rho_{A'B'})$ comes from the construction: the private state is unitarily equivalent to $\psi \otimes \rho$ (where ψ is maximally entangled state of dimension d_k^2), and the entropy is invariant under unitary transformations, additive and zero for pure states. In the equality (A4) we add the unitary transformations to $\rho_{A'B'}$ which is assured by mentioned property of entropy: $S(\rho_{A'B'}) = \sum_i \frac{1}{d_k} S(U_i \rho_{A'B'} U_i^\dagger)$. We then observe that $S(B') - S(A'B')_{[U_i \rho_{A'B'} U_i^\dagger]}$ is nothing but the coherent information of σ_i for each i . Hence the final formula involves average value of the coherent information evaluated for states $\sigma_i \equiv U_i \rho_{A'B'} U_i^\dagger$. ■

Proof of theorem 1. Below we present a sequence of inequalities, that altogether allow to prove the theorem.

$$\theta \geq R^{\rightarrow}(\gamma_{d_k, d_s}) \geq E_D^{\rightarrow}(\gamma_{d_k, d_s}) \geq \log_2 d_k + \sum_i \frac{1}{d_k} I_{\text{coh}}(A')B')_{\sigma_i} \geq \log_2 d_k - \log_2 d_s, \quad (\text{A8})$$

The first inequality comes from our assumption that γ_{d_k, d_s} is an $(\theta, \log_2 d_k)$ -good one-way secure network scheme. The second inequality is supported by the fact, that one can distill $R^{\rightarrow}(\rho)$ singlets and use them for teleportation. One of methods to repeat key is to distill E_D^{\rightarrow} of pure entanglement between H and A and H and B, respectively. This is followed by entanglement swapping protocol [8]. The third inequality comes from Eq. (14). The final inequality is due to corollary 1.

Thanks to the above inequality (A8), we can upper bound the density of the private key as follows:

$$D(\gamma_{d_k, d_s}) = \frac{\log_2 d_k}{\log_2 \dim_H} = \frac{\log_2 d_k}{\log_2 d_k + \log_2 d_s} \quad (\text{A9})$$

$$\leq \frac{\log_2 d_k}{2 \log_2 d_k - \theta} = \frac{1}{2 - \frac{\theta}{\log_2 d_k}}. \quad (\text{A10})$$

From Eq. (11), we have

$$V(S_{\gamma_{d_k, d_s}}) \geq M(\gamma_{d_k, d_s}) \left(1 - \frac{1}{2 - \frac{\theta}{\log_2 d_k}} \right) \quad (\text{A11})$$

$$\approx M(\gamma_{d_k, d_s}) \left(\frac{1}{2} - \frac{\theta}{4 \log_2 d_k} + O(\theta^2) \right) \approx_{\theta \approx 0} \frac{1}{2} M(\gamma_{d_k, d_s}), \quad (\text{A12})$$

what ends the proof. \blacksquare

Proof of theorem 2. Because $\theta \geq R^{\leftrightarrow}(\rho) \geq E_D(\rho) \geq I_{\text{coh}}(H)A$, and it has nonpositive coherent information $I_{\text{coh}}(H)A$ [27], thus distillable key has to fulfill:

$$\begin{aligned} K_D(\rho) &\leq E_{\text{sq}}(\rho) \leq \frac{1}{2} I(\rho) = \frac{1}{2} S(H) + \frac{1}{2} I_{\text{coh}}(H)A \\ &\leq \frac{1}{2} (\log_2 d_k + \log_2 d_s) + \theta, \end{aligned} \quad (\text{A13})$$

where $E_{\text{sq}}(\rho_{A,B}) = \inf_{\rho^{ABE} \in S_{\text{Ext}}} \frac{1}{2} I(A; B|E)$ is the squashed entanglement [39], and the next inequality is by the definition of E_{sq} . Owing to the fact that $K_D(\rho) \geq \eta$, we obtain

$$\eta \leq \frac{1}{2} (\log_2 d_k + \log_2 d_s) + \theta, \quad (\text{A14})$$

$$\mathcal{D}(\rho) \leq \frac{\eta}{\log_2 d_H} \leq \frac{\frac{1}{2} \log_2 d_H + \theta}{\log_2 d_H} = \frac{1}{2} + \frac{\theta}{\log_2 d_H}, \quad (\text{A15})$$

$$V(S_\rho) \geq M(\rho) \left(\frac{1}{2} - \frac{\theta}{\log_2 d_H} \right) \approx_{\theta \approx 0} \frac{1}{2} M(\rho). \quad (\text{A16})$$

Proof of observation 2. The first inequality comes from the result of Christandl and Ferrara [19]. There is

$$R^{\rightarrow}(\gamma) \leq E_D^{\rightarrow}(\gamma \otimes \gamma). \quad (\text{A17})$$

The distillable entanglement is upper bounded by the log-negativity:

$$E_D^{\rightarrow}(\gamma \otimes \gamma) \leq \log_2 \|\gamma^\Gamma \otimes \gamma^\Gamma\| = 2 \log_2 \|\gamma^\Gamma\|, \quad (\text{A18})$$

where equality comes from the additivity of the log-negativity. We upper bound log-negativity as follows:

$$\|\gamma^\Gamma\| = \left\| \frac{1}{d_k} \sum_{ij} |ij\rangle \langle ji| \otimes X_{ij}^\Gamma \right\| \quad (\text{A19})$$

$$= \left\| \frac{1}{d_k} \sum_i |ii\rangle \langle ii| \otimes X_{ii}^\Gamma + \frac{1}{d_k} \sum_{i \neq j} |ij\rangle \langle ji| \otimes X_{ij}^\Gamma \right\| \quad (\text{A20})$$

$$\leq \left\| \frac{1}{d_k} \sum_i |ii\rangle \langle ii| \otimes X_{ii}^\Gamma \right\| + \left\| \frac{1}{d_k} \sum_{i \neq j} |ij\rangle \langle ji| \otimes X_{ij}^\Gamma \right\| \quad (\text{A21})$$

$$= 1 + \|\gamma^\Gamma - \hat{\gamma}^\Gamma\|. \quad (\text{A22})$$

The last equality is obtained due to the fact that $X_{ii} \in PPT$. Finally, because logarithm is strictly increasing, we have $2 \log_2 \|\gamma^\Gamma\| \leq 2 \log_2 (1 + \|\gamma^\Gamma - \hat{\gamma}^\Gamma\|)$, and hence

$$E_D^{\rightarrow}(\gamma \otimes \gamma) \leq 2 \log_2 (1 + \|\gamma^\Gamma - \hat{\gamma}^\Gamma\|). \quad (\text{A23})$$

This implies by virtue of Eq. (A17):

$$R^{\rightarrow}(\gamma) \leq 2 \log_2 (1 + \|\gamma^\Gamma - \hat{\gamma}^\Gamma\|). \quad (\text{A24})$$

Proof of observation 3. By direct calculations we have

$$\|\gamma^\Gamma - \hat{\gamma}^\Gamma\| \quad (\text{A25})$$

$$= \left\| \sum_{i,j} \frac{1}{d_k} |ij\rangle \langle ij| \otimes X_{ij}^\Gamma - \sum_i \frac{1}{d_k} |ii\rangle \langle ii| \otimes X_{ii}^\Gamma \right\| \quad (\text{A26})$$

$$= \left\| \sum_{i \neq j} \frac{1}{d_k} |ij\rangle \langle ij| \otimes X_{ij}^\Gamma \right\| \quad (\text{A27})$$

$$= \sum_{i \neq j} \frac{1}{d_k} \|X_{ij}^\Gamma\|. \quad (\text{A28})$$

Proof of lemma 1. A pdit γ_{d_k, d_s} has $d_k^2 - d_k$ off-diagonal block elements X_{ij} , and $\|X_{ij}\| \geq 0$. From observation 3 we have that $\sum_{i \neq j} \frac{1}{d_k} \|X_{ij}^\Gamma\| \leq \epsilon$, for some small $\epsilon \geq \|\gamma_{d_k, d_s}^\Gamma - \hat{\gamma}_{d_k, d_s}^\Gamma\|$. Then among those block elements there clearly has to be a one such that $\frac{1}{d_k} \|X_{i_0, j_0}^\Gamma\| \leq \frac{\epsilon}{d_k^2 - d_k}$ as a property of mean value, hence

$$\|X_{i_0, j_0}^\Gamma\| \leq \frac{\epsilon d_k}{d_k^2 - d_k} = \frac{\epsilon}{d_k - 1}. \quad (\text{A29})$$

We know from [23], that $\|X_{ij}\| \leq d_s \|X_{ij}^\Gamma\|$. Hence, for arbitrary i and j , we have

$$d_s \|X_{ij}^\Gamma\| \geq \|X_{ij}\| = \|U_i \sigma U_j^\dagger\| = \|\sigma\| = 1. \quad (\text{A30})$$

In particular $1 \leq d_s \|X_{i_0, j_0}^\Gamma\|$. Then $1 \leq d_s \frac{\epsilon}{d_k - 1}$ and finally $d_s \geq \frac{d_k - 1}{\epsilon}$. \blacksquare

Proof of theorem 3. From observation 2: $R^{\rightarrow}(\gamma_{(d_k, d_s)}) \leq 2 \log_2 (1 + \epsilon)$. Further from irreducibility of $\gamma_{(d_k, d_s)}$, we have that $K_D(\gamma_{(d_k, d_s)}) = \log_2 d_k$, and the lower bound for $V(S_{\gamma_{(d_k, d_s)}})$ we obtain in the following way

$$\begin{aligned} \mathcal{D}(\gamma_{(d_k, d_s)}) &= \frac{K_D(\gamma_{(d_k, d_s)})}{\log_2 d_k + \log_2 d_s} \\ &= \frac{\log_2 d_k}{\log_2 d_k + \log_2 d_s} \leq \frac{\log_2 d_k}{\log_2 d_k + \log_2 \frac{d_k - 1}{\epsilon}}. \end{aligned} \quad (\text{A31})$$

Thus

$$\begin{aligned} V(S_{\gamma_{(d_k, d_s)}}) &\geq M_{\gamma_{(d_k, d_s)}} \left(1 - \frac{\log_2 d_k}{\log_2 d_k + \log_2 \frac{d_k - 1}{\epsilon}} \right) \\ &\approx_{\epsilon \rightarrow 0} M(\gamma_{(d_k, d_s)}), \end{aligned} \quad (\text{A32})$$

where the first inequality is a consequence of lemma 1. \blacksquare

Proof of proposition 1. In this proof, partial transposition Γ and the operation of $\text{diag}(\cdot)$ are assumed to be evaluated

in computational basis. Furthermore we assume $\|A_{01,10}^\Gamma\| \leq \epsilon$. Using the results in Ref. [23], we know

$$\|\rho - \gamma_{(2,d_s)}\| \leq \epsilon \Rightarrow \|A_{0011}^\Gamma\| \leq \epsilon. \quad (\text{A33})$$

We define a projection

$$\Pi := (|00\rangle\langle 00| + |11\rangle\langle 11|) \otimes I_{d_s^2}. \quad (\text{A34})$$

Notice that $\Pi\gamma_{(2,d_s)}\Pi = \gamma_{(2,d_s)}$, and let us define subnormalized state

$$\rho_\Pi^\Gamma := \Pi\rho^\Gamma\Pi. \quad (\text{A35})$$

From one of assumptions, we have

$$\|\rho_\Pi^\Gamma - \text{diag}(\rho_\Pi^\Gamma)\| = 2\|A_{01,10}^\Gamma\| \leq 2\epsilon, \quad (\text{A36})$$

where $\text{diag}(\cdot)$ refers to an operation that projects the key part to its diagonal, i.e, it acts in the following way:

$$\text{diag}\left(\sum_{i,j,k,l} |ij\rangle\langle kl| \otimes A_{ij,kl}\right) := \sum_{i,j} |ij\rangle\langle ij| \otimes A_{ij,ij}. \quad (\text{A37})$$

We define a CPTP operation ϕ and corresponding Kraus operators:

$$\phi(\rho) := \sum_{i=1}^2 K_i\rho K_i^\dagger, \quad (\text{A38})$$

$$K_1 := |01\rangle\langle 01| \otimes I_{d_s^2}, \quad (\text{A39})$$

$$K_2 := |10\rangle\langle 10| \otimes I_{d_s^2}. \quad (\text{A40})$$

We employ the above to upper bound some traces of certain diagonal block of ρ^Γ . Since the trace norm is nonincreasing under CPTP maps and for $i \neq j$, we have $X_{ij,ij} = 0$:

$$\epsilon \geq \|\rho - \gamma_{(2,d_s)}\| \geq \|\phi(\rho) - \phi(\gamma_{(2,d_s)})\| \quad (\text{A41})$$

$$= \||01\rangle\langle 01| \otimes A_{01,01} + |10\rangle\langle 10| \otimes A_{10,10}\| \quad (\text{A42})$$

$$= \|A_{01,01}\| + \|A_{10,10}\| = \text{Tr}A_{01,01} + \text{Tr}A_{10,10} \quad (\text{A43})$$

$$= \text{Tr}A_{01,01}^\Gamma + \text{Tr}A_{10,10}^\Gamma, \quad (\text{A44})$$

where we used a property that the trace of hermitean positive semidefinite matrix is invariant under partial transpose. We use now Eqs. (A41)–(A44) to lower bound the following quantity:

$$\text{Tr}(\Pi\rho^\Gamma\Pi) = \text{Tr}A_{00,00}^\Gamma + \text{Tr}A_{11,11}^\Gamma \quad (\text{A45})$$

$$= 1 - \text{Tr}A_{01,01}^\Gamma + \text{Tr}A_{10,10}^\Gamma \geq 1 - \epsilon. \quad (\text{A46})$$

As a byproduct notice that

$$\|\rho_\Pi^\Gamma\| = \text{Tr}(\rho_\Pi^\Gamma) \equiv \text{Tr}(\Pi\rho^\Gamma\Pi) \geq 1 - \epsilon. \quad (\text{A47})$$

We employ now the ‘‘gentle measurement lemma’’ [40–42], saying that for all positive semidefinite operators σ , and $0 \leq H \leq 1$, one has

$$\|\sigma - \sqrt{H}\sigma\sqrt{H}\| \leq 2\sqrt{\text{Tr}(\sigma)}\sqrt{\text{Tr}(\sigma(I-H))}. \quad (\text{A48})$$

Since Π is a projector, and ρ^Γ is normalized, from Eqs. (A45),(A46), and (A48), we find

$$\|\rho^\Gamma - \rho_\Pi^\Gamma\| \leq 2\sqrt{1 - \text{Tr}(\Pi\rho^\Gamma\Pi)} \leq 2\sqrt{\epsilon}, \quad (\text{A49})$$

where we used cyclic property of the trace. Using the triangle inequality twice, the fact that $\|\rho_\Pi^\Gamma\| \equiv \|\text{diag}(\rho_\Pi^\Gamma)\|$, and inequalities in (A36), (A47), and (A49), we obtain

$$\left\|\rho^\Gamma - \frac{\text{diag}(\rho_\Pi^\Gamma)}{\|\rho_\Pi^\Gamma\|}\right\| \leq \|\rho^\Gamma - \text{diag}(\rho_\Pi^\Gamma)\| \quad (\text{A50})$$

$$+ \left\|\text{diag}(\rho_\Pi^\Gamma) - \frac{\text{diag}(\rho_\Pi^\Gamma)}{\|\rho_\Pi^\Gamma\|}\right\| \quad (\text{A51})$$

$$= \|\rho^\Gamma - \rho_\Pi^\Gamma + (\rho_\Pi^\Gamma - \text{diag}(\rho_\Pi^\Gamma))\| + (1 - \|\rho_\Pi^\Gamma\|) \quad (\text{A52})$$

$$\leq \|\rho^\Gamma - \rho_\Pi^\Gamma\| + \|\rho_\Pi^\Gamma - \text{diag}(\rho_\Pi^\Gamma)\| + \epsilon \quad (\text{A53})$$

$$\leq 2\sqrt{\epsilon} + 2\epsilon + \epsilon = 2\left(\sqrt{\epsilon} + \frac{3}{2}\epsilon\right). \quad (\text{A54})$$

From the Refs. [15,22], two-way repeater rate is upper bounded in the following way:

$$R^{\leftrightarrow}(\rho) \leq K_D(\rho^\Gamma) \leq E_r(\rho^\Gamma). \quad (\text{A55})$$

While employing asymptotic continuity of the relative entropy of entanglement E_r [43,44], we obtain

$$\left|E_r(\rho^\Gamma) - E_r\left(\frac{\text{diag}(\rho_\Pi^\Gamma)}{\|\rho_\Pi^\Gamma\|}\right)\right| \leq \xi \log_2 \dim_H(\rho^\Gamma) + (1 + \xi)h\left(\frac{\xi}{1 + \xi}\right) \quad (\text{A56})$$

$$\Rightarrow E_r(\rho^\Gamma) \leq E_r\left(\frac{\text{diag}(\rho_\Pi^\Gamma)}{\|\rho_\Pi^\Gamma\|}\right) + \xi \log_2 \dim_H(\rho^\Gamma) + (1 + \xi)h\left(\frac{\xi}{1 + \xi}\right), \quad (\text{A57})$$

where $\xi = 2(\sqrt{\epsilon} + \frac{3}{2}\epsilon)$. From Eq. (A55), we have then

$$R^{\leftrightarrow}(\rho) \leq E_r\left(\frac{\text{diag}(\rho_\Pi^\Gamma)}{\|\rho_\Pi^\Gamma\|}\right) + \xi \log_2 \dim_H(\rho^\Gamma) \quad (\text{A58})$$

$$+ (1 + \xi)h\left(\frac{\xi}{1 + \xi}\right). \quad (\text{A59})$$

Blocks of $\text{diag}(\rho^\Gamma)$ are separable by assumption. Since nonzero blocks of $\text{diag}(\rho_\Pi^\Gamma)$ are identical to corresponding blocks of $\text{diag}(\rho^\Gamma)$ they are also separable. This implies that the relative entropy of entanglement of $\frac{\text{diag}(\rho_\Pi^\Gamma)}{\|\rho_\Pi^\Gamma\|}$, from its definition reads 0. Knowing that $d_k = 2$ and that dimension of matrix is invariant under partial transpose, we obtain an upper bound.

$$R^{\leftrightarrow}(\rho) \leq 2\left(\sqrt{\epsilon} + \frac{3}{2}\epsilon\right)(1 + \log_2 d_s) \quad (\text{A60})$$

$$+ (1 + 2\sqrt{\epsilon} + 3\epsilon)h\left(\frac{2\sqrt{\epsilon} + 3\epsilon}{1 + 2\sqrt{\epsilon} + 3\epsilon}\right). \quad (\text{A61})$$

Proof of theorem 4. We work under an assumption that $\frac{1}{2(d_s+1)} \leq \|\rho - \gamma_{(2,d_s)}\| \leq \epsilon < \frac{1}{2}$. The first step is to upper bound key rate with relative entropy (see Ref. [22]):

$$K_D(\rho) \leq E_r(\rho). \quad (\text{A62})$$

Then we make use of asymptotic continuity of quantum relative entropy [43,44].

$$|E_r(\rho) - E_r(\gamma_{(2,d_s)})| \leq \frac{\epsilon}{2} \log_2 \dim_H(\rho) + \left(1 + \frac{\epsilon}{2}\right) h\left(\frac{\frac{\epsilon}{2}}{1 + \frac{\epsilon}{2}}\right) \quad (\text{A63})$$

$$\Rightarrow E_r(\rho) \leq E_r(\gamma_{(2,d_s)}) + \frac{\epsilon}{2} \log_2 \dim_H(\rho) + \left(1 + \frac{\epsilon}{2}\right) h\left(\frac{\frac{\epsilon}{2}}{1 + \frac{\epsilon}{2}}\right). \quad (\text{A64})$$

Since $E_r(\gamma_{(d_k,d_s)}) \leq \log_2 d_k$ [22], by combining Eqs. (A62) and (A64), we have

$$K_D(\rho) \leq \log_2 d_k + \frac{\epsilon}{2} \log_2 \dim_H \rho + \left(1 + \frac{\epsilon}{2}\right) h\left(\frac{\frac{\epsilon}{2}}{1 + \frac{\epsilon}{2}}\right). \quad (\text{A65})$$

From Ref. [23], we know that $\|\rho - \gamma_{(2,d_s)}\| \geq \frac{1}{2(d_s+1)}$, what together with the initial condition ($\epsilon < \frac{1}{2}$) yields

$$\log_2 d_s \geq \log_2 \left(\frac{1 - 2\epsilon}{2\epsilon}\right). \quad (\text{A66})$$

The overhead of the scheme is then lower bounded

$$V(\rho) = M(\rho) \left(1 - \frac{K_D^{\rightarrow}(\rho)}{\log_2 \dim_H(\rho)}\right) \quad (\text{A67})$$

$$\geq M(\rho) \left(1 - \frac{\log_2 d_k + \frac{\epsilon}{2} \log_2 \dim_H(\rho) + \left(1 + \frac{\epsilon}{2}\right) h\left(\frac{\frac{\epsilon}{2}}{1 + \frac{\epsilon}{2}}\right)}{\log_2 \dim_H(\rho)}\right) \quad (\text{A68})$$

$$\geq M(\rho) \left(1 - \frac{\log_2 d_k + \left(1 + \frac{\epsilon}{2}\right) h\left(\frac{\frac{\epsilon}{2}}{1 + \frac{\epsilon}{2}}\right)}{\log_2 d_k + \log_2 \left(\frac{1-2\epsilon}{2\epsilon}\right)} - \frac{\epsilon}{2}\right) \quad (\text{A69})$$

$$= M(\rho) \left(1 - \frac{1 + \left(1 + \frac{\epsilon}{2}\right) h\left(\frac{\frac{\epsilon}{2}}{1 + \frac{\epsilon}{2}}\right)}{1 + \log_2 \left(\frac{1-2\epsilon}{2\epsilon}\right)} - \frac{\epsilon}{2}\right). \quad (\text{A70})$$

Where we used $\dim_H(\rho) = \dim_H(\gamma_{(2,d_s)})$ and $d_k = 2$.

Now we have to design an appropriate lower bound on K_D . Following arguments of Ref. [45], the operation of privacy squeezing does not increase the trace distance $\|\rho^{\text{ps}} - \gamma_{(2,d_s)}^{\text{ps}}\| \leq \epsilon$. Moreover after this operation private state (strictly irreducible in this case) turns into one of the two Bell states $\gamma_{(2,d_s)}^{\text{ps}} \equiv \psi$. In general, the following inequalities hold:

$$K_D^{\rightarrow}(\rho^{\text{ps}}) \leq K_D^{\rightarrow}(\rho) \leq K_D(\rho). \quad (\text{A71})$$

On the other hand due to lemma V.3. in Ref. [46], both one-way and two-way key rates are lower bounded with

$$1 - 8\epsilon \log_2 \dim_H(\gamma_{(2,d_s)}^{\text{ps}}) - 4h(\epsilon) \leq K_D^{\rightarrow}(\rho^{\text{ps}}). \quad (\text{A72})$$

From Eqs. (A71) and (A72), and the fact $\dim_H(\rho^{\text{ps}}) = 2$, we obtain

$$K_D(\rho) \geq \eta := 1 - 8\epsilon - 4h(\epsilon). \quad (\text{A73})$$

Form proposition 1, the rate of the repeater is upper bounded with

$$R^{\leftrightarrow}(\rho) \leq \theta := 2\left(\sqrt{\epsilon} + \frac{3}{2}\epsilon\right)(1 + \log_2 d_s) \quad (\text{A74})$$

$$+ (1 + 2\sqrt{\epsilon} + 3\epsilon)h\left(\frac{2\sqrt{\epsilon} + 3\epsilon}{1 + 2\sqrt{\epsilon} + 3\epsilon}\right). \quad (\text{A75})$$

Notation 6. We denote projectors $P_{i,j}$ and P_i for $i \neq j$

$$P_{i,j} = |ii\rangle\langle ii| + |jj\rangle\langle jj|, \quad (\text{A76})$$

$$P_i = |ii\rangle\langle ii|. \quad (\text{A77})$$

The following identities hold.

Fact 1. We have the following identities:

$$(P_{i,j} \otimes I) \left(\sum_{ijkl} |ij\rangle\langle kl| \otimes A_{ij,kl} \right) (P_{i,j} \otimes I) \quad (\text{A78})$$

$$= |ii\rangle\langle ii| \otimes A_{ii,ii} + |ii\rangle\langle jj| \otimes A_{ii,jj} \quad (\text{A79})$$

$$+ |jj\rangle\langle ii| \otimes A_{jj,ii} + |jj\rangle\langle jj| \otimes A_{jj,jj} \quad (\text{A80})$$

and also

$$(P_i \otimes I) \left(\sum_{ijkl} |ij\rangle\langle kl| \otimes A_{ij,kl} \right) (P_i \otimes I) \quad (\text{A81})$$

$$= |ii\rangle\langle ii| \otimes A_{ii,ii}. \quad (\text{A82})$$

Notation 7. For the proofs of observation 4 we abuse the notation denoting $\frac{1}{d_k} X_{ii,jj} \rightarrow X_{ii,jj}$, $P_{i,j} \otimes I \rightarrow P_{i,j}$, and $P_i \otimes I \rightarrow P_i$ for conciseness.

Proof of observation 4. We start with proving first inequality (34). Using the contractivity of the trace norm, we have

$$\|\rho - \gamma\| \leq \epsilon \Rightarrow \|P_{i,j}\rho P_{i,j} - P_{i,j}\gamma P_{i,j}\| \leq \epsilon. \quad (\text{A83})$$

Thus

$$\| |ii\rangle\langle ii| \otimes (A_{ii,ii} - X_{ii,ii}) \quad (\text{A84})$$

$$+ |ii\rangle\langle jj| \otimes (A_{ii,jj} - X_{ii,jj}) \quad (\text{A85})$$

$$+ |jj\rangle\langle ii| \otimes (A_{jj,ii} - X_{jj,ii})$$

$$+ |jj\rangle\langle jj| \otimes (A_{jj,jj} - X_{jj,jj}) \| \leq \epsilon. \quad (\text{A86})$$

Using again the norm contractivity property and projector P_i , we have

$$\epsilon \geq \|\rho - \gamma\| \geq \|P_i \rho P_i - P_i \gamma P_i\| \quad (\text{A87})$$

$$= \sum_i \|A_{ii,ii} - X_{ii,ii}\| \geq \|A_{ii,ii} - X_{ii,ii}\|. \quad (\text{A88})$$

Now we want to prove that

$$\|A_{ii,jj} - X_{ii,jj}\| \leq \epsilon. \quad (\text{A89})$$

Let us express the matrix from LHS of (A86) as follows:

$$M = D + \hat{A}, \quad (\text{A90})$$

where M is a matrix, D are diagonal elements and \hat{A} are anti-diagonal elements. Note that $\|D\| \leq \epsilon$ as $\|M\| \leq \epsilon$.

We get then

$$\|M\| = \|D + \hat{A}\| \geq \|D\| - \|\hat{A}\| \tag{A91}$$

$$\Rightarrow \|\hat{A}\| \leq \|M\| + \|D\| \leq \|M\| + \epsilon \leq 2\epsilon. \tag{A92}$$

We note then that

$$\|\hat{A}\| = \left\| \begin{matrix} 0 & \hat{A}_{ii,jj} \\ \hat{A}_{ii,jj}^\dagger & 0 \end{matrix} \right\| = 2\|\hat{A}_{ii,jj}\|,$$

hence

$$\|\hat{A}\| = 2\|A_{ii,jj} - X_{ii,jj}\| \leq 2\epsilon, \tag{A93}$$

$$\|A_{ii,jj} - X_{ii,jj}\| \leq \epsilon. \tag{A94}$$

Finally, applying the reverse triangle inequality to Eq. (A94) and having $\|X_{ii,jj}\| = \frac{1}{d_k}$,

$$\|A_{ii,jj}\| - \frac{1}{d_k} \leq \epsilon \Rightarrow \|A_{ii,jj}\| \geq \frac{1}{d_k} - \epsilon. \tag{A95}$$

Now we prove the second inequality (35). Consider an incomplete von Neumann measurement

$$\{K_{ij}\} = \{|ij\rangle\langle ij| \otimes I\}. \tag{A96}$$

Using $\|\rho - \gamma\| \leq \epsilon$ and contractivity of norm, we obtain

$$\left\| \sum_{ij} K_{ij} \rho K_{ij}^\dagger - \sum_{ij} K_{ij} \gamma K_{ij}^\dagger \right\| \leq \epsilon, \tag{A97}$$

$$\left\| \sum_{ij} |ij\rangle\langle ij| \otimes A_{ij,ij} - \sum_i |ii\rangle\langle ii| \otimes X_{ii,ii} \right\| \leq \epsilon. \tag{A98}$$

For $i \neq j$ let $X_{ij,ij} = 0$, then

$$\left\| \sum_{ij} |ij\rangle\langle ij| \otimes A_{ij,ij} - \sum_{ij} |ij\rangle\langle ij| \otimes X_{ij,ij} \right\| \leq \epsilon, \tag{A99}$$

$$\sum_{ij} \||ij\rangle\langle ij| \otimes (A_{ij,ij} - X_{ij,ij})\| \leq \epsilon. \tag{A100}$$

$$\sum_{i \neq j} \|A_{ij,ij} - X_{ij,ij}\| + \sum_i \|A_{ii,ii} - X_{ii,ii}\| \leq \epsilon. \tag{A101}$$

Employing the aforementioned condition that $X_{ij,ij}$ vanish and non-negativity of the trace norm, we obtain

$$\sum_{i \neq j} \|A_{ij,ij}\| \leq \epsilon. \tag{A102}$$

Proof of lemma 2. We know that $\rho^\Gamma \geq 0$. Firstly we construct, a projector on certain $2 \times d_s$ dimensional subspace of $\rho^\Gamma \geq 0$.

$$\Pi_0 = (|ij\rangle\langle ij| + |ji\rangle\langle ji|) \otimes I_{d_s^2}, \quad i \neq j. \tag{A103}$$

Having in mind that $\rho = \sum_{ijkl} |ij\rangle\langle kl| \otimes A_{ij,kl}$, we perform the projection and obtain

$$\Pi_0 \rho^\Gamma \Pi_0 = \begin{bmatrix} A_{ij,ij}^\Gamma & 0 & 0 & A_{ii,jj}^\Gamma \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ A_{ii,jj}^\Gamma & 0 & 0 & A_{ji,ji}^\Gamma \end{bmatrix} \geq 0, \tag{A104}$$

where we used that $A_{jj,ii}^\Gamma = (A_{ii,jj}^\Gamma)^\dagger$, what is a consequence of ρ^Γ being Hermitian. Indeed $\Pi_0 \rho^\Gamma \Pi_0$ is positive semidefinite since Π_0 is a Kraus operator. In what follows, we construct a unitary transformation based on singular value decomposition of $A_{ii,jj}^\Gamma = S \Sigma V$.

$$U = |ij\rangle\langle ij| \otimes S^\dagger + |ji\rangle\langle ji| \otimes V. \tag{A105}$$

Note that $\text{Tr} \Sigma = \|A_{ii,jj}^\Gamma\|$. In the next step, we perform a specific privacy squeezing operation on ρ^Γ :

$$\rho_{\text{ps.}}^\Gamma = \text{Tr}_{A'B'} U \Pi_0 \rho^\Gamma \Pi_0 U^\dagger. \tag{A106}$$

What yields following form of a privacy squeezed matrix, which is positive semidefinite,

$$\rho_{\text{ps.}}^\Gamma = \begin{bmatrix} \|A_{ij,ij}\| & 0 & 0 & \|A_{ii,jj}^\Gamma\| \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \|A_{ii,jj}^\Gamma\| & 0 & 0 & \|A_{ji,ji}\| \end{bmatrix} \geq 0. \tag{A107}$$

Where we used a property of diagonal blocks $\|A_{ij,ij}\| = \text{Tr} A_{ij,ij}^\Gamma = \text{Tr} A_{ij,ij}^\Gamma = \|A_{ij,ij}^\Gamma\|$. Using a basic fact known for positive matrices we have the following dependence between its elements:

$$\|A_{ii,jj}^\Gamma\| \leq \frac{\|A_{ij,ij}\| + \|A_{ji,ji}\|}{2}. \tag{A108}$$

Now we are going to use observation 4. Since the smallest component of the sum is always smaller than an average, we have

$$2\epsilon \geq 2 \sum_{i \neq j} \|A_{ij,ij}\| = \sum_{i \neq j} (\|A_{ij,ij}\| + \|A_{ji,ji}\|) \tag{A109}$$

$$= d_k(d_k - 1) \sum_{i \neq j} \frac{(\|A_{ij,ij}\| + \|A_{ji,ji}\|)}{d_k(d_k - 1)} \tag{A110}$$

$$\geq d_k(d_k - 1) \min_{i \neq j} (\|A_{ij,ij}\| + \|A_{ji,ji}\|). \tag{A111}$$

Since Eq. (A108) is true for all $i \neq j$, we use the smallest element denoted with $i_0 \neq j_0$. Hence form (A108),

$$\|A_{i_0 i_0, j_0 j_0}^\Gamma\| \leq \frac{\epsilon}{d_k(d_k - 1)}. \tag{A112}$$

By observation 4, $\forall_{i \neq j}$ we have $\|A_{ii,jj}\| \geq \frac{1}{d_k} - \epsilon$ and $\|A_{i_0 i_0, j_0 j_0}^\Gamma\| \leq \frac{\epsilon}{d_k^2 - d_k}$. Owing to the fact that under partial transposition the trace norm can not increase by more than the dimension of the matrix (here d_s) [23], we have

$$\frac{1}{d_k} - \epsilon \leq \|A_{i_0 i_0, j_0 j_0}\| \leq d_s \|A_{i_0 i_0, j_0 j_0}^\Gamma\| \leq \frac{\epsilon}{d_k^2 - d_k} d_s, \tag{A113}$$

thus,

$$1 - \epsilon d_k \leq \frac{d_k \epsilon}{d_k^2 - d_k} d_s = \frac{\epsilon}{d_k - 1} d_s, \tag{A114}$$

and finally,

$$d_s \geq \left(\frac{d_k - 1}{\epsilon} \right) (1 - \epsilon d_k). \tag{A115}$$

Proof of corollary 2. The proof is straightforward consequence of lemma 2. Since the implication stated in Lemma

2 is true for any ϵ that $\epsilon \geq \|\rho - \gamma\|$, we denote with ϵ_0 the one that saturates it. We have the following implication:

$$\|\rho - \gamma_{d_k, d_s}\| = \epsilon_0 \Rightarrow d_s \geq \left(\frac{d_k - 1}{\epsilon_0} \right) (1 - \epsilon_0 d_k). \quad (\text{A116})$$

We immediately obtain the following lower bound:

$$\|\rho - \gamma_{d_k, d_s}\| \geq \frac{d_k - 1}{d_s + d_k(d_k - 1)}. \quad (\text{A117})$$

Proof of corollary 3. We notice that equation (A112) is true also for $d_k = 2$. Since in this dimension there is only a single choice of $i_0 \neq j_0$ (up to Hermitian conjugate), we have:

$$\|A_{00,11}^\Gamma\| \leq \frac{\epsilon}{d_k(d_k - 1)} = \frac{\epsilon}{2}. \quad (\text{A118})$$

Proof of proposition 2. This proof follows the same steps as the proof of proposition 1. Partial transposition Γ and the operation of $\text{diag}(\cdot)$ are assumed to be evaluated in computational basis. Furthermore we assume that $\sum_{i \neq j} \|A_{ij,ji}^\Gamma\| \leq \epsilon$. We work under an assumption that $\frac{d_k - 1}{d_s + d_k(d_k - 1)} \leq \|\rho - \gamma_{(d_k, d_s)}\| \leq \epsilon < \frac{1}{d_k}$.

We define a projection and subnormalized state ρ_Π^Γ ,

$$\Pi := \sum_{i=0}^{d_k-1} |ii\rangle\langle ii| \otimes I_{d_s^2}, \quad (\text{A119})$$

$$\rho_\Pi^\Gamma := \Pi \rho^\Gamma \Pi. \quad (\text{A120})$$

We notice then that

$$\|\rho_\Pi^\Gamma - \text{diag}(\rho_\Pi^\Gamma)\| = \sum_{i \neq j} \|A_{ij,ji}^\Gamma\| \leq \epsilon, \quad (\text{A121})$$

Where operation of $\text{diag}(\cdot)$ is defined in Eq. (A37).

We anticipate now and calculate the following quantity using equation (A102) again:

$$\text{Tr}(\Pi \rho^\Gamma \Pi) = \sum_{i=0}^{d_k-1} A_{ii,ii}^\Gamma \quad (\text{A122})$$

$$= \sum_{i=0}^{d_k-1} A_{ii,ii} = 1 - \sum_{i \neq j} A_{ij,ij} \geq 1 - \epsilon. \quad (\text{A123})$$

As a byproduct we notice that

$$\|\rho_\Pi^\Gamma\| = \text{Tr}(\rho_\Pi^\Gamma) = \text{Tr}(\Pi \rho^\Gamma \Pi) \geq 1 - \epsilon. \quad (\text{A124})$$

We employ now the ‘‘gentle measurement lemma’’ [40–42], saying that for all positive semidefinite operators σ , and $0 \leq H \leq 1$, one has

$$\|\sigma - \sqrt{H} \sigma \sqrt{H}\| \leq 2\sqrt{\text{Tr}(\sigma)} \sqrt{\text{Tr}(\sigma(I - H))}. \quad (\text{A125})$$

Since Π is a projector, and ρ^Γ is normalized, from Eqs. (A122), (A123), and (A125), we find

$$\|\rho^\Gamma - \rho_\Pi^\Gamma\| \leq 2\sqrt{1 - \text{Tr}(\Pi \rho^\Gamma \Pi)} \leq 2\sqrt{\epsilon}, \quad (\text{A126})$$

where we used cyclic property of the trace. Using the triangle inequality twice, the fact that $\|\rho_\Pi^\Gamma\| \equiv \|\text{diag}(\rho_\Pi^\Gamma)\|$, and

inequalities in Eqs. (A121), (A124), and (A126):

$$\left\| \rho^\Gamma - \frac{\text{diag}(\rho_\Pi^\Gamma)}{\|\rho_\Pi^\Gamma\|} \right\| \leq \|\rho^\Gamma - \text{diag}(\rho_\Pi^\Gamma)\| \quad (\text{A127})$$

$$+ \left\| \text{diag}(\rho_\Pi^\Gamma) - \frac{\text{diag}(\rho_\Pi^\Gamma)}{\|\rho_\Pi^\Gamma\|} \right\| \quad (\text{A128})$$

$$= \|\rho^\Gamma - \rho_\Pi^\Gamma + (\rho_\Pi^\Gamma - \text{diag}(\rho_\Pi^\Gamma))\| + (1 - \|\rho_\Pi^\Gamma\|) \quad (\text{A129})$$

$$\leq \|\rho^\Gamma - \rho_\Pi^\Gamma\| + \|\rho_\Pi^\Gamma - \text{diag}(\rho_\Pi^\Gamma)\| + \epsilon \quad (\text{A130})$$

$$\leq 2\sqrt{\epsilon} + \epsilon + \epsilon = 2(\sqrt{\epsilon} + \epsilon). \quad (\text{A131})$$

This upper bound is tighter than the corresponding one for a pbrit from proposition 1 due to application of corollary 3.

From the Refs. [15,22], the two-way repeater rate is upper bounded in the following way:

$$R^{\leftrightarrow}(\rho) \leq K_D(\rho^\Gamma) \leq E_r(\rho^\Gamma). \quad (\text{A132})$$

While employing asymptotic continuity of the relative entropy of entanglement E_r [43,44], we obtain

$$\left| E_r(\rho^\Gamma) - E_r\left(\frac{\text{diag}(\rho_\Pi^\Gamma)}{\|\rho_\Pi^\Gamma\|}\right) \right| \leq \xi \log_2 \dim_H(\rho^\Gamma) + (1 + \xi)h\left(\frac{\xi}{1 + \xi}\right) \quad (\text{A133})$$

$$\Rightarrow E_r(\rho^\Gamma) \leq E_r\left(\frac{\text{diag}(\rho_\Pi^\Gamma)}{\|\rho_\Pi^\Gamma\|}\right) + \xi \log_2 \dim_H(\rho^\Gamma) + (1 + \xi)h\left(\frac{\xi}{1 + \xi}\right), \quad (\text{A134})$$

where $\xi = 2(\sqrt{\epsilon} + \epsilon)$. Since dimension of a matrix is invariant under the partial transpose we have now:

$$R^{\leftrightarrow}(\rho) \leq E_r\left(\frac{\text{diag}(\rho_\Pi^\Gamma)}{\|\rho_\Pi^\Gamma\|}\right) + \xi \log_2 \dim_H(\rho) \quad (\text{A135})$$

$$+ (1 + \xi)h\left(\frac{\xi}{1 + \xi}\right). \quad (\text{A136})$$

Blocks of $\text{diag}(\rho^\Gamma)$ are separable from assumption. Since nonzero blocks $\text{diag}(\rho_\Pi^\Gamma)$ are identical to corresponding blocks of $\text{diag}(\rho^\Gamma)$ they are also separable. This implies that the relative entropy of entanglement of $\frac{\text{diag}(\rho_\Pi^\Gamma)}{\|\rho_\Pi^\Gamma\|}$, from its definition reads 0, hence

$$R^{\leftrightarrow}(\rho) \leq 2(\sqrt{\epsilon} + \epsilon) \dim_H(\rho) \quad (\text{A137})$$

$$+ (1 + 2\sqrt{\epsilon} + 2\epsilon)h\left(\frac{\sqrt{\epsilon} + \epsilon}{\frac{1}{2} + \sqrt{\epsilon} + \epsilon}\right). \quad (\text{A138})$$

Proof of theorem 5. We work under assumption that $\frac{d_k - 1}{d_s + d_k(d_k - 1)} \leq \|\rho - \gamma_{(d_k, d_s)}\| \leq \epsilon < \frac{1}{d_k}$.

The first step is to upper bound key rate with relative entropy (see Ref. [22]).

$$K_D(\rho) \leq E_r(\rho). \quad (\text{A139})$$

Then we make use of asymptotic continuity of quantum relative entropy [43,44].

$$\begin{aligned} |E_r(\rho) - E_r(\gamma_{(d_k, d_s)})| &\leq \frac{\epsilon}{2} \log_2 \dim_H(\rho) \\ &+ \left(1 + \frac{\epsilon}{2}\right) h\left(\frac{\frac{\epsilon}{2}}{1 + \frac{\epsilon}{2}}\right) \end{aligned} \quad (\text{A140})$$

$$\begin{aligned} \Rightarrow E_r(\rho) &\leq E_r(\gamma_{(d_k, d_s)}) + \frac{\epsilon}{2} \log_2 \dim_H(\rho) \\ &+ \left(1 + \frac{\epsilon}{2}\right) h\left(\frac{\frac{\epsilon}{2}}{1 + \frac{\epsilon}{2}}\right). \end{aligned} \quad (\text{A141})$$

Since $E_r(\gamma_{(d_k, d_s)}) \leq \log_2 d_k$ [22], by combining Eqs. (A139) and (A141), we have:

$$K_D(\rho) \leq \log_2 d_k + \frac{\epsilon}{2} \log_2 \dim_H \rho + \left(1 + \frac{\epsilon}{2}\right) h\left(\frac{\frac{\epsilon}{2}}{1 + \frac{\epsilon}{2}}\right). \quad (\text{A142})$$

From Lemma 2, we know that $d_s \geq (\frac{d_k-1}{\epsilon})(1 - \epsilon d_k)$. We assume RHS to be positive, which together with the initial condition yields:

$$\log_2 d_s \geq \log_2 \left(\frac{d_k - 1}{\epsilon}\right) + \log_2(1 - \epsilon d_k). \quad (\text{A143})$$

The overhead of the scheme is then lower bounded as follows:

$$\begin{aligned} V(\rho) &= M(\rho) \left(1 - \frac{K_D(\rho)}{\log_2 \dim_H(\rho)}\right) \\ &\geq M(\rho) \left(1 - \frac{\log_2 d_k + \frac{\epsilon}{2} \log_2 \dim_H(\rho) + \left(1 + \frac{\epsilon}{2}\right) h\left(\frac{\frac{\epsilon}{2}}{1 + \frac{\epsilon}{2}}\right)}{\log_2 \dim_H(\rho)}\right) \end{aligned} \quad (\text{A144})$$

$$\quad (\text{A145})$$

$$= M(\rho) \left(1 - \frac{\log_2 d_k + \left(1 + \frac{\epsilon}{2}\right) h\left(\frac{\frac{\epsilon}{2}}{1 + \frac{\epsilon}{2}}\right) - \frac{\epsilon}{2}}{\log_2 d_k + \log_2 d_s} - \frac{\epsilon}{2}\right) \quad (\text{A146})$$

$$\geq M(\rho) \left(1 - \frac{\log_2 d_k + \left(1 + \frac{\epsilon}{2}\right) h\left(\frac{\frac{\epsilon}{2}}{1 + \frac{\epsilon}{2}}\right)}{\log_2 d_k + \log_2 \left(\frac{d_k-1}{\epsilon}\right) + \log_2(1 - \epsilon d_k)} - \frac{\epsilon}{2}\right). \quad (\text{A147})$$

Now we have to find an appropriate lower bound on K_D . Following arguments of Ref. [45] the operation of privacy squeezing does not increase the trace distance $\|\rho^{\text{ps}} - \gamma_{(d_k, d_s)}^{\text{ps}}\| \leq \epsilon$, in a similar manner the key rate $K_D^{\rightarrow}(\rho^{\text{ps}}) \leq K_D^{\rightarrow}(\rho)$. Moreover after this operation private state (strictly irreducible in that case) turns into maximally entangled state of dimension d_k^2 .

$$K_D^{\rightarrow}(\rho^{\text{ps}}) \leq K_D^{\rightarrow}(\rho) \leq K_D(\rho). \quad (\text{A148})$$

On the other hand due to results in Ref. [46] and the fact that $K_D^{\rightarrow}(\rho^{\text{ps}}) = \log_2 \dim_H(\rho^{\text{ps}})$ both one-way and two-way keys are lower bounded

$$\log_2 d_k - 8\epsilon \log_2 \dim_H(\gamma_{(d_k, d_s)}^{\text{ps}}) - 4h(\epsilon) \leq K_D^{\rightarrow}(\rho^{\text{ps}}). \quad (\text{A149})$$

From Eqs. (A148) and (A149), and the fact $\dim_H(\gamma_{(d_k, d_s)}^{\text{ps}}) = d_k$, we obtain

$$K_D(\rho) \geq \eta := \log_2 d_k - 8\epsilon \log_2 d_k - 4h(\epsilon). \quad (\text{A150})$$

Form proposition 2, the key repeater rate is upper bounded with

$$\begin{aligned} R^{\leftrightarrow}(\rho) &\leq \theta := 2(\sqrt{\epsilon} + \epsilon) \dim_H(\rho) \\ &+ (1 + 2\sqrt{\epsilon} + 2\epsilon) h\left(\frac{\sqrt{\epsilon} + \epsilon}{\frac{1}{2} + \sqrt{\epsilon} + \epsilon}\right). \end{aligned} \quad (\text{A151})$$

$$\quad (\text{A152})$$

■

[1] S. Wiesner, *Sigact News* **15**, 78 (1983).
 [2] C. H. Bennett and G. Brassard, in *Proceedings of the IEEE International Conference on Computers, Systems and Signal Processing* (IEEE Computer Society Press, New York, Bangalore, India, 1984), pp. 175–179.
 [3] J. Preskill, *Quantum* **2**, 79 (2018).
 [4] W. Dür, H.-J. Briegel, J. I. Cirac, and P. Zoller, *Phys. Rev. A* **59**, 169 (1999).
 [5] S. Muralidharan, L. Li, J. Kim, N. Lütkenhaus, M. D. Lukin, and L. Jiang, *Sci. Rep.* **6**, 20463 (2016).
 [6] M. Zwerger, A. Pirker, V. Dunjko, H. J. Briegel, and W. Dür, *Phys. Rev. Lett.* **120**, 030503 (2018).
 [7] S. Wehner, D. Elkouss, and R. Hanson, *Science* **362**, eaam9288 (2018).
 [8] M. Żukowski, A. Zeilinger, M. A. Horne, and A. K. Ekert, *Phys. Rev. Lett.* **71**, 4287 (1993).
 [9] C. H. Bennett, G. Brassard, S. Popescu, B. Schumacher, J. A. Smolin, and W. K. Wootters, *Phys. Rev. Lett.* **76**, 722 (1996).
 [10] W. K. Wootters and W. H. Zurek, *Nature (London)* **299**, 802 (1982).
 [11] S. Suzuki and R. V. Meter, in *Proceedings 2015 Workshop on Security of Emerging Networking Technologies, San Diego, California* (Internet Society, 2015).
 [12] L.-A. Wu and D. A. Lidar, *Quant. Info. Process.* **5**, 69 (2006).
 [13] D. Mayers and A. Yao, [arXiv:quant-ph/9809039](https://arxiv.org/abs/quant-ph/9809039).
 [14] R. Colbeck, Ph.D. thesis, University of Cambridge, 2009.
 [15] S. Bäuml, M. Christandl, K. Horodecki, and A. Winter, *Nat. Commun.* **6**, 6908 (2015).
 [16] M. Horodecki, P. Horodecki, and R. Horodecki, *Phys. Rev. Lett.* **80**, 5239 (1998).
 [17] C. H. Bennett, D. P. DiVincenzo, J. A. Smolin, and W. K. Wootters, *Phys. Rev. A* **54**, 3824 (1996).
 [18] K. Horodecki, M. Horodecki, P. Horodecki, and J. Oppenheim, *Phys. Rev. Lett.* **94**, 160502 (2005).
 [19] M. Christandl and R. Ferrara, *Phys. Rev. Lett.* **119**, 220506 (2017).

- [20] S. Pirandola, R. Laurenza, C. Ottaviani, and L. Banchi, *Nat. Commun.* **8**, 15043 (2017).
- [21] K. Tamaki, H.-K. Lo, W. Wang, and M. Lucamarini, [arXiv:1805.05511](https://arxiv.org/abs/1805.05511).
- [22] K. Horodecki, M. Horodecki, P. Horodecki, and J. Oppenheim, *IEEE Trans. Inf. Theory* **55**, 1898 (2009).
- [23] P. Badziag, K. Horodecki, M. Horodecki, J. Jenkinson, and S. J. Szarek, *Phys. Rev. A* **90**, 012301 (2014).
- [24] R. Horodecki and P. Horodecki, *Phys. Lett. A* **194**, 147 (1994).
- [25] N. J. Cerf and C. Adami, *Phys. Rev. Lett.* **79**, 5194 (1997).
- [26] K. Horodecki, P. Ćwikliński, A. Rutkowski, and M. Studziński, *New J. Phys.* **20**, 083021 (2018).
- [27] I. Devetak and A. Winter, *Proc. R. Soc. London A* **461**, 207 (2005).
- [28] M. Christandl, Ph.D. thesis, University of Cambridge, 2006.
- [29] D. Leung, K. Li, G. Smith, and J. A. Smolin, *Phys. Rev. Lett.* **113**, 030502 (2014).
- [30] G. Vidal and R. F. Werner, *Phys. Rev. A* **65**, 032314 (2002).
- [31] M. B. Plenio, *Phys. Rev. Lett.* **95**, 090503 (2005).
- [32] K. Dobek, M. Karpiński, R. Demkowicz-Dobrzański, K. Banaszek, and P. Horodecki, *Phys. Rev. Lett.* **106**, 030501 (2011).
- [33] J. S. Kim and B. C. Sanders, *Lett. Math. Phys.* **92**, 67 (2010).
- [34] F. G. Brandão and J. Oppenheim, *Phys. Rev. Lett.* **108**, 040504 (2012).
- [35] Y. Dulek, C. Schaffner, and F. Speelman, in *Advances in Cryptology—CRYPTO 2016*, Lecture Notes in Computer Science, Vol. 9816 (Springer, Berlin, Heidelberg, 2016), pp. 3–32.
- [36] G. Alagic, Y. Dulek, C. Schaffner, and F. Speelman, in *Advances in Cryptology—ASIACRYPT 2017*, Lecture Notes in Computer Science, Vol. 10624 (Springer, Cham, 2017), pp. 438–467.
- [37] B. van der Vecht, X. Coiteux-Roy, and B. Skoric, [arXiv:2006.02476](https://arxiv.org/abs/2006.02476).
- [38] K. Horodecki, R. P. Kosteki, R. Salazar, and M. Studziński, *Phys. Rev. A* **102**, 012615 (2020).
- [39] M. Christandl, Diploma thesis, Institute for Theoretical Computer Science, ETH Zurich, 2002.
- [40] A. Winter, *IEEE Trans. Inf. Theory* **45**, 2481 (1999).
- [41] T. Vidick and H. Yuen, [arXiv:1608.04814](https://arxiv.org/abs/1608.04814).
- [42] T. Ogawa and H. Nagaoka, [arXiv:quant-ph/0208139](https://arxiv.org/abs/quant-ph/0208139).
- [43] M. J. Donald and M. Horodecki, *Phys. Lett. A* **264**, 257 (1999).
- [44] A. Winter, *Commun. Math. Phys.* **347**, 291 (2016).
- [45] K. Horodecki, Ł. Pankowski, M. Horodecki, and P. Horodecki, *IEEE Trans. Inf. Theory* **54**, 2621 (2008).
- [46] M. L. Nowakowski, *J. Phys. A: Math. Gen.* **49**, 385301 (2016).

Universal Limitations on Quantum Key Distribution over a NetworkSiddhartha Das^{1,*}, Stefan Bäuml^{2,†}, Marek Winzewski^{3,4} and Karol Horodecki^{4,5}¹*Centre for Quantum Information & Communication (QuIC), École polytechnique de Bruxelles, Université libre de Bruxelles, Brussels, B-1050, Belgium*²*ICFO-Institut de Ciències Fòniques, The Barcelona Institute of Science and Technology, Avinguda Carl Friedrich Gauss 3, 08860 Castelldefels (Barcelona), Spain*³*Institute of Theoretical Physics and Astrophysics, National Quantum Information Centre, Faculty of Mathematics, Physics and Informatics, University of Gdańsk, Wita Stwosza 57, 80-308 Gdańsk, Poland*⁴*International Centre for Theory of Quantum Technologies (ICTQT), University of Gdańsk, 80-308 Gdańsk, Poland*⁵*Institute of Informatics, National Quantum Information Centre, Faculty of Mathematics, Physics and Informatics, University of Gdańsk, Wita Stwosza 57, 80-308 Gdańsk, Poland* (Received 30 September 2020; revised 15 July 2021; accepted 23 July 2021; published 22 October 2021)

We consider the distribution of secret keys, both in a bipartite and a multipartite (conference) setting, via a quantum network and establish a framework to obtain bounds on the achievable rates. We show that any multipartite private state—the output of a protocol distilling secret key among the trusted parties—has to be genuinely multipartite entangled. In order to describe general network settings, we introduce a multiplex quantum channel, which links an arbitrary number of parties where each party can take the role of sender only, receiver only, or both sender and receiver. We define asymptotic and nonasymptotic local quantum operations and classical communication-assisted secret-key-agreement (SKA) capacities for multiplex quantum channels and provide strong and weak converse bounds. The structure of the protocols we consider, manifested by an adaptive strategy of secret-key and entanglement [Greenberger–Horne–Zeilinger (GHZ) state] distillation over an arbitrary multiplex quantum channel, is generic. As a result, our approach also allows us to study the performance of quantum key repeaters and measurement-device-independent quantum key distribution (MDI-QKD) setups. For teleportation-covariant multiplex quantum channels, we get upper bounds on the SKA capacities in terms of the entanglement measures of their Choi states. We also obtain bounds on the rates at which secret key and GHZ states can be distilled from a finite number of copies of an arbitrary multipartite quantum state. We are able to determine the capacities for MDI-QKD setups and rates of GHZ-state distillation for some cases of interest.

DOI: [10.1103/PhysRevX.11.041016](https://doi.org/10.1103/PhysRevX.11.041016)Subject Areas: Atomic and Molecular Physics,
Photonics, Quantum Physics,
Quantum Information**I. INTRODUCTION**

Quantum communication over a network is a pertinent issue from both fundamental and application aspects [1–7]. With technological advancement [8–11], and concerns for privacy [7,12], there is a need for determining protocols and criteria for secret communication among multiple trusted parties in a network. Quantum key distribution (QKD) provides unconditional security for generating secure,

random bits among trusted parties against a quantum eavesdropper, i.e., an eavesdropper that is only limited by the laws of quantum mechanics. Secret key agreement (SKA) among multiple allies is called conference key agreement [13,14]. Conference key agreement can be achieved if all parties involved share a Greenberger–Horne–Zeilinger (GHZ) state [15]. As in the case of bipartite QKD, however, there exists a larger class of states, known as multipartite private states [14], which can provide conference keys by means of local measurements by the parties.

Given the global efforts towards a so-called quantum internet [3,16,17], as well as quantum key distribution over long distances [18,19], it is thus pertinent to establish security criteria and benchmarks on key distribution and entanglement generation capabilities over a quantum network. A quantum network is a complex structure as it

*das.seed@gmail.com

†stefan.baeuml@icfo.eu

Published by the American Physical Society under the terms of the Creative Commons Attribution 4.0 International license. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.

inherits various setups of different quantum channels with particular alignment due to local environmental conditions. One of the biggest obstacles in building this structure is the attenuation of the signal, which cannot be amplified by cloning or broadcasting because of its inherent quantum nature. The signal decays exponentially with distance over an optical fiber [20], and also, the interaction with the environment makes it difficult to preserve entanglement for a long time [10]. Hence, even obtaining a metropolitan-scale quantum network remains a challenge. To overcome these problems, there is a global effort in building technology of quantum repeaters [11,21–23] that could act as relay stations for long-distance quantum communication [7,19].

Some of the first protocols to be performed once a quantum network is available will likely be bipartite as well as multipartite secret key agreement. Securing the network is a necessity for these QKD protocols to be free of loopholes. A number of spectacular attacks on implementations are based on inaccuracy (inefficiency) of detectors of polarized light [24–26]. Based on the idea of entanglement swapping, a novel protocol known as measurement-device-independent QKD (MDI-QKD) [27,28] was introduced, which does not require the honest parties to detect an incoming quantum signal, thus avoiding the problem of detector inefficiencies. This idea has drawn enormous theoretical and experimental attention over the last few years in terms of analyzing achievable key rates for such a scheme with various noise models and performing experiments with current technologies [29–39].

Given the broad interest in implementing such technologies, understanding the fundamental limitations on the key rates achievable in scenarios such as quantum networks and quantum repeaters, as well as setups for MDI-QKD, is an important task. Seminal papers [40,41] on upper bounds on secret key distillation from states, along with results from Refs. [42–46], have led to notable recent progress in the aforementioned direction, for two parties over point-to-point channels assisted by local quantum operations and classical communication (LOCC) [47–50]. Building upon these works, further progress has been made in restricted network settings, e.g., between two parties over bidirectional [51–53], broadcast [54–56], multiple access, and interference quantum channels [54], as well as quantum repeaters [50,57] and networks consisting of point-to-point [58–60] or broadcast channels [61].

In this work, we aim to provide a unifying framework to derive upper bounds on the key rates, both in bipartite and conference settings, achievable in a broad range of different scenarios, including but not limited to broadcast, multiple access, interference channels, repeaters, some MDI-QKD setups, and more general network scenarios. For that purpose, we introduce a multiplex quantum channel, i.e., a multipartite quantum process that connects parties, each playing one of three possible roles—both sender and receiver, only sender, or only receiver. A multiplex

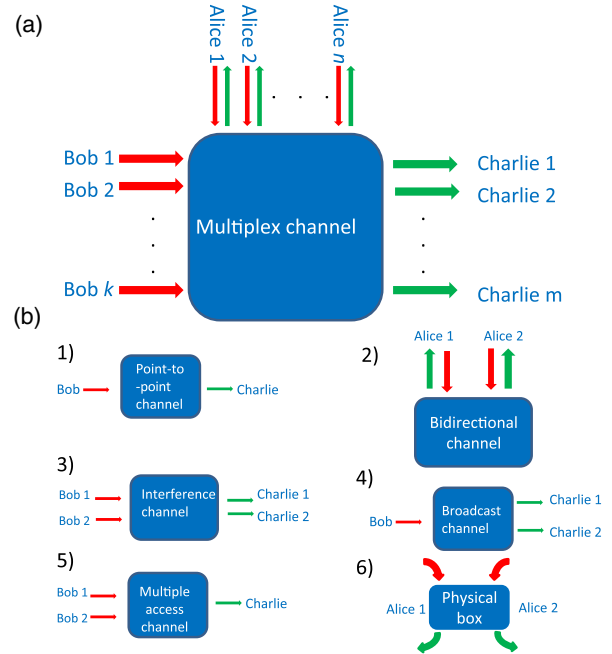


FIG. 1. Pictorial illustration of the universal nature of a multiplex quantum channel from which all other network quantum channels arise, where red and green arrows show inputs and outputs to channels, respectively; see Sec. III B for definitions.

quantum channel is the most general form of a memoryless multipartite quantum channel in a communication network setting. All other network quantum channels can be seen as a special case of this channel (see Fig. 1 for certain common examples). Even the physical setups of MDI-QKD and key repeaters can be described as special cases of multiplex quantum channels (see Fig. 2). In general, the input and output systems on which such a channel acts can be discrete

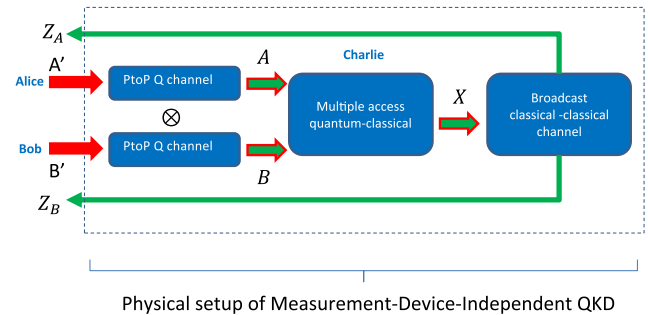


FIG. 2. Graphical depiction of a quantum-to-classical multiplex channel $\mathcal{N}_{A'B' \rightarrow Z_A Z_B}^{\text{MDI}}$ as a bidirectional channel, which is a composition of three elementary multiplex channels. We show a pair of point-to-point channels from Alice to Charlie, and from Bob to Charlie composed of a multiple access quantum-to-classical channel (quantum instrument) performed by Charlie, followed by a broadcast classical channel back to Alice and Bob. The green arrows with red boundaries are the outputs of one multiplex channel, which are, at the same time, inputs to the other channel, hence the coloring.

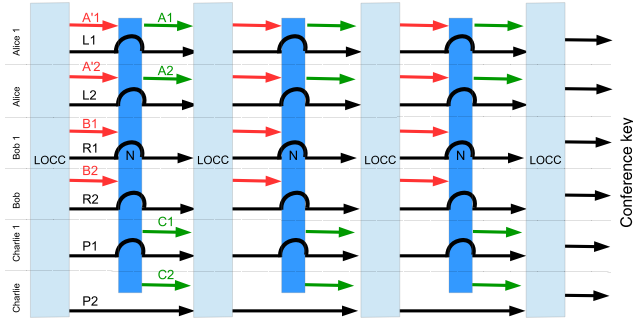


FIG. 3. Example of an LOCC-assisted secret-key-agreement protocol among six parties—Alice 1, Alice 2, Bob 1, Bob 2, Charlie 1, and Charlie 2—using the multiplex channel \mathcal{N} three times. Inputs into \mathcal{N} are depicted in red, outputs in green, and reference systems in black. Alice 1 and 2 enter systems into and receive systems from \mathcal{N} , Bob 1 and 2 only enter systems, and Charlie 1 and 2 only receive systems. In the end, the six parties obtain a six-partite conference key.

(finite-dimensional) or continuous variable (infinite-dimensional) quantum systems.

Next, we introduce secret-key-distribution protocols over multiplex quantum channels with LOCC assistance between users, as shown in Fig. 3, which provides a unifying framework to evaluate performances of various seemingly different QKD protocols. In particular, we describe a general paradigm of QKD protocols where a fixed number of trusted allies are connected over a multiplex quantum channel \mathcal{N} . In these protocols, the allies are allowed to perform LOCC between each use of \mathcal{N} to generate, in the end, a key that is secure against any eavesdropper that satisfies the laws of quantum mechanics. This so-called quantum eavesdropper can have access to all environment parts, including the isometric extension to channel \mathcal{N} .

Our main technical result consists of a metaconverse bound on the one-shot conference key agreement capacity of a multiplex quantum channel, from which we can obtain a number of weak as well as strong converse bounds for the many uses of the multiplex quantum channel, including adaptive and nonadaptive strategies. As our results work in the nonasymptotic setting of a finite number of channel uses, we believe them to be of wide practical interest.

In particular, as an important observation, we show that key repeater protocols, as well as commonly used setups for MDI-QKD, are special cases of LOCC-assisted secret key agreement via a multiplex quantum channel. Whereas bounds on the key rates in such scenarios can also be obtained from a number of earlier results—e.g., from Refs. [50,58,60]—our framework allows for a higher level of specificity in the setups, e.g., by taking into consideration the lack of quantum memory or a particular kind of noisy measurement that is performed in the relay station. Thus, our framework allows us to obtain tighter bounds

than those in Refs. [50,58,60] and even to compute MDI-QKD capacities of certain photon-based practical prototypes that use the so-called dual-rail encoding scheme. This approach provides important tools for benchmarking the performance of such experimentally relevant protocols.

When considering conference key agreement, the pivotal observation we arrive at is that multipartite quantum states with directly accessible secret bits, also called (multipartite) private states [14,62], are genuinely multipartite entangled. This fact also allows us to derive nonasymptotic upper bounds on the secret key distillation from a finite number of copies of a multipartite quantum state.

Our work showcases the topology-dependent and yet universal nature of entanglement measures based on sandwiched Rényi relative entropies [63,64], of which relative entropy is a special case. These entanglement measures provide upper bounds on the secret key rate over an arbitrary multiplex quantum channel, which was first shown for bipartite states in Ref. [40]. The entanglement measures are topology dependent because the upper bound's argument depends (only) on the partition of quantum systems held by trusted allies based on their roles in the network channel. The results are based on the observation that multipartite private states are necessarily genuinely multipartite entangled.

The structure of this paper is as follows. We begin with a brief overview of the main results and briefly mention some important prior results along the direction of our work in Sec. II, respectively. We introduce notations and review basic definitions and relevant prior results in Sec. III. In Sec. IV, we introduce and discuss the properties of entanglement measures for the multiplex quantum channel. We show that genuine multipartite entanglement is a necessary criterion for secrecy. In Sec. V, we introduce LOCC-assisted secret-key-agreement protocols over an arbitrary multiplex quantum channel. We derive upper bounds on the maximum achievable rate for conference key agreement over finite uses of multiplex quantum channels. In Sec. VI, we leverage our bounds to provide nontrivial upper bounds on other quantum key distribution schemes such as measurement-device-independent quantum key distribution and quantum key repeaters. In Sec. VII, we derive lower bounds on the secret-key-agreement capacity over an arbitrary multiplex quantum channel. In Sec. VIII, we derive upper bounds on the number of secret key bits that can be distilled via LOCC among trusted parties sharing a finite number of copies of multipartite quantum states. We provide concluding remarks and open questions in Sec. IX.

II. SUMMARY OF THE MAIN RESULTS

In the following, we provide a brief overview of our main results. Regarding technique, our focus is on multipartite private states, which are the most general class of states that provide the quantum conference key directly (i.e., without

distillation) by local measurements. Such states are of the form [14]

$$\gamma_{SK}^{\overrightarrow{}} := U_{SK}^{\text{tw}} (\Phi_{\vec{K}}^{\text{GHZ}} \otimes \omega_{\vec{S}}) (U_{SK}^{\text{tw}})^{\dagger}, \quad (1)$$

where $\vec{K} = K_1, \dots, K_N$ denotes the so-called key part—i.e., the systems that the N parties involved have to measure in order to obtain conference—and $\vec{S} = S_1, \dots, S_N$ denotes the so-called shield systems, which the parties have to keep secure from the eavesdropper. Also, Φ^{GHZ} is an N -partite GHZ state, ω is some density operator, and U^{tw} is a specifically constructed bipartite unitary operation known as twisting.

We show that states of this form are necessarily genuinely multipartite entangled (GME); i.e., they cannot be expressed as a convex sum of product states no matter with respect to which partition the states are products. To show this, we define a multipartite privacy test, i.e., a dichotomic measurement $\{\Pi', \mathbb{1} - \Pi'\}$ such that any ϵ -approximate multipartite private state ρ with fidelity $F(\rho, \gamma) \geq 1 - \epsilon$ passes the test with success probability $\text{Tr}[\Pi' \rho] \geq 1 - \epsilon$. We then show that any biseparable state σ cannot pass the privacy test with probability larger than $1/K$, where $\log K$ is the number of conference key bits obtainable by measuring (the key part of) γ . Namely, we show that $\text{Tr}[\Pi' \sigma] \leq 1/K$ for all biseparable σ .

As a means of distributing bipartite or multipartite private states among the users, e.g., in a future quantum version of the internet [3,17], we introduce multiplex quantum channels that connect a number of parties that have one of three possible roles—that of only sender, only receiver, or both sender and receiver. We denote senders as Bob 1, ..., Bob k , and their inputs as B_1, \dots, B_k ; receivers as Charlie 1, ..., Charlie m , and their inputs as C_1, \dots, C_m ; and parties that are both senders and receivers as Alice 1, ..., Alice n , with respective inputs A'_1, \dots, A'_n and outputs A_1, \dots, A_n . See also Fig. 1. To describe such channels, we use the notation $\mathcal{N}_{\vec{A}' \vec{B} \rightarrow \vec{A} \vec{C}}$, where, for sake of brevity, we have introduced $\vec{A} := A_1, \dots, A_n$, etc. Furthermore, $:\vec{A}:$ denotes the partition $A_1 : \dots : A_n$ and $:\vec{A} : \vec{B}:$ stands for $A_1 : \dots : A_n : B_1 : \dots : B_k$, etc.

By interleaving the uses of a multiplex quantum channel with LOCC among the parties, we provide a general framework to describe a number of different quantum protocols. The idea is to construct a multiplex quantum channel in such a way that its use, interleaved by LOCC, simulates the protocol. For example, in a MDI-QKD setup, where Alice 1 and Alice 2 send states to the central measurement unit using respective channels $\mathcal{N}^{1,2}$, we can define a (bipartite) multiplex quantum channel of the form

$$\mathcal{N}_{A'_1 A'_2 \rightarrow A_1 A_2}^{\text{MDI}} := \mathcal{B}_{X \rightarrow A_1 A_2} \circ \mathcal{M}_{A''_1 A''_2 \rightarrow X} \circ \mathcal{N}_{A'_1 \rightarrow A''_1}^1 \otimes \mathcal{N}_{A'_2 \rightarrow A''_2}^2. \quad (2)$$

Here, $\mathcal{M}_{A''_1 A''_2 \rightarrow X}$ is the quantum channel performing the central measurement, and $\mathcal{B}_{X \rightarrow A_1 A_2}$ is a classical broadcast channel sending the result back to Alice 1 and Alice 2. Other examples include multipartite MDI-QKD and secret-key-agreement protocols over quantum network laced with key repeaters [50,57].

Generalizing results for point-to-point [48–50] and bidirectional [51–53] channels, we derive divergence-based measures for the entangling abilities of multiplex quantum channels and show that they provide upper bounds on their secret-key-agreement capacities. The measures we introduce are of the following form:

$$\mathbf{E}_r(\mathcal{N}) := \sup_{\tau \in \text{FS}(\vec{L} \vec{A}' : \vec{R} \vec{B} :)} \mathbf{E}_r(\vec{L} \vec{A}' : \vec{R} \vec{C} :)_{\mathcal{N}(\tau)}, \quad (3)$$

where $r = \text{E}$ or $r = \text{GE}$ (E and GE denote entanglement and genuine entanglement, respectively) and FS denotes the set of fully separable states (see Secs. IV A and IV B). Here, \vec{L}, \vec{R} denote ancillary systems that are kept by the respective parties. For any partition $:\vec{X}:$, we have defined \mathbf{E}_r as the divergence from the convex set \mathbf{S}_{E} of fully separable or the convex set \mathbf{S}_{GE} of biseparable states, measured by some divergence \mathbf{D} :

$$\mathbf{E}_r(\vec{X})_{\rho} := \inf_{\sigma \in \mathbf{S}_r(\vec{X})} \mathbf{D}(\rho \| \sigma). \quad (4)$$

Our main results are the following upper bounds on secret-key-agreement capacities of a multiplex quantum channel, i.e., on the maximum rates at which multipartite private states can be obtained by using the channel as well as some free operations. In the one-shot case of a multiplex quantum channel with classical preprocessing and post-processing (cPPP), we have the following weak converse result: For any fixed $\epsilon \in (0, 1)$, the achievable region of cPPP-assisted secret key agreement over a multiplex channel \mathcal{N} satisfies

$$P_{\text{cPPP}}^{(1,\epsilon)}(\mathcal{N}) \leq E_{h,\text{GE}}^{\epsilon}(\mathcal{N}), \quad (5)$$

where $E_{h,\text{GE}}^{\epsilon}(\mathcal{N})$ is the ϵ -hypothesis-testing relative entropy of genuine multipartite entanglement of the multiplex channel \mathcal{N} , which is based on the ϵ -hypothesis-testing divergence [65]. In the case of many channel uses, interleaved by LOCC, we can also show the following strong converse bound:

$$P_{\text{LOCC}}(\mathcal{N}) \leq E_{\text{max},E}(\mathcal{N}), \quad (6)$$

where $E_{\text{max},E}(\mathcal{N})$ is the max-relative entropy of entanglement of the multiplex channel \mathcal{N} , which is based on the max-relative entropy [46]. In the case of finite-dimensional Hilbert spaces, we can also get a strong converse result in terms of the regularized relative entropy,

$$P_{\text{LOCC}}(\mathcal{N}) \leq E_{R,E}^{\infty}(\mathcal{N}). \quad (7)$$

If \mathcal{N} is teleportation-simulable [48,66]—i.e., it can be simulated by a resource state and an LOCC operation—the bounds on $P_{\text{LOCC}}(\mathcal{N})$ reduce to the relative entropy of entanglement of the resource state. Our upper bounds on the secret-key-agreement capacities are also upper bounds on the multipartite quantum capacities, where our goal is to distill GHZ states.

Our technique allows us to compute upper bounds on the rates achievable in MDI-QKD scenarios. For instance, we consider a dual-rail scheme based on single photons [67] to determine bounds on the MDI-QKD rates for two users. In this case, the channels between the users and the relay station are describable by erasure channels \mathcal{E}_i . We obtain the MDI-QKD capacity

$$\tilde{P}_{\text{LOCC}}(\mathcal{N}_{\bar{A} \rightarrow \bar{Z}}^{\text{MDI}, \mathcal{E}}) = q\eta_1\eta_2, \quad (8)$$

where η_i 's are the parameters of the erasure channels connecting users to the relay station and q is the probability of success of the Bell measurement at the relay station (see Sec. VID for a precise model of the MDI-QKD setup). Dependence on η_i allows us to consider the rate-distance trade-off. We also determine upper bounds on the maximum rates for the MDI-QKD setups, where the quantum channels from the users to the relay station are depolarizing and dephasing channels.

We also provide lower bounds on the secret-key-agreement rates of multiplex quantum channels that can be achieved by cppp. Our protocols are based on Devetak-Winter (DW) [68] and generalize the lower bound for multipartite states presented in Ref. [14], as well as the bound for point-to-point quantum channels presented in Ref. [69] to multiplex quantum channels. Our first lower bound is a direct extension of the result for states given in Ref. [14]. The idea is to choose a so-called distributing party that performs the (directed) DW protocol with all remaining parties. The achievable rate is then the worst-case DW rate achievable between the distributing party and any other party. Furthermore, we maximize over all choices for the distributing party. Our second protocol is a variation, where we have a directed chain of parties in which each party performs the DW protocol with the next party in the chain. The obtainable rate is given by the “weakest link,” i.e., the lowest DW rate, in the chain, and we maximize over all possible permutations of the parties in the chain.

In the case of a bidirectional network, i.e., a network in which all nodes are connected with their neighbors by a product of point-to-point channels in opposite directions, we provide a tighter bound based on spanning trees. The idea is to find the lowest DW rate in a spanning tree among any pair of the parties and maximize this quantity among all spanning trees. We provide an example where this protocol achieves a higher rate than the previous ones and show that

the lower bound can be computed with polynomial complexity.

Finally, we show that the techniques developed in previous sections can also be applied to upper bound the rates at which the conference key can be distilled from multipartite quantum states. In particular, we provide an upper bound on the one-shot distillable conference key in terms of the hypothesis-testing relative entropy with respect to biseparable states. Our bound reads

$$K_D^{(1,e)}(\rho) \leq E_{h,\text{GE}}^e(\rho). \quad (9)$$

Using a particular construction of biseparable states, we provide bounds on this quantity for a number of examples, such as (multiple copies of) GHZ and W states, as well as dephased or depolarized GHZ and W states. We also provide an upper bound on the asymptotic distillable conference key, which is given by the regularized relative entropy with respect to biseparable states,

$$K_D(\rho) \leq E_{\text{GE}}^{\infty}(\rho), \quad (10)$$

which is a generalization of the bipartite bound given in Ref. [62].

A. Relation to prior works

We briefly sketch some of the major developments that provide upper bounds on the key distillation capacities from states or via an LOCC-assisted secret-key-agreement protocol over a quantum channel. We then compare our bounds on the SKA capacities with those mentioned in prior works.

Conditions and bounds on the distillable key of bipartite states were provided in Refs. [40,41,62]. The former is in terms of the relative entropy of entanglement [43,44], and the latter is in terms of the squashed entanglement [70] (cf. Refs. [71,72]). These results were generalized to the conference key in Refs. [14,73], respectively.

For an LOCC-assisted secret-key-agreement protocol over a point-to-point channel, Ref. [47] provides a weak converse bound in terms of the squashed entanglement, which is generalized to the distribution of bipartite and multipartite private states via broadcast channels in Ref. [55]. In the case of tele-covariant channels (see Sec. VC), Ref. [48] provides a weak converse bound and Ref. [49] a strong converse bound in terms of the relative entropy of entanglement. This bound has been generalized to the distribution of multiple pairs of bipartite private states via broadcast channels [54,56], as well as multiple-access and interference channels [54].

For arbitrary point-to-point channels, a strong converse bound in terms of the max-relative entropy of entanglement [46] is provided in Ref. [50]. Recently, another strong converse bound in terms of the regularized relative entropy was provided in Ref. [74]. For bidirectional channels,

strong converse bounds in terms of the max-relative entropy of entanglement, which reduce to the relative entropy of entanglement for tele-covariant channels, have been provided in Refs. [51–53].

In the case where the bipartite key is distributed between two parties using a quantum key repeater, bounds have been provided in Ref. [50] when quantum communication takes place over a point-to-point channel. Bounds on rates, at which bipartite and multipartite keys for networks of point-to-point or broadcast channels can be obtained, have been provided in Refs. [58–60,75,76] and [61], respectively. Also, bounds on the rates obtainable in key repeaters that are in terms of entanglement measures of the input states have been obtained in Refs. [57,77].

In an LOCC-assisted conference key agreement protocol, the use of a multiplex quantum channel is interleaved with LOCC among trusted parties. For this scenario, we derive strong converse bounds in terms of the max-relative entropy entanglement for arbitrary multiplex channels. In the case of finite channel dimensions, we also derive bounds in terms of the regularized relative entropy of entanglement. In the case of tele-covariant channels, we obtain bounds in terms of the relative entropy of entanglement. In general, our bounds are not comparable with the squashed entanglement bounds provided in Refs. [47,55]. We are able to retain the results of Refs. [48–50,74] when multiplex channels are assumed to be point-to-point channels. Our bounds in terms of the max-relative entropy are a direct generalization of the bounds on bidirectional channels presented in Refs. [51–53]; thus, we retain those results. By using the recent results in Ref. [74], we further provide bounds in terms of the regularized relative entropy of entanglement, which can provide an improvement.

Concerning quantum key repeaters as well as setups of MDI-QKD, upper bounds on the achievable key rates can be obtained from results bounding key rates achievable in quantum networks, e.g., the one presented in Ref. [60] and subsequently used in Ref. [78] or the ones presented in Refs. [50,58]. However, we note that by designing the right kind of multiplex channel, we can make more specific assumptions on the operations performed at the relay stations and thus obtain tighter bounds. For example, we could design a multiplex channel for a protocol that does not use a quantum memory at the relay station or that performs a particular imperfect measurement at the relay station. The bounds given in Refs. [50,58,60], on the other hand, would bound the key rates of a repeater or MDI-QKD setup by finding the weakest link between the nodes, i.e., only taking into consideration limitations arising from imperfect point-to-point channels linking Alice and Bob with the central relay station, while assuming unlimited quantum memory at the nodes as well as the possibility to perform perfect measurements, resulting in looser bounds. Hence, the bounds given in Refs. [50,58,60] basically reduce to the minimum of the capacities of the

two point-to-point channels, whereas our bounds represent the limitation arising from both imperfect channels and imperfect node operations, which is an important factor when benchmarking experimental implementations.

As for conference key distillation from multipartite states, we provide tighter bounds than those presented in Ref. [14]. As a GHZ state is a special case of a multipartite private state, our bounds can also be applied to the distillation of GHZ states from any pure or mixed multipartite entangled state, both in the asymptotic and finite copies regimes. There are a number of results concerned with computing and bounding rates of multipartite entanglement transformation, including those in Refs. [79–87]. As an example, we consider the nonasymptotic distillation of a tripartite conference key from noisy and noiseless W states and compare our results with Ref. [80].

III. PRELIMINARIES

In this section, we introduce notations and review basic concepts and standard definitions to be used frequently in later sections.

A. Notations and definitions

We consider quantum systems associated with separable Hilbert spaces. We study both discrete and continuous variable quantum systems; therefore, the associated Hilbert spaces can be finite or infinite dimensional. For a composite quantum system AB in a state ρ_{AB} , the reduced state $\text{Tr}_B[\rho_{AB}]$ of system A is denoted as ρ_A . We denote the identity operator as $\mathbb{1}$. Let $\vec{A}' := \{A'_a\}_{a \in \mathcal{A}}$, $\vec{A} := \{A_a\}_{a \in \mathcal{A}}$, $\vec{B} = \{B_b\}_{b \in \mathcal{B}}$, $\vec{C} = \{C_c\}_{c \in \mathcal{C}}$, $\vec{K} = \{K_i\}_{i=1}^M$ denote sets (compositions) of quantum systems, where $\mathcal{A}, \mathcal{B}, \mathcal{C}$ are finite sets of symbols such that $|\mathcal{A}| + |\mathcal{B}| + |\mathcal{C}| = M$ for some natural number $M \geq 2$. We consider M trusted allies $\vec{\mathbf{X}}_i := \{\mathbf{A}_a\}_{a \in \mathcal{A}} \cup \{\mathbf{B}_b\}_{b \in \mathcal{B}} \cup \{\mathbf{C}_c\}_{c \in \mathcal{C}}$. Also, $\vec{L}\vec{A}$ denotes the set $\{L_a A_a\}_{a \in \mathcal{A}}$, where L_a is a reference system of A_a held by \mathbf{A}_a , and the same follows for $\vec{R}\vec{B}$, $\vec{P}\vec{C}$, and $\vec{S}\vec{K}$. A quantum state $\rho_{\vec{A}}$ denotes a joint state of a system formed by composition of all A_a . We use \vec{A} to denote partition with respect to each system in the set \vec{A} as they are held by separate entities, and the same follows for $\vec{L}\vec{A} : \vec{R}\vec{B} : \dots$. Each separate element in a set is held by a separate party, in general. For example, let us consider $\vec{A} = \{A_1, A_2, A_3\}$ for $|\mathcal{A}| = 3$; then, \vec{A} also depicts the composite system $A_1 A_2 A_3$, and \vec{A} denotes the partition $A_1 : A_2 : A_3$ between each subsystem A_a of \vec{A} . In a conference key agreement protocol, each pair K_i, S_i of key and shield systems belongs to the respective trusted party \mathbf{X}_i fully secure from Eve, while all $A'_a, A_a, B_b, C_c, K_i, S_i$ are physically inaccessible to Eve.

Let $\Phi_{\vec{K}}^{\text{GHZ}}$ denote an M -partite GHZ state and $\Phi_{\vec{L}|\vec{A}}^+$ an Einstein–Podolsky–Rosen (EPR) state [88], also called a

maximally entangled state, where maximal entanglement is between \vec{L} and \vec{A} . It should be noted that $\Phi_{L|\vec{A}}^+ = \bigotimes_{a \in \mathcal{A}} \Phi_{L_a|A_a}^+$, where

$$\Phi_{L_a|A_a}^+ = \frac{1}{d} \sum_{i,j=0}^{d-1} |i, i\rangle \langle j, j|_{L_a A_a} \quad (11)$$

for an orthonormal basis $\{|i\rangle\}_i$, where $d = \min\{|L_a|, |A_a|\}$. (Without loss of generality, one may assume an EPR state of an even-dimensional qudit system to be a tensor product of EPR states of qubit systems.)

A quantum channel $\mathcal{M}_{B \rightarrow C}$ is a completely positive, trace-preserving map that acts on trace-class operators defined on the Hilbert space \mathcal{H}_B and uniquely maps them to trace-class operators defined on the Hilbert space \mathcal{H}_C . For a channel $\mathcal{M}_{A \rightarrow B}$ with A and B as input and output systems, its Choi state $J_{LB}^{\mathcal{M}}$ is equal to $\mathcal{M}(\Phi_{LA}^+)$.

A measurement channel $\mathcal{M}_{A' \rightarrow AX}$ is a quantum instrument whose action is expressed as

$$\mathcal{M}_{A' \rightarrow AX}(\cdot) = \sum_x \mathcal{E}_{A' \rightarrow A}^x(\cdot) \otimes |x\rangle \langle x|_X, \quad (12)$$

where each \mathcal{E}^x is a completely positive, trace-nonincreasing map such that \mathcal{M} is a quantum channel and X is a classical register that stores measurement outcomes. A classical register (system) X can be represented with a set of orthogonal quantum states $\{|x\rangle \langle x|_X\}_{x \in \mathcal{X}}$ defined on the Hilbert space \mathcal{H}_X .

An LOCC channel $\mathcal{L}_{\vec{A}' \rightarrow \vec{B}}$ can be written as $\sum_{x \in \mathcal{X}} (\bigotimes_{y \in \mathcal{Y}} \mathcal{E}_{A'_y \rightarrow B_y}^{y,x})$, where $\vec{A}' = \{A'_y\}_y$ and $\vec{B} = \{B_y\}_y$ are sets of inputs and outputs, respectively, and $\{\mathcal{E}^{y,x}\}_x$ is a set of completely positive, trace-nonincreasing maps for each y such that \mathcal{L} is a quantum channel (cf. Ref. [89]). A LOCC channel does not increase the value of entanglement monotonies and is deemed as a free operation in the resource theory of entanglement [14,62,89].

A quantity is called a generalized divergence [90,91] if it satisfies the following monotonicity (data-processing) inequality for all density operators ρ and σ and quantum channels \mathcal{N} :

$$\mathbf{D}(\rho||\sigma) \geq \mathbf{D}(\mathcal{N}(\rho)||\mathcal{N}(\sigma)). \quad (13)$$

Examples include the quantum relative entropy [92]

$$D(\rho||\sigma) := \text{Tr}[\rho \log_2(\rho - \sigma)], \quad (14)$$

for $\text{supp}(\rho) \subseteq \text{supp}(\sigma)$ —otherwise it is ∞ —as well as the sandwiched Rényi relative entropy [63,64], which is denoted as $\tilde{D}_\alpha(\rho||\sigma)$ and defined for states ρ, σ , and $\forall \alpha \in (0, 1) \cup (1, \infty)$ as

$$\tilde{D}_\alpha(\rho||\sigma) := \frac{1}{\alpha - 1} \log_2 \text{Tr}[(\sigma^{(1-\alpha)/2\alpha} \rho \sigma^{(1-\alpha)/2\alpha})^\alpha], \quad (15)$$

but it is set to $+\infty$ for $\alpha \in (1, \infty)$ if $\text{supp}(\rho) \not\subseteq \text{supp}(\sigma)$. In the limit $\alpha \rightarrow 1$, the sandwiched Rényi relative entropy converges to the quantum relative entropy; in the limit $\alpha \rightarrow \infty$, it converges to the max-relative entropy [64], which is defined as [46,93]

$$D_{\max}(\rho||\sigma) := \inf\{\lambda \in \mathbb{R} : \rho \leq 2^\lambda \sigma\}, \quad (16)$$

and if $\text{supp}(\rho) \not\subseteq \text{supp}(\sigma)$, then $D_{\max}(\rho||\sigma) = \infty$. Another generalized divergence is the ε -hypothesis-testing divergence [65,94], defined as

$$D_h^\varepsilon(\rho||\sigma) := -\log_2 \inf_{\Lambda: 0 \leq \Lambda \leq 1} \{\text{Tr}[\Lambda \sigma] : \text{Tr}[\Lambda \rho] \geq 1 - \varepsilon\}, \quad (17)$$

for $\varepsilon \in [0, 1]$ and density operators ρ, σ . For a more detailed description and other examples of the generalized divergences like the trace distance $\|\rho - \sigma\|_1$ and negative of fidelity $-F(\rho, \sigma)$ and their properties, see the Appendix A.

B. Multiplex quantum channels

We now formally define a general form of network channel that encompasses all other known multiplex quantum channels possible in communication or information processing settings [see Fig. 1(a) and Appendix B]. To the best of our knowledge, there is not such a general form of network channel in the literature of quantum communication and computation.

Definition 1: Consider the multipartite quantum channel $\mathcal{N}_{\vec{A}' \vec{B} \rightarrow \vec{A} \vec{C}}$, where each pair A'_a, A_a is held by a respective party \mathbf{A}_a and each B_b, C_c are held by parties $\mathbf{B}_b, \mathbf{C}_c$, respectively. While A_a is both the sender and receiver to the channel, \mathbf{B}_b is only a sender, and \mathbf{C}_c is only a receiver to the channel. Such a quantum channel is referred to as the multiplex quantum channel. Any two different systems need not be of the same size, in general. The sets \mathcal{A}, \mathcal{B} , or \mathcal{C} can be empty in such a way that there is at least one input to the channel and one output from the channel.

Definition 1 includes all scenarios depicted in Fig. 1 (see Appendix B). For example, for a point-to-point channel from Bob to Charlie the set $\mathcal{A} = \emptyset$ and the sets \mathcal{B} and \mathcal{C} are singleton sets.

Also, any physical box with quantum or classical inputs and quantum or classical outputs is a type of multiplex quantum channel. We may not have an exact description of what is going on inside the box except that the undergoing process is physical, i.e., described by quantum mechanics. Physical computational devices like a physical black box (oracle) and quantum circuit [95] are also examples of multiplex quantum channels.

C. Conference key and private states

There are two usual approaches to studying secret key distillation. A direct approach is to consider purifications of states where the purifying system is accessible to Eve and all allied parties are allowed to perform local operations and public communication (LOPC). In this approach, we have Eve and M allied parties. Another approach is to consider the private states defined below, where all allied parties perform LOCC. We need not consider Eve explicitly in the paradigm of private states, and it is assumed that purifications of states are accessible to Eve. Both approaches are known to be equivalent [40]. We discuss the equivalence of these two approaches in more detail in Sec. V.

We now review the properties of conference key states discussed in Ref. [14]. Conference key states are a multipartite generalization of the secret key shared between two parties.

Definition 2: A conference key state $\gamma_{\bar{K}E}^c$, with $|K_i| = K$ for all $i \in [M] := 1, \dots, M$, is defined as

$$\begin{aligned} & \mathcal{D}_{K_1} \otimes \mathcal{D}_{K_2} \otimes \dots \otimes \mathcal{D}_{K_M}(\gamma_{\bar{K}E}^c) \\ & := \frac{1}{K} \sum_{k \in \mathcal{K}} |k\rangle\langle k|_{K_1} \otimes |k\rangle\langle k|_{K_2} \otimes \dots \otimes |k\rangle\langle k|_{K_M} \otimes \sigma_E, \end{aligned} \quad (18)$$

where σ_E is a state of the system E , which is accessible to an eavesdropper Eve, $\mathcal{D}(\cdot) = \sum_{k \in \mathcal{K}} |k\rangle\langle k|(\cdot)|k\rangle\langle k|$ is a projective measurement channel, and $\{|k\rangle_{K_i}\}_{k \in \mathcal{K}}$ forms an orthonormal basis for each $i \in [M]$.

A conference key state $\gamma_{\bar{K}E}^c$ has $\log_2 K$ secret bits (key) that are readily accessible.

A state $\rho_{\bar{K}E}$ is called an ε -approximate conference key state, for $\varepsilon \in [0, 1]$, if there exists a conference key state $\gamma_{\bar{K}E}^c$ such that [14]

$$F(\gamma_{\bar{K}E}^c, \rho_{\bar{K}E}) \geq 1 - \varepsilon. \quad (19)$$

Definition 3: A state $\gamma_{\bar{S}M}$, with $|K_i| = K$ for all $i \in [M]$, is called a (M -partite) private state if and only if

$$\gamma_{\bar{S}M} := U_{\bar{S}K}^{\text{tw}}(\Phi_{\bar{K}}^{\text{GHZ}} \otimes \omega_{\bar{S}})(U_{\bar{S}K}^{\text{tw}})^\dagger, \quad (20)$$

where $U_{\bar{S}K}^{\text{tw}} := \sum_{\vec{k} \in \mathcal{K}^{\times M}} |\vec{k}\rangle\langle \vec{k}|_{\bar{K}} \otimes U_{\bar{S}}^{\vec{k}}$ is called a twisting unitary operator for some unitary operator $U_{\bar{S}}^{\vec{k}}$ and ω is some density operator [14].

It should be noted that $\gamma_{\bar{S}M}$ has at least $\log_2 K$ secret (key) bits (see Ref. [62] for a discussion of when the private state has exactly $\log_2 K$ bits). Similar to a conference key state, a state $\rho_{\bar{S}M}$ is called an ε -approximate private state for $\varepsilon \in [0, 1]$ if there exists a private state $\gamma_{\bar{S}M}$ such that [14]

$$F(\gamma_{\bar{S}M}, \rho_{\bar{S}M}) \geq 1 - \varepsilon. \quad (21)$$

Any state extension (including purification) $\gamma_{\bar{S}KE}$ of such a private state (20) necessarily has the following form [14]:

$$\gamma_{\bar{S}KE} := U_{\bar{S}KE}^{\text{tw}}(\Phi_{\bar{K}} \otimes \omega_{\bar{S}E})(U_{\bar{S}KE}^{\text{tw}})^\dagger, \quad (22)$$

where $\omega_{\bar{S}E}$ is a state extension of the density operator $\omega_{\bar{S}}$.

It follows from Theorem IV.1 of Ref. [14] that $F(\gamma_{\bar{K}E}^c, \rho_{\bar{K}E}) \geq 1 - \varepsilon$ implies $F(\gamma_{\bar{S}K}, \rho_{\bar{S}K}) \geq 1 - \varepsilon$, and the converse is also true; i.e., $F(\gamma_{\bar{S}K}, \rho_{\bar{S}K}) \geq 1 - \varepsilon$ implies $F(\gamma_{\bar{K}E}^c, \rho_{\bar{K}E}) \geq 1 - \varepsilon$.

It is known that all perfect private states have nonlocal correlations [96].

IV. ENTANGLEMENT AND PRIVACY TEST

This section introduces frameworks for the resource theories of multipartite entanglement for the multipartite quantum channels (see Refs. [51,53,97,98] for the discussion on bipartite channels).

A. Multipartite entanglement

Here, we provide a short overview of the relevant definitions. For a detailed review of the topic, see Ref. [99]. A pure n -partite state that can be written as a tensor product $|\psi_1\rangle \otimes |\psi_2\rangle \otimes \dots \otimes |\psi_m\rangle$ is called m -separable. If $m < n$, there are partitions of the set of all the parties into two, with respect to which the state is entangled. If $n = m$, the pure state is said to be fully separable. If there is no bipartition with respect to which the pure state is a product state, it is called genuinely n -partite entangled.

An arbitrary n -partite state is m -separable if it can be written as the following convex composition:

$$\rho_{m\text{-sep}} = \sum_{x \in \mathcal{X}} p_X(x) |\psi_1^x\rangle\langle \psi_1^x| \otimes |\psi_2^x\rangle\langle \psi_2^x| \otimes \dots \otimes |\psi_m^x\rangle\langle \psi_m^x|, \quad (23)$$

where $p_X(x)$ is a probability distribution. The m -separable states form a convex set. Note, however, that the subsystems with respect to which the elements of the decomposition have to be products can differ.

A mixed n -partite state is considered GME if any decomposition into pure states contains at least one genuinely n -partite entangled pure state; i.e., the state is not biseparable. Let a free set $\mathbf{F}(\vec{A})$ denote the set of all fully separable and biseparable states of system \vec{A} for $\mathbf{F} = \text{FS}$ and $\mathbf{F} = \text{BS}$, respectively. Both the sets FS and BS are convex. We note that while FS is preserved under an LOCC operation and tensor product, BS is preserved under LOCC but not under the tensor product, i.e., $\rho_{A^{(x)}}^{(x)} \in \text{BS}(\vec{A}^{(x)})$

for $x \in [2]$ but $\rho^{(1)} \otimes \rho^{(2)}$ need not belong to $\text{BS}(\overrightarrow{A^{(1)}A^{(2)}})$. We refer to biseparable quantum states whose biseparability is preserved under tensor products, i.e., $\rho \xrightarrow{A^{(x)}} \in \text{BS}(\overrightarrow{A^{(x)}})$ and $\rho^{(1)} \otimes \dots \otimes \rho^{(n)} \in \text{BS}(\overrightarrow{A^{(1)} \dots A^{(n)}})$ for all $n \in \mathbb{N}$, as tensor-stable biseparable states.

B. Entanglement measures

It is pertinent to quantify the resourcefulness of states and channels. The bounds on the capacities that we obtain are in terms of these quantifiers. It is desirable for entanglement quantifiers to be non-negative, to attain their minimum for the free states (and separable channels, respectively), and to be monotone under the action of LOCC.

Definition 4: The generalized divergence of entanglement \mathbf{E}_E or GME \mathbf{E}_{GE} of an arbitrary state $\rho_{\vec{A}}$ is defined as [100]

$$\mathbf{E}_r(\overrightarrow{A})_\rho := \inf_{\sigma \in \mathbf{F}(\overrightarrow{A})} \mathbf{D}(\rho_{\vec{A}} \| \sigma_{\vec{A}}), \quad (24)$$

when $\mathbf{F} = \text{FS}$ or $\mathbf{F} = \text{BS}$ for $r = E$ or $r = GE$, respectively, where $\mathbf{D}(\rho \| \sigma)$ denotes the generalized divergence.

The following definition of the entanglement measure of a multiplex channel generalizes the notion of entangling power of bipartite quantum channels [101] (see also Refs. [51,53,102]).

Definition 5: The entangling power of a multiplex channel $\mathcal{N}_{\overrightarrow{A'} \overrightarrow{B'} \overrightarrow{A} \overrightarrow{C}}$ with respect to entanglement measure \mathbf{E}_r [Eq. (24)] is defined as the maximum possible gain in the entanglement \mathbf{E}_r when a quantum state is acted upon by the given channel \mathcal{N} ,

$$\mathbf{E}_r^p(\mathcal{N}) := \sup_{\rho} [\mathbf{E}_r(\overrightarrow{LA} : \overrightarrow{R} : \overrightarrow{PC})_{\mathcal{N}(\rho)} - \mathbf{E}_r(\overrightarrow{LA'} : \overrightarrow{RB} : \overrightarrow{P})_\rho], \quad (25)$$

where optimization is over all possible input states $\rho_{\overrightarrow{LA'} \overrightarrow{RB} \overrightarrow{P}}$.

Another way to quantify the entanglement measure of a multiplex channel is the following (see Ref. [53] for the bidirectional channel).

Definition 6: The generalized divergence of entanglement $\mathbf{E}_E(\mathcal{N})$ or GME $\mathbf{E}_{GE}(\mathcal{N})$ of a multiplex channel $\mathcal{N}_{\overrightarrow{A'} \overrightarrow{B'} \overrightarrow{A} \overrightarrow{C}}$ is

$$\mathbf{E}_r(\mathcal{N}) := \sup_{\rho \in \mathbf{FS}(\overrightarrow{LA'} : \overrightarrow{RB} : \overrightarrow{P})} \mathbf{E}_r(\overrightarrow{LA} : \overrightarrow{R} : \overrightarrow{C})_{\mathcal{N}(\rho)}, \quad (26)$$

for $r = E$ or $r = GE$, respectively, where $\mathbf{E}_r(\overrightarrow{A})_\rho$ is defined in Eq. (24) and GME stands for genuinely multipartite entanglement.

For $r = E$, the entanglement measure in Eq. (24) is called ε -hypothesis-testing relative entropy of entanglement $E_{h,E}^\varepsilon$, max-relative entropy of entanglement $E_{\max,E}$, sandwiched Rényi relative entropy of entanglement $\tilde{E}_{\alpha,E}$, or relative entropy of entanglement E_E when the generalized divergence is the ε -hypothesis-testing relative entropy, max-relative entropy, sandwiched Rényi relative entropy, or relative entropy, respectively. For $r = GE$, the entanglement measure in Eq. (24) is called ε -hypothesis-testing relative entropy of GME $E_{h,GE}^\varepsilon$, max-relative entropy of GME $E_{\max,GE}$, sandwiched Rényi relative entropy of GME $\tilde{E}_{\alpha,GE}$, or relative entropy of GME when the generalized divergence E_{GE} is the ε -hypothesis-testing relative entropy, max-relative entropy, sandwiched Rényi relative entropy, or relative entropy, respectively. We follow the same procedure for the nomenclature of entanglement measures of channels.

We note that the sets FS, BS are convex. Using the data-processed triangle inequality [50] and the argument from the proof of Proposition 2 in Ref. [51], we arrive at the following lemma.

Lemma 1: The entangling power of a multiplex channel $\mathcal{N}_{\overrightarrow{A'} \overrightarrow{B'} \overrightarrow{A} \overrightarrow{C}}$ with respect to the max-relative entropy of entanglement $E_{\max,E}$ is equal to the max-relative entropy of entanglement of the channel \mathcal{N} ,

$$E_{\max,E}^p(\mathcal{N}) = E_{\max,E}(\mathcal{N}). \quad (27)$$

Using a recent result on relative entropies [103], we can also obtain a result for the relative entropy of entanglement. Let us first define the regularized relative entropy of entanglement of a multiplex channel $\mathcal{N}_{\overrightarrow{A'} \overrightarrow{B'} \overrightarrow{A} \overrightarrow{C}}$ as

$$E_R^\infty(\mathcal{N}) := \inf_{\Lambda \in \text{LOCC}} D^\infty(\mathcal{N}_{\overrightarrow{A'} \overrightarrow{B'} \overrightarrow{A} \overrightarrow{C}} \| \Lambda_{\overrightarrow{A'} \overrightarrow{B'} \overrightarrow{A} \overrightarrow{C}}), \quad (28)$$

where $D^\infty(\mathcal{N} \| \mathcal{M}) := \lim_{n \rightarrow \infty} (1/n) D(\mathcal{N}^{\otimes n} \| \mathcal{M}^{\otimes n})$ and

$$D(\mathcal{N} \| \mathcal{M}) := \max_{\phi_{\overrightarrow{LA'} \overrightarrow{RB} \overrightarrow{P}}} D(\mathcal{N}_{\overrightarrow{A'} \overrightarrow{B'} \overrightarrow{A} \overrightarrow{C}}(\phi) \| \mathcal{M}_{\overrightarrow{A'} \overrightarrow{B'} \overrightarrow{A} \overrightarrow{C}}(\phi)), \quad (29)$$

where $L \simeq A'$, $R \simeq B$ and $P \simeq C$. We now show the following relation between the regularized relative entropy of entanglement and the relative entropy of entanglement.

Lemma 2: For finite-dimensional Hilbert spaces, the entangling power of a multiplex channel $\mathcal{N}_{\overrightarrow{A'} \overrightarrow{B'} \overrightarrow{A} \overrightarrow{C}}$ with respect to the relative entropy of entanglement E_E is less than or equal to the regularized relative entropy of entanglement of the channel \mathcal{N} ,

$$E_E^p(\mathcal{N}) \leq E_E^\infty(\mathcal{N}). \quad (30)$$

Proof.—Let $\rho_{\overrightarrow{LA'} \overrightarrow{RB} \overrightarrow{P}}$ be a state and let $\sigma' \in \text{FS}(\overrightarrow{LA'} : \overrightarrow{RB} : \overrightarrow{P})$. Let $\Lambda_{\overrightarrow{A'} \overrightarrow{B'} \overrightarrow{A} \overrightarrow{C}}$ be an LOCC channel. Then, the following inequality holds:

$$E_E(\vec{LA}:\vec{R}:\vec{PC})_{\mathcal{N}(\rho)} \leq D\left(\mathcal{N}_{\vec{A}'\vec{B}\rightarrow\vec{A}\vec{C}}(\rho_{\vec{L}\vec{A}'\vec{R}\vec{B}\vec{P}})\|\Lambda_{\vec{A}'\vec{B}\rightarrow\vec{A}\vec{C}}(\sigma'_{\vec{L}\vec{A}'\vec{R}\vec{B}\vec{P}})\right). \quad (31)$$

Applying the chain rule from Ref. [103], we find that

$$\begin{aligned} & D\left(\mathcal{N}_{\vec{A}'\vec{B}\rightarrow\vec{A}\vec{C}}(\rho_{\vec{L}\vec{A}'\vec{R}\vec{B}\vec{P}})\|\Lambda_{\vec{A}'\vec{B}\rightarrow\vec{A}\vec{C}}(\sigma'_{\vec{L}\vec{A}'\vec{R}\vec{B}\vec{P}})\right) \\ & \leq D\left(\rho_{\vec{L}\vec{A}'\vec{R}\vec{B}\vec{P}}\|\sigma'_{\vec{L}\vec{A}'\vec{R}\vec{B}\vec{P}}\right) + D^\infty(\mathcal{N}_{\vec{A}'\vec{B}\rightarrow\vec{A}\vec{C}}\|\Lambda_{\vec{A}'\vec{B}\rightarrow\vec{A}\vec{C}}). \end{aligned}$$

Since the above holds for arbitrary, fully separable states $\sigma'_{\vec{L}\vec{A}'\vec{R}\vec{B}\vec{P}}$ and arbitrary LOCC channels $\Lambda_{\vec{A}'\vec{B}\rightarrow\vec{A}\vec{C}}$, we arrive at

$$E_E(\vec{LA}:\vec{R}:\vec{PC})_{\mathcal{N}(\rho)} \leq E_E(\vec{LA}:\vec{R}\vec{B}:\vec{P})_\rho + E_E^\infty(\mathcal{N}), \quad (32)$$

finishing the proof. \blacksquare

Remark 1: It suffices to optimize $E_{h,E}^e(\mathcal{N})$, $E_{h,GE}^e(\mathcal{N})$, $E_{\max,E}(\mathcal{N})$, $E_{E,E}(\mathcal{N})$, and $E_{\max,GE}(\mathcal{N})$ of a multiplex channel \mathcal{N} over all pure input states; i.e., $\rho \in \text{FS}(\vec{LA}:\vec{R}\vec{B})$ is a pure state in Eq. (26) for $E_{h,E}^e(\mathcal{N})$, $E_{h,GE}^e(\mathcal{N})$, $E_{\max,E}(\mathcal{N})$, $E_{E,E}(\mathcal{N})$, $E_{\max,GE}(\mathcal{N})$. This reduction follows from the quasiconvexity of the max-relative entropy [93] and ε -hypothesis-testing relative entropy [104], as well as the convexity of the relative entropy of entanglement [44]. Namely, the maximum of a (quasi)convex function over a convex set will be attained on a boundary point. The boundary points of the set of fully separable density matrices are given by the fully separable pure states.

C. Multipartite privacy test

A γ -privacy test corresponding to γ_{SK}^- is defined as the dichotomic measurement [49] $\{\Pi_{SK}^\gamma, \mathbb{1} - \Pi_{SK}^\gamma\}$, where $\Pi_{SK}^\gamma := U_{SK}^{\text{tw}}(\Phi_{\vec{K}} \otimes \mathbb{1}_{\vec{S}})(U_{SK}^{\text{tw}})^\dagger$.

Using the properties of fidelity and form of the test measurement, we arrive at the following proposition.

Proposition 1: If a state ρ_{SK}^- is ε approximate to γ_{SK}^- , i.e., $F(\rho_{SK}^-, \gamma_{SK}^-) \geq 1 - \varepsilon$, then ρ_{KS}^- passes the γ -privacy test with success probability $1 - \varepsilon$, i.e.,

$$\text{Tr}[\Pi_{SK}^\gamma \rho_{SK}^-] \geq 1 - \varepsilon. \quad (33)$$

Proof.—

$$\begin{aligned} & \text{Tr}[\Pi_{SK}^\gamma \rho_{SK}^-] \\ & = \langle \Phi_{\vec{K}}^{\text{GHZ}} |_{\vec{K}} \text{Tr}_{\vec{S}}[(U_{SK}^{\text{tw}})^\dagger \rho_{SK}^- U_{SK}^{\text{tw}}] | \Phi_{\vec{K}}^{\text{GHZ}} \rangle_{\vec{K}} \end{aligned} \quad (34)$$

$$= F(\Phi_{\vec{K}}^{\text{GHZ}}, \text{Tr}_{\vec{S}}[(U_{SK}^{\text{tw}})^\dagger \rho_{SK}^- U_{SK}^{\text{tw}}]) \quad (35)$$

$$\geq F(\Phi_{\vec{K}}^{\text{GHZ}} \otimes \omega_{\vec{S}}, (U_{SK}^{\text{tw}})^\dagger \rho_{SK}^- U_{SK}^{\text{tw}}) \quad (36)$$

$$= F(U_{SK}^{\text{tw}} \Phi_{\vec{K}}^{\text{GHZ}} \otimes \omega_{\vec{S}} (U_{SK}^{\text{tw}})^\dagger, \rho_{SK}^-) \quad (37)$$

$$= F(\gamma_{SK}^-, \rho_{SK}^-) \geq 1 - \varepsilon. \quad (38)$$

We employ proof arguments similar to the bipartite case of Eq. (281) in Ref. [62] to arrive at the following theorem, which implies that all private states are necessarily GME states. This is a strict generalization of Eq. (281) in Ref. [62], as a direct generalization would be the same statement for fully separable states instead of biseparable states (cf. Ref. [14]). See Appendix C for the proof. \blacksquare

Theorem 1: A biseparable state $\sigma_{SK}^- \in \text{BS}(\vec{SK})$ can never pass any γ -privacy test with probability greater than $1/K$, i.e.,

$$\text{Tr}[\Pi_{SK}^\gamma \sigma_{SK}^-] \leq \frac{1}{K}. \quad (39)$$

V. CONFERENCE KEY AGREEMENT PROTOCOL

In this section, we give a formal description of a secret-key-agreement protocol for multiple trusted parties, i.e., a conference key agreement protocol.

We consider an LOCC-assisted secret-key-agreement protocol among M trusted allies $\{\mathbf{X}_i\}_{i=1}^M$ over a multiplex quantum channel $\mathcal{N}_{\vec{A}'\vec{B}\rightarrow\vec{A}\vec{C}}$, where each pair A'_a, A_a is held by trusted party \mathbf{A}_a and each B_b, C_c is held by trusted parties $\mathbf{B}_b, \mathbf{C}_c$, respectively. The environment part E of an isometric extension $U_{\vec{A}'\vec{B}\rightarrow\vec{A}\vec{C}E}^{\mathcal{N}}$ of the channel \mathcal{N} is accessible to Eve, along with all classical information communicated among \mathbf{X}_i while performing LOCC. All other quantum systems that are locally available to \mathbf{X}_i are said to be secure from Eve; i.e., even if local operations during LOCC are noisy, purifying quantum systems are still within labs of trusted allies, which are off limits for Eve. This assumption is justifiable because \mathbf{X}_i 's can always abandon performing local operations that would leak information to Eve. In an LOCC-assisted protocol, the uses of the multiplex channel \mathcal{N} are interleaved with LOCC channels.

In the first round, all \mathbf{X}_i perform LOCC \mathcal{L}^1 to generate a state $\rho_1 \in \text{FS}(\overrightarrow{L^{(1)}A^{(1)}}; \overrightarrow{R^{(1)}B^{(1)}}; \overrightarrow{P^{(1)}})$. All \mathbf{A}_a and \mathbf{B}_b input respective systems to multiplex channel $\mathcal{N}_{A^{(1)}B^{(1)} \rightarrow C^{(1)}}^1$ and let $\tau_1 := \mathcal{N}^1(\rho)$ be the output state after the first use \mathcal{N}^1 of the multiplex channel. In the second round, an LOCC \mathcal{L}^2 is performed on τ_1 , and then, the second use \mathcal{N}^2 of the multiplex channel is employed on $\rho_2 := \mathcal{L}^2(\tau_1)$. In the third round, an LOCC \mathcal{L}^3 is performed on $\tau_2 := \mathcal{N}^2(\rho_2)$, and then, the third use \mathcal{N}^3 of the multiplex channel is employed on $\rho_3 := \mathcal{L}^3(\tau_2)$. Successively, we continue this procedure for n rounds, where an \mathcal{L} acts on the output state of the previous round, after which the multiplex channel is performed on the resultant state. Finally, after the n th round, an LOCC \mathcal{L}^{n+1} is performed as a decoding channel, which generates the final state ω_{SK}^- .

It can be concluded from the equivalence between private states and CK states that any protocol of the above form can be purified, i.e., by considering isometric extensions of all channels (LOCC and \mathcal{N}) (the proof arguments are the same as for the purified protocol for LOCC-assisted secret key agreement [51]). At the end of the purified protocol, Eve possesses all the environment systems E^n from isometric extension $U^{\mathcal{N}}$ of each use of the multiplex channel \mathcal{N} along with coherent copies Y^{n+1} of the classical data exchanged among trusted parties \mathbf{X}_i during performances of $n+1$ LOCC channels, whereas each trusted party \mathbf{X}_i possesses the key system K_i and the shield system S_i , which consist of all local reference systems, after the action of the decoder. The state at the end of the protocol is a pure state $\omega_{SK}^-_{Y^{n+1}E^n}$ with $F(\gamma_{SK}^-, \omega_{SK}^-) \geq 1 - \epsilon$. Such a protocol is called an (n, K, ϵ) LOCC-assisted secret-key-agreement protocol. The rate P of a given (n, K, ϵ) protocol is equal to the number of conference (secret) bits generated per channel use:

$$P := \frac{1}{n} \log_2 K. \quad (40)$$

A rate P is achievable if for $\epsilon \in (0, 1)$, $\delta > 0$, and sufficiently large n , there exists an $(n, 2^{n(P-\delta)}, \epsilon)$ LOCC-assisted secret-key-agreement protocol. The LOCC-assisted secret-key-agreement capacity $\hat{P}_{\text{LOCC}}(\mathcal{N})$ of a multiplex quantum channel \mathcal{N} is defined as the supremum of all achievable rates.

A rate P is called a strong converse rate for LOCC-assisted secret key agreement if for all $\epsilon \in [0, 1)$, $\delta > 0$, and sufficiently large n , there does not exist an $(n, 2^{n(P+\delta)}, \epsilon)$ LOCC-assisted secret-key-agreement protocol. The strong converse LOCC-assisted secret-key-agreement capacity

$\tilde{P}_{\text{LOCC}}(\mathcal{N})$ is defined as the infimum of all strong converse rates.

The following inequality is a direct consequence of the definitions:

$$\hat{P}_{\text{LOCC}}(\mathcal{N}) \leq \tilde{P}_{\text{LOCC}}(\mathcal{N}). \quad (41)$$

We can also consider the whole development discussed above for conference key agreement assisted only by cppp communication; i.e., all parties are allowed only two LOCC channels, one for encoding and the other for decoding. A (n, K, ϵ) cppp-assisted secret-key-agreement protocol over \mathcal{N} is the same as a $(1, K, \epsilon)$ LOCC-assisted secret-key-agreement protocol over channel $\mathcal{N}^{\otimes n}$, and for $n = 1$, both protocols are the same. The cppp-assisted secret-key-agreement capacity \hat{P}_{cppp} of the channel \mathcal{N} is always less than or equal to \hat{P}_{LOCC} ,

$$\hat{P}_{\text{cppp}}(\mathcal{N}) \leq \hat{P}_{\text{LOCC}}(\mathcal{N}). \quad (42)$$

Let $\hat{P}_{\text{cppp}}^{\mathcal{N}}(n, \epsilon)$ be the maximum rate such that $(n, 2^{nP}, \epsilon)$ cppp-assisted secret key agreement is achievable for any given \mathcal{N} .

Remark 2: It should be noted that the maximum rate at which secret keys can be distilled using the LOCC- or cppp-assisted protocol over a multiplex channel \mathcal{N} is never less than the maximum rate at which the GHZ state can be distilled using the LOCC- or cppp-assisted protocol over a given channel \mathcal{N} , respectively. This statement holds because the GHZ state is a special private state from which secret bits are readily accessible to trusted allies.

Remark 3: Different physical constraints can be invoked in communication protocols to define constrained protocols and associated capacities. For instance, we can invoke energy constraints on input states and detectors to get energy-constrained protocols and respective capacities (cf. Refs. [105,106]).

A. Privacy from a single use of a multiplex channel

Let $\hat{P}_{\text{cppp}}^{\mathcal{N}}(n, \epsilon)$ denote the maximum rate P such that the (n, K, ϵ) conference key agreement protocol is achievable for any \mathcal{N} using cppp. The following bound holds for the one-shot secret-key-agreement rate of a multiplex quantum channel \mathcal{N} (see Appendix D 1 for the proof).

Theorem 2: For any fixed $\epsilon \in (0, 1)$, the achievable region of cppp-assisted secret key agreement over a single use of the multiplex channel $\mathcal{N}_{\vec{A} \vec{B} \rightarrow \vec{A} \vec{C}}$ satisfies

$$\hat{P}_{\text{cppp}}^{\mathcal{N}}(1, \epsilon) \leq E_{h, \text{GE}}^{\epsilon}(\mathcal{N}), \quad (43)$$

where

$$E_{h,\text{GE}}^\varepsilon(\mathcal{N}) := \sup_{\psi \in \text{FS}(\overrightarrow{LA'}:\overrightarrow{RB}:)} \inf_{\sigma \in \text{BS}(\overrightarrow{LA}:\overrightarrow{R}:\overrightarrow{C}:)} D_h^\varepsilon(\mathcal{N}(\psi) \parallel \sigma) \quad (44)$$

is the ε -hypothesis-testing relative entropy of genuine entanglement of the multiplex channel \mathcal{N} . It suffices to optimize over pure input states $\psi \in \text{FS}(\overrightarrow{LA'}:\overrightarrow{RB}:)$.

We can conclude from the above theorem that

$$\hat{P}_{\text{cppp}}^\mathcal{N}(n, \varepsilon) \leq \frac{1}{n} E_{h,\text{GE}}^\varepsilon(\mathcal{N}^{\otimes n}), \quad (45)$$

which leads to the following corollaries.

Corollary 1: A weak converse bound on the cppp-assisted secret-key-agreement capacity of a multiplex channel \mathcal{N} is given by

$$\hat{P}_{\text{cppp}}(\mathcal{N}) = \inf_{\varepsilon \in (0,1)} \liminf_{n \rightarrow \infty} \hat{P}_{\text{cppp}}^\mathcal{N}(n, \varepsilon) \quad (46)$$

$$\leq E_{\text{GE}}^\infty(\mathcal{N}). \quad (47)$$

Corollary 2: Consider a class of multiplex channels $\mathcal{N}_{\overrightarrow{A'}\overrightarrow{B} \rightarrow \overrightarrow{A}\overrightarrow{C}}$ such that for all pure input states $\psi \in \text{FS}(\overrightarrow{LA'}:\overrightarrow{RB}:\overrightarrow{P})$, the output states $\mathcal{N}(\psi)$ are tensor-stable biseparable states with respect to the partition $\overrightarrow{LA}:\overrightarrow{RB}:\overrightarrow{PC}$. The cppp-assisted secret-key-agreement capacities for such a class of multiplex channels are zero.

B. Strong converse bounds on LOCC-assisted private capacity of multiplex channel

We now derive converse and strong converse bounds on an LOCC-assisted secret-key-agreement protocol over a multiplex channel \mathcal{N} .

For an LOCC-assisted secret-key-agreement protocol, by employing Theorem 1 and generalizing the proof arguments of Theorem 2 in Ref. [51] (see also Ref. [50]) to the multiplex scenario, we get the following converse bound (proof in Appendix D 2).

Theorem 3: For a fixed $n, K \in \mathbb{N}, \varepsilon \in (0, 1)$, the following bound holds for an (n, K, ε) protocol for LOCC-assisted secret key agreement over a multiplex $\mathcal{N}_{\overrightarrow{A'}\overrightarrow{B} \rightarrow \overrightarrow{A}\overrightarrow{C}}$:

$$\frac{1}{n} \log_2 K \leq E_{\text{max},E}(\mathcal{N}) + \frac{1}{n} \log_2 \left(\frac{1}{1 - \varepsilon} \right), \quad (48)$$

where the max-relative entropy of entanglement $E_{\text{max},E}(\mathcal{N})$ of the multiplex channel \mathcal{N} is

$$E_{\text{max},E}(\mathcal{N}) := \sup_{\psi \in \text{FS}(\overrightarrow{LA'}:\overrightarrow{RB}:)} \inf_{\sigma \in \text{FS}(\overrightarrow{LA}:\overrightarrow{R}:\overrightarrow{C}:)} D_{\text{max}}(\mathcal{N}(\psi) \parallel \sigma)$$

and it suffices to optimize over pure states ψ .

Remark 4: The bound in Eq. (48) can also be rewritten as

$$1 - \varepsilon \leq 2^{-n(P - E_{\text{max},E}(\mathcal{N}))}, \quad (49)$$

where we have $P = (1/n) \log_2 K$. Thus, if the secret-key-agreement rate P is strictly greater than the max-relative entropy of entanglement $E_{\text{max},E}(\mathcal{N})$ of the (multiplex) channel \mathcal{N} , then the fidelity of the distillation $(1 - \varepsilon)$ decays exponentially fast to zero in the number of channel uses.

An immediate corollary of the above remark is the following strong converse statement.

Corollary 3: The strong converse LOCC-assisted secret-key-agreement capacity of a multiplex channel \mathcal{N} is bounded from above by its max-relative entropy of entanglement:

$$\tilde{P}_{\text{LOCC}}(\mathcal{N}) \leq E_{\text{max},E}(\mathcal{N}). \quad (50)$$

We also have another upper bound on the private capacity of a multiplex channel $\mathcal{N}_{\overrightarrow{A'}\overrightarrow{B} \rightarrow \overrightarrow{A}\overrightarrow{C}}$ with finite-dimensional input and output systems in terms of the regularized relative entropy instead of the max-relative entropy (proof in Appendix D 3).

Theorem 4: For finite Hilbert space dimensions, the asymptotic LOCC-assisted secret-key-agreement capacity of a multiplex channel $\mathcal{N}_{\overrightarrow{A'}\overrightarrow{B} \rightarrow \overrightarrow{A}\overrightarrow{C}}$ is bounded by its regularized relative entropy of entanglement:

$$\tilde{P}_{\text{LOCC}}(\mathcal{N}) \leq E_E^\infty(\mathcal{N}). \quad (51)$$

C. Teleportation-simulable and tele-covariant multiplex channels

For a class of multipartite quantum channels obeying certain symmetries, such as teleportation-simulability [66], the LOCC assistance does not enhance secret-key-agreement capacity, and the original protocol can be reduced to a cppp-assisted secret-key-agreement protocol. This observation for secret communication between two parties over the point-to-point teleportation-simulable channel was first made in Ref. [48].

Definition 7: A multipartite quantum channel $\mathcal{N}_{\overrightarrow{A'}\overrightarrow{B} \rightarrow \overrightarrow{A}\overrightarrow{C}}$ is teleportation simulable with the associated resource state $\theta_{\overrightarrow{LA}\overrightarrow{R}\overrightarrow{C}}$, where $R_b \simeq B_b$ for all $b \in \mathcal{B}$ and $L_a \simeq A'_a$ for all $a \in \mathcal{A}$, if for all input states $\rho_{\overrightarrow{A'}\overrightarrow{B}}$ the following identity holds:

$$\mathcal{N}_{\vec{A}'\vec{B}'\vec{A}\vec{C}}(\rho_{\vec{A}'\vec{B}'}) = \mathcal{T}_{\vec{A}'\vec{L}\vec{A}\vec{B}\vec{R}\vec{C}\vec{A}\vec{C}}(\rho_{\vec{A}'\vec{B}'} \otimes \theta_{\vec{L}\vec{A}\vec{R}\vec{C}}) \quad (52)$$

for some LOCC channel \mathcal{T} with input partition $:\vec{A}'\vec{L}\vec{A}:\vec{B}\vec{R}:\vec{C}$: and output partition $:\vec{A}:\vec{C}$:

Covariant channels.—For each $a \in \mathcal{A}$ and $b \in \mathcal{B}$, let \mathcal{G}_a and \mathcal{G}_b be finite groups of respective sizes G_a and G_b with respective unitary representations $g_a \rightarrow U_{A'_a}(g_a)$ and $g_b \rightarrow U_{B_b}(g_b)$ for all group elements g_a and g_b . Let $W_{A'_a}^{\vec{g}}$ and $W_{B_b}^{\vec{g}}$ be unitary representations for all $a \in \mathcal{A}$ and $c \in \mathcal{C}$, where $\vec{g} = \{g_a, g_b\}_{a,b}$. A multiplex quantum channel $\mathcal{N}_{\vec{A}'\vec{B}'\vec{A}\vec{C}}$ is *covariant* with respect to these representations if the following relation holds for all input states $\rho_{\vec{A}'\vec{B}'}$ and group elements $g_a \in \mathcal{G}_a$ and $g_b \in \mathcal{G}_b$ for all $a \in \mathcal{A}$ and $b \in \mathcal{B}$:

$$\begin{aligned} \mathcal{N}_{\vec{A}'\vec{B}'\vec{A}\vec{C}}\left(\left(\bigotimes_{a \in \mathcal{A}} U_{A'_a}^{g_a} \otimes \bigotimes_{b \in \mathcal{B}} U_{B_b}^{g_b}\right)(\rho_{\vec{A}'\vec{B}'})\right) \\ = \left(\bigotimes_{a \in \mathcal{A}} W_{A'_a}^{\vec{g}} \otimes \bigotimes_{c \in \mathcal{C}} W_{C_c}^{\vec{g}}\right)(\mathcal{N}_{\vec{A}'\vec{B}'\vec{A}\vec{C}}(\rho_{\vec{A}'\vec{B}'})), \quad (53) \end{aligned}$$

where we have used the notation $\mathcal{U}(\cdot) := U(\cdot)U^\dagger$ for unitaries U .

Definition 24: A quantum channel $\mathcal{N}_{\vec{A}'\vec{B}'\vec{A}\vec{C}}$ is called *tele-covariant* if it is covariant with respect to groups $\{\mathcal{G}_a\}_{a \in \mathcal{A}}$ and $\{\mathcal{G}_b\}_{b \in \mathcal{B}}$ that have representations as unitary one-designs; i.e., for all $a \in \mathcal{A}$ and $b \in \mathcal{B}$ as well as states $\rho_{A'_a}$ and ρ_{B_b} , it holds that $(1/G_a) \sum_{g_a \in \mathcal{G}_a} U_{A'_a}^{g_a}(\rho_{A'_a}) = \mathbb{1}/|A_a|$ and $(1/G_b) \sum_{g_b \in \mathcal{G}_b} U_{B_b}^{g_b}(\rho_{B_b}) = \mathbb{1}/|B_b|$, respectively.

The following observation follows from the definition of tele-covariant channels.

Remark 5: Tele-covariance of a channel is with respect to the groups and their unitary representations on the input and output Hilbert spaces of the channel. If associated unitary representations for the tele-covariant channels \mathcal{N}^1 and \mathcal{N}^2 are, respectively, the same on the output Hilbert spaces of \mathcal{N}^1 that are also the input Hilbert spaces for \mathcal{N}^2 , then the composition channel $\mathcal{N} = \mathcal{N}^2 \circ \mathcal{N}^1$ is also tele-covariant.

A quantum channel obtained by the tensor product (superoperation “ \otimes ,” which physically means parallel uses) of tele-covariant channels is also a tele-covariant channel.

The following theorem generalizes the developments in Refs. [51,107–109] (see Appendix D 4 for the proof):

Theorem 5: If a multipartite channel $\mathcal{N}_{\vec{A}'\vec{B}'\vec{A}\vec{C}}$ is tele-covariant, then it is teleportation-simulable with resource state (52) as its Choi state, i.e., $\theta_{\vec{L}\vec{A}\vec{R}\vec{C}} = \mathcal{N}(\Phi_{\vec{L}\vec{R}|\vec{A}'\vec{B}'})$.

Following the approach in Refs. [48,62], we obtain the following theorem:

Theorem 6: The LOCC-assisted secret-key-agreement capacity of a multiplex quantum channel $\mathcal{N}_{\vec{A}'\vec{B}'\vec{A}\vec{C}}$, which is teleportation-simulable with resource state $\theta_{\vec{L}\vec{A}\vec{R}\vec{C}}$, is upper bounded as

$$\hat{P}_{\text{LOCC}}(\mathcal{N}) \leq E_{\text{GE}}^\infty(\vec{L}\vec{A}:\vec{R}:\vec{C})_\theta, \quad (54)$$

where $E^\infty(\vec{A}:\vec{B})_\rho$ is the regularized relative entropy of entanglement of state $\rho_{\vec{A}\vec{B}}$.

For the proof, see Appendix D 4. Using the above theorem, we immediately get the following.

Corollary 4: For a multiplex quantum channel $\mathcal{N}_{\vec{A}'\vec{B}'\vec{A}\vec{C}}$, which is teleportation-simulable with a tensor-stable biseparable resource state, it holds that $\hat{P}_{\text{LOCC}}(\mathcal{N}) = 0$.

Let us note that unlike in Refs. [48,62], which deals with the bipartite relative entropy of entanglement, we do not trivially get a nonregularized bound, which is due to the fact that the definition of biseparability is not tensor stable. If we consider the relative entropy of entanglement with respect to fully separable states, however, we can employ the proof argument of Theorem 4 in Ref. [51] and arrive at the following theorem:

Theorem 7: For a fixed $n, K \in \mathbb{N}, \varepsilon \in (0, 1)$, the following bound holds for an (n, M, ε) protocol for LOCC-assisted secret key agreement over a multiplex teleportation-simulable quantum channel $\mathcal{N}_{\vec{A}'\vec{B}'\vec{A}\vec{C}}$ with the associated resource state $\theta_{\vec{L}\vec{A}\vec{R}\vec{C}}, \forall \alpha > 1$,

$$\frac{1}{n} \log_2 K \leq \tilde{E}_{\alpha, \varepsilon}(\vec{L}\vec{A}:\vec{R}:\vec{C})_\theta + \frac{\alpha}{n(\alpha - 1)} \log_2 \left(\frac{1}{1 - \varepsilon} \right). \quad (55)$$

For the proof, see Appendix D 4. Setting $\alpha = 1 + (1/\sqrt{n})$ and letting $n \rightarrow \infty$, we obtain the following:

Corollary 5: The LOCC-assisted secret-key-agreement capacity of a multiplex channel $\mathcal{N}_{\vec{A}'\vec{B}'\vec{A}\vec{C}}$, which is teleportation-simulable with the resource state $\theta_{\vec{L}\vec{A}\vec{R}\vec{C}}$, is upper bounded as

$$\hat{P}_{\text{LOCC}}(\mathcal{N}) \leq E(\vec{L}\vec{A}:\vec{R}:\vec{C})_\theta, \quad (56)$$

where $E(\vec{A}:\vec{B})_\rho$ is the relative entropy of entanglement of state $\rho_{\vec{A}\vec{B}}$; this bound is also a strong converse bound.

VI. APPLICATION TO OTHER PROTOCOLS

In this section, we exploit the general nature of an LOCC-assisted secret-key-agreement protocol over a multiplex quantum channel. We derive upper bounds on the rates for two-party and conference key distribution for a number of seemingly different protocols that are of wide interest. Such seemingly different quantum key distribution and conference key agreement protocols can be shown to be special types of LOCC-assisted secret-key-agreement protocol over some particular multiplex quantum channels. In particular, we identify protocols like measurement-device-independent quantum key distribution, both in the

bipartite [27,28] and conference setting [30,110,111], as well as for quantum key repeaters, i.e., generalized quantum repeaters with the goal of distributing private states [50,57,77] to be special types of LOCC-assisted secret-key-agreement protocol over some particular multiplex quantum channels. We are able to derive upper bounds on the rates achieved in these protocols by exploiting our results in the previous section. Furthermore, as EPR or GHZ states are special cases of bipartite or multipartite private states, respectively, the same holds for LOCC-assisted quantum communication protocols, where the goal is to distill EPR or GHZ states. By providing a unified approach to such a diverse class of private communication setup, we contribute to a better understanding of limitations on respective protocols. These limitations provide benchmarks on experimental realizations of private communication protocols.

A. Measurement-device-independent QKD

Measurement-device-independent (MDI) QKD is a form of QKD, where the honest parties, Alice and Bob, trust their state preparation but do not trust the detectors [27,28]. In a typical setup of MDI-QKD, such as the ones described in Refs. [27,28], Alice and Bob locally prepare states that they send to a relay station, which might be in the hands of Eve, using channels $\mathcal{N}_{A' \rightarrow A}^1$ and $\mathcal{N}_{B' \rightarrow B}^2$. At the relay station, a joint measurement of the systems AB is performed, e.g., in the Bell basis, the results of which are classical values that are then communicated to Alice and Bob. Alice and Bob use the relay many times and perform classical postprocessing.

A way to incorporate such protocols in our scenario is to identify Alice and Bob as two trusted parties and include the measurement performed by the relay, as well as channels $\mathcal{N}^{1,2}$, into a bipartite quantum-classical (qc) channel

$$\mathcal{N}_{A'B' \rightarrow Z_A Z_B}^{\text{MDI}} := \mathcal{B}_{X \rightarrow Z_A Z_B} \circ \mathcal{M}_{AB \rightarrow X} \circ \mathcal{N}_{A' \rightarrow A}^1 \otimes \mathcal{N}_{B' \rightarrow B}^2, \quad (57)$$

where $\mathcal{M}_{AB \rightarrow X}$ is the quantum instrument (channel) performing a POVM $\{\Lambda^x\}_x$ and writing the output x into a classical register X and $\mathcal{B}_{X \rightarrow Z_A Z_B}$ a classical broadcast channel sending input x to Z_A and Z_B . Registers Z_A and Z_B are received by Alice and Bob, respectively. The channel $\mathcal{N}_{A'B' \rightarrow Z_A Z_B}^{\text{MDI}}$ is a multiplex channel that is a composition of multiplex channels (see Fig. 2).

Application of Theorem 3 for arbitrary systems and Theorem 4 for finite-dimensional systems (as well as the results of Ref. [51–53]) then provides bounds on the achievable key rate in terms of $E_{\max, E}(\mathcal{N}_{A'B' \rightarrow Z_A Z_B}^{\text{MDI}})$ and $E_E^\infty(\mathcal{N}_{A'B' \rightarrow Z_A Z_B}^{\text{MDI}})$, respectively, which can be seen as measures of the entangling capabilities of the measurement $\{\Lambda^x\}_x$. The multiplex quantum channel $\mathcal{N}_{A'B' \rightarrow Z_A Z_B}^{\text{MDI}}$ is tele-covariant if $\mathcal{N}_{1,2}$ as well as \mathcal{M} are tele-covariant, and the

bound reduces to the relative entropy of entanglement of the Choi state of $\mathcal{N}_{A'B' \rightarrow Z_A Z_B}^{\text{MDI}}$.

B. Measurement-device-independent conference key agreement

The concept of MDI-QKD has also been generalized to the multipartite setting [30,110,111]. We assume a setup of MDI conference agreement, where a number of trusted parties A_i , for $i \in [n]$, locally prepare states that they send to a central relay via channels $\mathcal{N}_{A'_1 \rightarrow A_1}^1, \dots, \mathcal{N}_{A'_n \rightarrow A_n}^n$. At the relay, a joint measurement is performed on $A_1 A_2 \dots A_n$, the result of which is broadcast back to the trusted parties. It is straightforward to generalize Eq. (57) to the multipartite case and apply Theorems 3 and 4 (or Theorem 5 for tele-covariant channels) to obtain bounds on the conference key rates.

C. Quantum key repeater

Let us now consider the quantum key repeater. In its simplest setup, there are three parties: Alice, Bob, and Charlie. Alice and Bob are trusted parties who wish to establish a cryptographic key, whereas Charlie is assumed to be cooperative but is not trusted. One could think of Charlie as a telecom provider. There are two quantum channels, $\mathcal{N}_1^{A \rightarrow C_A}$ from Alice to Charlie and $\mathcal{N}_2^{B \rightarrow C_B}$ from Bob to Charlie. Alice and Bob are not connected by a quantum channel and are assumed not to have any pre-shared entanglement. Instead, Alice and Bob locally prepare quantum states, e.g., two singlets $\Phi_{A R_A}^+$ and $\Phi_{B R_B}^+$, and both send a subsystem to Charlie, using the respective channels. This is then followed by an entanglement swapping operation [112], where Charlie performs a joint measurement on the $C_A C_B$ subsystem and communicates the result to Bob, who then performs a unitary on his reference system R_B , which should create entanglement that can be used for a cryptographic key, between Alice and Bob. The key has to be secure even in the case where Charlie's information falls into the hands of Eve.

If the channels $\mathcal{N}_1^{A \rightarrow C_A}$ and $\mathcal{N}_2^{B \rightarrow C_B}$ are too noisy, it might be necessary to use them multiple times and perform an entanglement purification or error-correction protocol before applying the swapping operation. Whereas early quantum repeater protocols [21,22] make use of entanglement purification protocols that require two-way classical communication, between Alice and Charlie and between Charlie and Bob, it is also possible to use error correction that only requires one-way classical communication. Such protocols are known as second- and third-generation repeater protocols (see Ref. [23] and references therein).

By using a large enough number of repeater stations, the key can, in principle, be distributed across arbitrarily long distances. A way to extend a basic three-party repeater protocol to arbitrarily long repeater chains is known as

nested purification [22]. More advanced schemes using error correction and one-way communication have also been developed [23].

As in Refs. [50,57,77], we want to find upper bounds on the rates at which the key can be distributed. Depending on the repeater protocol, there are different ways in which we can describe a quantum key repeater as a multipartite channel and use our results to obtain such bounds. We now describe how a repeater can be described by a bipartite channel. For an alternative way to describe a repeater, we refer to Appendix E.

In order to describe a repeater as a bipartite channel, we consider two trusted parties, Alice and Bob, and a bipartite quantum-to-classical (qc) channel that takes two quantum (and possibly also classical) inputs from Alice and Bob and returns two classical outputs to Alice and Bob, respectively. Such an operation could include the channels from Alice to Charlie and from Bob to Charlie, the measurement performed by Bob, as well as classical communication of the measurement result from Charlie to Alice and Bob. It could also include an error-correction protocol that uses the channels from Alice to Charlie and from Bob to Charlie multiple times and makes use of one-way classical communication from Alice to Charlie and from Bob to Charlie. It is then followed by Charlie's measurement and classical communication to Alice and Bob. Alice and Bob are then allowed to perform LOCC among them but not including Charlie. In the case without error correction, we can define

$$\mathcal{N}_{AB \rightarrow XY}^{\text{repeater}} := \mathcal{M}_{C_A C_B \rightarrow XY} \circ \mathcal{N}_1^{A \rightarrow C_A} \otimes \mathcal{N}_2^{B \rightarrow C_B}, \quad (58)$$

where $\mathcal{M}_{C_A C_B \rightarrow XY}$ describes the measurement and sending of classical messages X and Y to Alice and Bob, respectively. If we add one-way error correction, we get a bipartite channel of the form

$$\mathcal{N}_{A^k B^k X' Y' \rightarrow XY}^{\text{repeater}} := \mathcal{M}_{\tilde{C}_A \tilde{C}_B \rightarrow XY} \circ \mathcal{E}_1^{X' A^k \rightarrow \tilde{C}_A} \otimes \mathcal{E}_2^{Y' B^k \rightarrow \tilde{C}_B}, \quad (59)$$

where $\mathcal{E}_1^{A^k \rightarrow \tilde{C}_A}$ includes k instances of the channel $\mathcal{N}_1^{A \rightarrow C_A}$, the transmission of the classical data X' obtained by Alice's part of the one-way error-correction protocol to Charlie, as well as Charlie's part of the error-correction protocol (Alice's part of the one-way error-correction protocol is included in the LOCC). Note that $\mathcal{E}_2^{Y' B^k \rightarrow \tilde{C}_B}$ is defined in the same way.

By recursively combining the bipartite channels $\mathcal{N}^{\text{repeater}}$, it is possible to derive a bipartite channel $\mathcal{N}^{\text{repeater chain}}$ between Alice and Bob that includes a repeater chain with an arbitrary amount of repeater stations.

Using the results of Refs. [51–53], or Theorem 4, we can obtain upper bounds for key repeater protocols that only involve one-way classical communication from Charlie to Alice and Bob, as considered in Refs. [57,77]. The bounds are given by $\min\{E_{\max,E}(\mathcal{N}^{\text{repeater chain}})\}$,

$E_E^\infty(\mathcal{N}^{\text{repeater chain}})$. By Remark 5, if $\mathcal{N}_{1,2}$ as well as \mathcal{M} are tele-covariant, so is $\mathcal{N}^{\text{repeater chain}}$. Hence, by Theorem 5, the bound reduces to the relative entropy of entanglement of the Choi state of $\mathcal{N}^{\text{repeater chain}}$. Note that, whereas the bounds in Refs. [57,77] only depend on the initial states shared by Alice and Charlie as well as Bob and Charlie, the formulation in terms of a bipartite channel can provide bounds that also depend on the measurement performed by Charlie, as well as operations performed during error correction. The new bounds take into account imperfect measurements and error correction, which provide an additional limitation on the obtainable rate in practical implementations. Our bounds can at least be shown to be comparable with the results of Refs. [57,77] under certain situations of practical interest. For example, our bound is certainly better when $\mathcal{N}_1^{A \rightarrow C_A}$ and $\mathcal{N}_2^{B \rightarrow C_B}$ are identity channels, allowing Alice and Charlie as well as Bob and Charlie to share maximally entangled states, whereas Charlie's measurement is noisy.

D. Limitations on some practical prototypes

In this section, we explore fundamental limitations on some practical prototypes for MDI-QKD protocols between two trusted parties. We first begin by considering photon-based prototypes for which a detailed discussion of the quantum system and transmission noise model can be found in Ref. [67]. In Appendix F, we consider MDI-QKD prototypes with qubit systems and transmission noise models depicted by dephasing or depolarizing channels.

We begin by considering a dual-rail scheme based on single photons to encode the qubits [113]. The dual-rail encoding of a qubit in two orthogonal optical modes can be represented in the computational basis of the qubit system, where only one of the two modes is occupied by a single photon and another mode is vacuum. When these optical modes are two polarization modes—horizontal and vertical—of the light, then we express eigenstates in the computational basis as $|H\rangle$ and $|V\rangle$ for horizontal and vertical polarization. It is also possible to consider frequency-offset modes instead of polarization modes for dual-rail encodings. We assume a noise model for the transmission of a photon through the optical fiber to be a pure-loss bosonic channel with transmissivity η . The inputs to the optical fiber are restricted to a single-photon subspace that is spanned by $|H\rangle$ and $|V\rangle$. The action of this pure-loss channel on a qubit encoded with our dual-rail scheme is identical to an erasure channel [114] \mathcal{E} with erasure parameter $1 - \eta$ and erasure state $|e\rangle$, where $|e\rangle$ is the vacuum state, i.e., zero photon in both modes. We note that an erasure channel is tele-covariant.

Two trusted parties \mathbf{A}_i , $i \in [2]$, use the above-mentioned polarization-based dual-rail photons to transmit their qubit systems to Charlie at the measurement-relay station, through the optical fibers with respective transmissivities

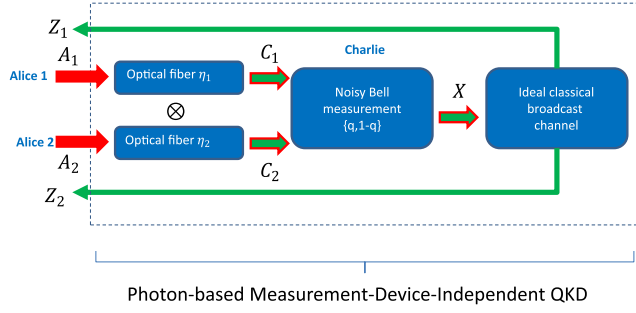


FIG. 4. Pictorial illustration of our photon-based MDI-QKD between two parties using the dual-rail encoding scheme.

η_i (see Fig. 4 for MDI-QKD). We make a simplistic noise model assumption on the measurement channel $\mathcal{M}_{C_i \rightarrow X}$ by Charlie: It can perform perfect qubit Bell measurement for bipartite MDI-QKD, respectively, with probability q , whereas with probability $1 - q$ for the failed measurement, we assume the relay station signals $|\perp\rangle\langle\perp|_X$ to the users. In addition, we can safely assume classical communication $X \rightarrow \vec{Z}$ among all parties to be clean (noiseless) as they do not require any quantum resource. Finally, for simplicity, we assume that error-correcting local operations for all parties can be made perfectly.

To calculate the upper bound on the MDI-QKD capacity, it suffices to consider the relative entropy of entanglement of the Choi state of the associated multiplex channel $\mathcal{N}_{\vec{A} \rightarrow \vec{Z}}^{\text{MDI}, \mathcal{E}}$ as it is tele-covariant. Notice that the action of the erasure channel $\mathcal{E}_{A_i \rightarrow C_i}$ on $D_i \in \{|H\rangle\langle H|_{A_i}, |H\rangle\langle V|_{A_i}, |V\rangle\langle H|_{A_i}, |V\rangle\langle V|_{A_i}\}$ is given as

$$\mathcal{E}_{A_i \rightarrow C_i}(D_i) = \eta_i D_i + (1 - \eta_i) \text{Tr}[D_i] |e\rangle\langle e|_{C_i}. \quad (60)$$

Then, the Choi state $J_{\vec{L}\vec{C}}^{\mathcal{E}}$ of $\bigotimes_{i=1}^2 \mathcal{E}_{A_i \rightarrow C_i}$ is

$$J_{\vec{L}\vec{C}}^{\mathcal{E}} = \bigotimes_{i=1}^2 \left(\eta_i \Phi_{L_i C_i}^+ + (1 - \eta_i) \frac{\mathbb{1}_{L_i}}{2} \otimes |e\rangle\langle e|_{C_i} \right). \quad (61)$$

For the bipartite MDI-QKD,

$$\begin{aligned} \mathcal{M}_{C_1 C_2 \rightarrow X}(\cdot) &= q \sum_{j=1}^4 \text{Tr}[\Phi^{(j)}(\cdot) \Phi^{(j)}] |j\rangle\langle j|_X \\ &+ (1 - q) \text{Tr}[\cdot] \otimes |\perp\rangle\langle\perp|_X, \end{aligned} \quad (62)$$

where $\{\Phi_{C_1 C_2}^{(j)}\}_{j=1}^4$ is the Bell measurement, which is a projective measurement. Here, $\{\Phi_{C_1 C_2}^{(j)}\}_{j=1}^4$ represents the set of maximally entangled states $\{\Phi^+, \Phi^-, \Psi^+, \Psi^-\}$ for two-qubit systems and $|\perp\rangle\langle\perp|_j$. We note that the Bell measurement is tele-covariant. Upon action of the

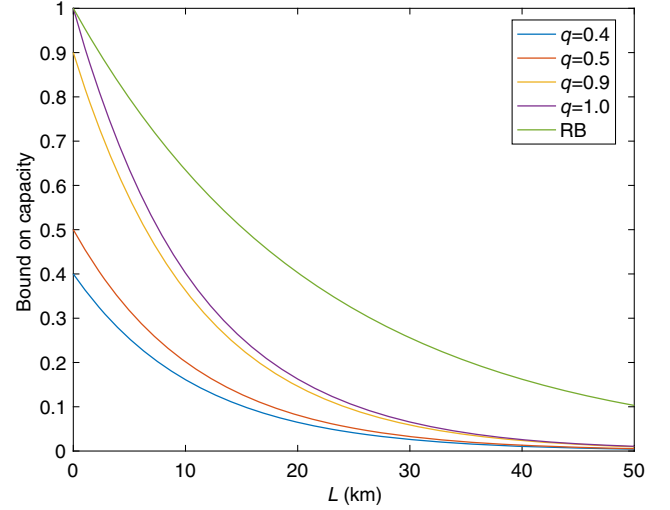


FIG. 5. Rate-distance trade-off comparison between our bound (64) (blue, red, yellow, and purple lines) and the RB bound (green line) for the MDI-QKD protocol for our photon-based prototype.

measurement channel $\mathcal{M}_{C_1 C_2 \rightarrow X}$ on the state $J_{L_1 L_2 C_1 C_2}^{\mathcal{E}}$ [Eq. (61)], the output state is essentially of the form (see Ref. [67])

$$q\eta_1\eta_2 \frac{1}{4} \sum_{j=1}^4 \Phi_{L_1 L_2}^{(j)} \otimes |j\rangle\langle j|_X + (1 - q\eta_1\eta_2) \frac{\mathbb{1}_{L_1 L_2}}{4} \otimes |\perp\rangle\langle\perp|_X, \quad (63)$$

which implies that the relative entropy of entanglement of the Choi state of $\mathcal{N}_{A_1 A_2 \rightarrow Z_1 Z_2}^{\text{MDI}, \mathcal{E}}$ is $q\eta_1\eta_2$. Employing Theorem 7, the bipartite MDI-QKD capacity for the given MDI-QKD prototype with erasure channels is

$$\tilde{P}_{\text{LOCC}}(\mathcal{N}^{\text{MDI}, \{\mathcal{E}_i\}_{i=1}^2}) = q\eta_1\eta_2, \quad (64)$$

as $q\eta_1\eta_2$ bits is an achievable rate for the given setup (see Refs. [48,49,105] for the private capacities of $\mathcal{E}_{A_i \rightarrow C_i}$). Notice that $q\eta_1\eta_2$ is a strong converse bound.

For bipartite MDI-QKD (see Fig. 4), using the results of Ref. [48,105], we get upper bound (RB) on the bipartite MDI-QKD capacity as $\min\{\eta_1, \eta_2\}$ (e.g., see Refs. [50,60]). This bound is always looser than our strong converse upper bound $q\eta_1\eta_2$ bits, for all practical purposes. In Fig. 5, we plot the rate-distance trade-off (secret key capacity versus distance L in km) for our bound in Eq. (64) when $n=2$, $\eta_1 = \eta_2 = \exp(-\alpha L)$, and $\alpha = (1/22 \text{ km})$ and compare it with the upper bound (RB) η_1 (since $\eta_1 = \eta_2$).

We note that, whereas there now exist variants of MDI-QKD schemes or setups that can achieve the repeaterless bound, e.g., Refs. [31,32,34], the dual-rail protocols we consider here, while being suboptimal, may be easier to implement practically. In particular, implementation of a

twin-field protocol requires long-distance phase stabilization, which can be challenging [115]. We showcase here the ability to get nontrivial upper bounds for a specific, suboptimal implementation of QKD schemes. These nontrivial upper bounds are derived from a universal framework, which illustrates the usefulness of the framework we have proposed.

VII. LOWER BOUNDS ON PRIVACY

In this section, we derive lower bounds on the secret-key-agreement rate of a multiplex channel achievable by means of cPPP, in the sense of Ref. [68]. This is a generalization of the lower bound presented in Ref. [14] from multipartite states to multiplex channels, as well as a generalization of the lower bounds on one-to-one channels presented in Ref. [69] to the multiplex case.

The DW protocol [68], which is considered with bipartite states, only uses one-way communication from Alice to Bob. In Ref. [69], which is concerned with one-to-one channels, direct and reverse scenarios are considered. The former corresponds to the case where the quantum channel and the classical communication are oriented in the same direction. The latter corresponds to the case where the two are oriented in opposite directions. In Ref. [14], the DW protocol is generalized to multipartite states by selecting one distributing party, which performs the DW protocol with all remaining parties simultaneously.

We now generalize this result to the setting of multiplex channels. We begin with a fully separable pure state $\phi^n \in \text{FS}(\overrightarrow{A^n L} : \overrightarrow{B^n R} : \overrightarrow{P})$. Here, the notation $\overrightarrow{X^n}$ means we consider n copies of all subsystems X_1, \dots, X_M . Application of n copies of the isometric extension of the multiplex channel $\mathcal{N}_{A^n B \rightarrow A^n C}$ results in a pure state $\psi^n_{\overrightarrow{A^n L} : \overrightarrow{R} : \overrightarrow{C^n P} : E^n}$. Let us now choose one party, \mathbf{X}_i , $i \in \{1, \dots, M\}$, as the distributing party. Party \mathbf{X}_i performs a POVM $\mathcal{Q} = \{Q_x\}$ with a corresponding random variable $X = \{x, p(x)\}$ on her subsystem, resulting in a classical-quantum-...-quantum (cq) state

$$\omega_{\text{cq}} = \sum_x p(x) |x\rangle\langle x|_X \otimes \omega^x, \quad (65)$$

where ω^x is the post-measurement state of the remaining parties and Eve. Party \mathbf{X}_i then processes X using classical channels $X \rightarrow Y$ and $Y \rightarrow Z$, where $Y = \{y, q(y)\}$ and $Z = \{z, r(z)\}$ are classical random variables. Here, Y is kept by party \mathbf{X}_i (to be used for the key), and Z is broadcast to all other trusted parties (and Eve). Upon receiving Z , the other parties then perform their respective POVMs, with the goal of estimating the key variable Y . Thus, as shown in Ref. [68], every trusted party \mathbf{X}_j , where $i \neq j \in \{1, \dots, M\}$, obtains a common key with \mathbf{X} at a rate $r_n^{i \rightarrow j}$ of

$$r_n^{i \rightarrow j} = \frac{1}{n} (I(Y : \mathbf{X}_j | Z)_{\tilde{\omega}_{\text{cq}}} - I(Y : E^n | Z)_{\tilde{\omega}_{\text{cq}}}), \quad (66)$$

where, in a slight abuse of notation, we use \mathbf{X}_j as a placeholder for $A_j^n L_j$, R_j , or $C_j^n P_j$, depending if \mathbf{X}_j is in $\{\mathbf{A}_a\}_a$, $\{\mathbf{B}_b\}_b$, or $\{\mathbf{C}_c\}_c$, respectively. The second and third cases correspond to the reverse and direct scenarios in Ref. [69], respectively, whereas

$$\tilde{\omega}_{\text{cq}} = \sum_{xyz} r(z|y) q(y|x) p(x) |xyz\rangle\langle xyz| \otimes \omega^x. \quad (67)$$

Equation (66) has to be maximized over all free input states $\phi^n \in \text{FS}(\overrightarrow{A^n L} : \overrightarrow{B^n R} : \overrightarrow{P})$, POVMs \mathcal{Q} , as well as classical channels $X \rightarrow Y$ and $Y \rightarrow Z$. As discussed in Ref. [14], a conference key among all trusted parties can be obtained at the worst-case rate between any pair $(\mathbf{X}_i, \mathbf{X}_j)$. We also have the freedom to choose the distributing party. Putting it all together, we can achieve the following rate of the conference key:

$$\hat{P}_{\text{cPPP}}^{\mathcal{N}} \geq \max_i \min_j \lim_{n \rightarrow \infty} \max_{\phi^n, \mathcal{Q}, \text{POVM}} r_n^{i \rightarrow j}, \quad (68)$$

with $\phi^n \in \text{FS}(\overrightarrow{A^n L} : \overrightarrow{B^n R} : \overrightarrow{P})$. Note that in the case of a single-sender-single-receiver channel $\mathcal{N} : B \rightarrow C$, this reduces to the maximum of the direct and reverse key rates presented in Ref. [69].

Next, we propose an alternative generalization of the DW protocol to the case of multipartite states and multiplex channels. The rough idea is that, instead of performing the DW protocol simultaneously with all other parties after her measurement, the distributing party performs a one-way protocol with a second party, who then performs a one-way protocol with a third party, and the iteration continues. In particular, the random variables obtained in all previous measurements can be passed on in every classical communication step, so a party can adapt her measurement depending on all previous measurements instead of the first measurement as in the protocol described in Ref. [14].

We now describe the protocol in detail: As before, we begin with a fully separable pure state $\phi^n \in \text{FS}(\overrightarrow{A^n L} : \overrightarrow{B^n R} : \overrightarrow{P})$ and apply n copies of the isometric extension of the multiplex channel $\mathcal{N}_{A^n B \rightarrow A^n C}$, resulting in a pure state $\psi^n_{\overrightarrow{A^n L} : \overrightarrow{R} : \overrightarrow{C^n P} : E^n}$.

Now, assume that we are given some permutation $\sigma : \{1, \dots, M\} \rightarrow \{\sigma(1), \dots, \sigma(M)\}$, which determines the order in which the parties participate in the protocol. Party $\mathbf{X}_{\sigma(1)}$ begins by performing a POVM $\mathcal{Q}^{(1)}$ on her share of ψ^n , i.e., on subsystem $A_{\sigma(1)}^n L_{\sigma(1)}$, $R_{\sigma(1)}$, or $C_{\sigma(1)}^n P_{\sigma(1)}$, depending on which kind of party $\mathbf{X}_{\sigma(1)}$ is. This results in a random variable $X^{(1)} = \{p_1(x_1), x_1\}$. The corresponding classical-quantum-...-quantum (cq) state is

$$\omega_{\text{cq}}^{(1)} = \sum_{x_1} p_1(x_1) |x_1\rangle \langle x_1|_{X^{(1)}} \otimes \omega^{x_1}. \quad (69)$$

Party $\mathbf{X}_{\sigma(1)}$ then performs classical channels $X^{(1)} \rightarrow Y^{(1)} \rightarrow Z^{(1)}$, keeping the random variable $Y^{(1)}$ and sending $Z^{(1)}$ to party $\mathbf{X}_{\sigma(2)}$. The corresponding cq state is then given by

$$\tilde{\omega}_{\text{cq}}^{(1)} = \sum_{x_1 y_1 z_1} r_1(z_1|y_1) q_1(y_1|x_1) p_1(x_1) |x_1 y_1 z_1\rangle \langle x_1 y_1 z_1| \otimes \omega^{x_1}, \quad (70)$$

where ω^{x_1} is the state of the remaining parties and Eve. Next, party $\mathbf{X}_{\sigma(2)}$ performs a POVM $\mathcal{Q}_{Z^{(1)}}$ on her share of ω^{x_1} , which provides the random variable $X^{(2)}$. Party $\mathbf{X}_{\sigma(2)}$ then performs classical channels $Z^{(1)} X^{(2)} \rightarrow Y^{(2)} \rightarrow Z^{(2)}$, keeps $Y^{(2)}$ for herself, and sends $Z^{(2)}$ to the next party $\mathbf{X}_{\sigma(3)}$, who applies the same procedure. The protocol is repeated until party $\mathbf{X}_{\sigma(M)}$ receives $Z^{(M-1)}$, followed by her POVM and postprocessing. The cq after $k \in \{1, \dots, M\}$ measurements and postprocessing steps is given by

$$\tilde{\omega}_{\text{cq}}^{(k)} = \sum_{\substack{x_1 \dots x_k \\ y_1 \dots y_k \\ z_1 \dots z_k}} \tilde{p}_{x_1 y_1 z_1 \dots x_k y_k z_k} |x_1 y_1 z_1 \dots x_k y_k z_k\rangle \langle x_1 y_1 z_1 \dots x_k y_k z_k| \otimes \omega^{x_1 \dots x_k}, \quad (71)$$

where we have defined, recursively,

$$\tilde{p}_{x_1 y_1 z_1 \dots x_k y_k z_k} = r_k(z_k|y_k) q_k(y_k|x_k z_{k-1}) p_k(x_k) \times \tilde{p}_{x_1 y_1 z_1 \dots x_{k-1} y_{k-1} z_{k-1}}. \quad (72)$$

Parties $\mathbf{X}_{\sigma(k)}$ and $\mathbf{X}_{\sigma(k+1)}$ can establish a key rate of [68]

$$r^{\sigma(k) \rightarrow \sigma(k+1)} = \frac{1}{n} (I(Y^{(k)} : \mathbf{X}_{\sigma(k+1)} | Z^{(k)})_{\tilde{\omega}_{\text{cq}}^{(k)}} - I(Y^{(k)} : E^n | Z^{(k)})_{\tilde{\omega}_{\text{cq}}^{(k)}}). \quad (73)$$

We can again maximize over all free input states, POVMs, as well as classical channels and consider the worst-case rate between any pair $(\mathbf{X}_i, \mathbf{X}_j)$. Furthermore, we have the freedom to choose the order of the parties. Putting it all together, we can achieve the following rate of the conference key:

$$\hat{P}_{\text{cPPP}}^{\mathcal{N}} \geq \max_{\sigma \in \text{perm}} \min_k \lim_{n \rightarrow \infty} \max_{\substack{\phi^n, \mathcal{Q}^{(1)}, \dots, \mathcal{Q}^{(k)} \text{ POVM} \\ X^{(1)} \rightarrow Y^{(1)} \rightarrow Z^{(1)} \\ X^{(2)} | Z^{(1)} \rightarrow Y^{(2)} \rightarrow Z^{(2)} \\ X^{(k)} | Z^{(k-1)} \rightarrow Y^{(k)} \rightarrow Z^{(k)}}} r^{\sigma(k) \rightarrow \sigma(k+1)}, \quad (74)$$

with $\phi^n \in \text{FS}(\overrightarrow{A^n L} : \overrightarrow{B^n R} : \overrightarrow{P} :)$.

A. Lower bound for bidirectional network via spanning tree

In this section, we observe that one can tighten the lower bounds presented in the previous section for a particular multiplex channel called the bidirectional network (BN). In the BN, each of the nodes is connected with its neighbors by product bidirectional channels, which are specific bidirectional channels that is a tensor product of two point-to-point channels directed in opposite ways from each other.

We first observe that BN is a particular case of a multiplex channel (call it \mathcal{N}). Indeed, in this case, all the parties are of type \mathcal{A} ; i.e., they can read and write. The rule is that each party represented in the network as a vertex v has $\text{deg}(v)$ of neighbors (see Ref. [116] for an introduction to graph theory). Each party is assumed to write to her neighbors and also receive from these neighbors some quantum data. We now present a tighter bound on the private capacity of \mathcal{N} based on the above exemplary graph.

To be more specific, the BN can be represented by a weighted, directed multigraph $G = (E, V)$ in which each edge $e_{ij} = (v_i, v_j) \in E$ represents a product bidirectional channel $\Lambda_{ij} = \Lambda_{i \rightarrow j} \otimes \Lambda_{j \rightarrow i}$ with weight $W: E \mapsto R_+$ such that $W(e_{ij}) = W(e_{ji}) = \mathcal{P}(\Lambda_{i \rightarrow j}) = \mathcal{P}(\Lambda_{j \rightarrow i})$ (this edge can be represented by two directed edges: one from v_i to v_j and the other vice versa; hence, the structure is directed multigraph). Each product bidirectional channel has in both directions the same private capacity (that, however, may differ for different channels). By convention, we consider edges with index $i > j$ only. The number of nodes in the network is denoted as $|V| := n$ and the number of edges as $|E| := m$.

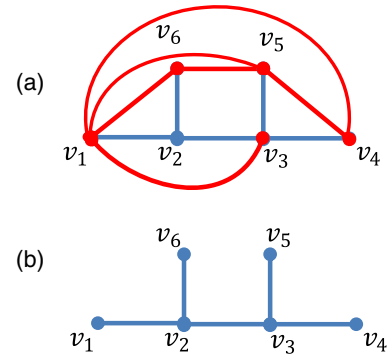


FIG. 6. (a) Exemplary graph. Red edges correspond to private capacity 1 and blue to private capacity 2. The first strategy for obtaining the conference key uses a vertex connected to all others and reaches the suboptimal rate $\min\{w(e_{ij}) : (v_1, v_6), (v_1, v_5), (v_1, v_4), (v_1, v_3), (v_1, v_2)\} = 1$. The same happens for any path, which inevitably has to pass through some red edge. The solution is a tree, which is a spanning tree of this graph, and it contains no red edge (b). Traversing the edges of this tree is equivalent to the breadth-first search.

As a motivation for the next consideration, for such multiplex channels, the bounds given in inequalities (68) and (74) above are not tight. We exemplify this on the graph presented in Fig. 6(a). Namely, we assume that each red edge of the graph G depicted there represents a (bidirectional) channel with private capacity 1, while each blue edge is with capacity equal to 2. We do not depict all other edges (connections) as they have zero private capacity by assumption. We now make two observations: (i) The approach of inequality (68) would yield overall secret key agreement at rate 1, as the only node connected to all others in G (v_1) contains (in fact, more than one) red edge. (ii) We observe, by direct inspection, that every path connecting all vertices also contains at least one red edge.

On the other hand, there is a set of vertices [depicted with edges in Fig. 6(b)] that forms the so-called spanning tree $T := (V_T, E_T) \subset G$ of the graph G . The spanning tree is an acyclic connected subgraph of G , and the word “spanning” refers to the fact that all the vertices of the graph G belong to V_T . It is easy to see that starting from any vertex of this tree, by the breadth-first search algorithm, one can visit all its edges, and one can obtain the conference key at rate 2 (see Ref. [117] for an introduction to algorithms).

As a generalization of this idea, one easily comes up with the following lower bound, which is the main result of this section:

$$\hat{P}_{\text{cPPP}}^{\mathcal{N}} \geq \max_{T \subseteq G} \min_{t \in V_T, t' \in N[t]} \lim_{n \rightarrow \infty} \max_{\substack{\phi^n, \mathcal{Q}^{(1)}, \dots, \mathcal{Q}^{(|V_T|)} \text{ POVM} \\ X^{(v_1)} \rightarrow Y^{(v_1)} \rightarrow Z^{(v_1)}, \\ X^{(N[v_1])} Z^{(v_1)} \rightarrow Y^{(N[v_1])} \rightarrow Z^{(N[v_1])} |_{\text{deg} \geq 2}, \\ X^{(N_2[v_1])} Z^{(N[v_1])} \rightarrow Y^{(N_2[v_1])} \rightarrow Z^{(N_2[v_1])} |_{\text{deg} \geq 2}, \\ X^{N_l[v_1]} Z^{N_l[v_1]} \rightarrow Y^{N_l[v_1]}}} r^{\sigma(t) \rightarrow \sigma(t')}, \quad (75)$$

where $1 \leq l \leq n$ is an index that counts how many times the breadth-first search needs to be invoked in order to traverse all the edges of the spanning tree T . For ease of notation, T is meant to be a rooted, without loss of generality, at vertex v_1 . By $N[v]$, we mean the proper neighborhood of the node v (i.e., the set of all vertices that are connected by a single edge with v). In a rooted tree, every vertex is reachable from the root vertex by a path. By $N_i[v_1]$, we mean the set of vertices reachable from vertex v_1 by a path of length i . Owing to this notation, $N[v_1] \equiv N_1[v_1]$, while all vertices achievable from v_1 by traversing two edges belong to $N_2[v_1]$ and so on.

The first inner maximization needs to be understood inductively. The first step is obvious: We begin with an arbitrary vertex $v_1 \in V_T$. The party X_{v_1} who is at node v_1 performs a POVM \mathcal{Q}_1 , which produces a random variable $X^{(v_1)}$. She processes this variable further to obtain $Y^{(v_1)}$ and sends a communication in the form of a variable $Z^{(v_1)}$. The latter variable is broadcast to all the next neighbors of v_1 , i.e., $N[v_1] \setminus \{v_1\}$. Furthermore, if at step $m-1$ the form of operations and communication between the nodes has concise notation $X^{S_m} Z^{S_{m-1}} \rightarrow Y^{S_m} \rightarrow Z^{S_m} |_{\text{deg} \geq 2}$, then the next level of nesting, i.e.,

$$X^{N[S_m]} Z^{S_m} \rightarrow Y^{N[S_m]} \rightarrow (Z^{N[S_m]} |_{\text{deg} \geq 2}), \quad (76)$$

has to be understood as a short notation of the following postprocessing at a number of nodes from the set $N[S_m] = \{s_1, \dots, s_r\}$ with $r = |N[S_m]|$:

$$\forall_{s_i \in N[S_m]} : \text{deg}(s_i) \geq 2 \quad X^{(s_i)} Z^{N[s_i] \cap p(s_i)} \rightarrow Y^{(s_i)} \rightarrow Z^{(s_i)}$$

$$\forall_{s_i \in N[S_m]} : \text{deg}(s_i) = 1 \quad X^{(s_i)} Z^{N[s_i] \cap p(s_i)} \rightarrow Y^{(s_i)},$$

where $p(s_i)$ denotes the parent vertex of the vertex s_i , that is, the unique vertex belonging to the neighborhood that is the closest to the root v_1 in terms of traversed edges.

The above description means that if some vertex of the tree is of degree equal to 1, it has no further children in the tree to pass useful information contained in the Z -type variable, while all vertices with larger degree than 1 need to broadcast appropriate data to their further neighbors in the tree.

We exemplify the lower bound given in inequality (75) with the broadcast network depicted on Fig. 6. Let us first focus on involved sets of vertices in the process of the breadth-first search over the tree T . The set of vertices of the spanning tree T reads $\{v_1, \dots, v_6\}$. As the root vertex, we choose v_1 . Next, $N_1[v_1] = \{v_2\}$, $N_2[v_1] = \{v_3, v_6\}$ and $N_3[v_1] = \{v_4, v_5\}$. In this case, the presented lower bound reads

$$\hat{P}_{\text{cPPP}}^{\mathcal{N}} \geq \max_{T \subseteq G} \min_{t \in V_T, t' \in N[t]} \lim_{n \rightarrow \infty} \max_{\substack{\phi^n, \mathcal{Q}^{(1)}, \dots, \mathcal{Q}^{(6)} \text{ POVM} \\ X^{(v_1)} \rightarrow Y^{(v_1)} \rightarrow Z^{(v_1)}, \\ X^{(v_2)} Z^{(v_1)} \rightarrow Y^{(v_2)} \rightarrow Z^{(v_2)}, \\ X^{(v_3)} Z^{(v_2)} \rightarrow Y^{(v_3)} \rightarrow Z^{(v_3)}, \\ X^{(v_6)} Z^{(v_2)} \rightarrow Y^{(v_6)}, \\ X^{(v_5)} Z^{(v_3)} \rightarrow Y^{(v_5)}, \\ X^{(v_4)} Z^{(v_3)} \rightarrow Y^{(v_4)}}} r^{\sigma(t) \rightarrow \sigma(t')}. \quad (77)$$

In Appendix G, we briefly comment on the complexity of finding a subgraph, which allows us to realize the

conference key agreement with the capacity indicated by the inequality (75).

VIII. KEY DISTILLATION FROM STATES

In this section, we concentrate on the subject of the distillation of secret keys from quantum states. An (n, K, ε) LOCC conference key distillation begins with M parties A_i for $i \in [M]$ sharing n copies of the M -partite quantum state $\rho_{\vec{A}}$, to which they apply an LOCC channel $\mathcal{L}_{A^{\otimes n} \rightarrow SK}$. The resulting output state satisfies the following condition:

$$F(\mathcal{L}_{A^{\otimes n} \rightarrow SK}(\rho_{\vec{A}}^{\otimes n}), \gamma_{KS}^-) \geq 1 - \varepsilon. \quad (78)$$

The one-shot secret-key-distillation rate from a single copy of a multipartite quantum state $K_D^{(1,\varepsilon)}$ is upper bounded as follows (cf. Sec. V).

Theorem 8: For any fixed $\varepsilon \in (0, 1)$, the achievable region of secret key agreement from a single copy of an arbitrary multipartite quantum state $\rho_{\vec{A}}$ satisfies

$$K_D^{(1,\varepsilon)}(\rho) \leq E_{h,GE}^\varepsilon(\vec{A})_\rho, \quad (79)$$

where

$$E_{h,GE}^\varepsilon(\vec{A})_\rho := \inf_{\sigma \in \text{BS}(\vec{A})} D_h^\varepsilon(\rho \| \sigma) \quad (80)$$

is the ε -hypothesis-testing relative entropy of genuine entanglement of multipartite state $\rho_{\vec{A}}$.

Proof.—The proof argument is the same as that of Theorem 2, so we omit the proof here. ■

In the asymptotic limit, the rate $K_D^{(n,\varepsilon)}$ satisfies

$$\inf_{\varepsilon > 0} \limsup_{n \rightarrow \infty} \frac{1}{n} K_D^{(n,\varepsilon)}(\rho^{\otimes n}) = K_D(\rho), \quad (81)$$

which follows directly from the definition of the secret key rate K_D [14].

Using the same argument as in the proof of Theorem 6 in Sec. V C, we can also get the following asymptotic bound, which is generalized in Theorem 9 of Ref. [62]:

Proposition 2: For an m -partite state $\rho_{\vec{A}}$, it holds that

$$K_D(\rho_{\vec{A}}) \leq E_{GE}^\infty(\rho_{\vec{A}}). \quad (82)$$

In general, to share the conference key, it is necessary for the honest parties to distill genuine multipartite entanglement.

Corollary 6: For a tensor-stable biseparable state $\rho_{\vec{A}}$, it holds that $K_D(\rho_{\vec{A}}) = 0$.

The above Corollary of Theorem 8 is precisely due to the infimum over biseparable states. However, already in the

tripartite setting, there are two nonequivalent families of three-partite genuinely entangled states, that is, Φ_M^{GHZ} -type and Φ_M^{W} -type states [79,85,118–121]. Both families of states contain states that are maximally entangled; however, they cannot be transformed with LOCC one into another at unit rate [81,84,86,87,122,123]. As the perfect Φ_M^{GHZ} state plays a role of the honest (or perfect) implementation of conference quantum key agreement protocols, the distillation of Φ_3^{GHZ} states from Φ_3^{W} states has been intensively studied [80–84,86]. In particular, recalling Example 11 of Ref. [80], it is known that one cannot transform a single Φ_3^{W} state into a Φ_3^{GHZ} state even in a probabilistic manner. However, according to Theorem 2 of Ref. [80], the calculated asymptotic rate for conversion from Φ_3^{W} to Φ_3^{GHZ} due to certain protocols is approximately 0.643 (per copy), which constitutes a lower bound for the general case. Another complementary lower bound has been provided in Ref. [86].

Surprisingly, in the one-shot regime, distillation of Φ_3^{GHZ} states from Φ_3^{W} states, and therefore of the secret key, is still possible. To accomplish this task, it is sufficient to consider the initial state as being made up of two copies of the Φ_3^{W} state. Then, using results in Ref. [81], it follows that we can obtain two Φ_2^+ states in two distinct bipartite systems with a probability that is arbitrarily close to $\frac{2}{3}$; having this in mind, one can obtain Φ_3^{GHZ} by employing ancilla and the entanglement swapping protocol [112]. In this way, we calculate a lower bound on the distillation of Φ_3^{GHZ} states from two copies of the Φ_3^{W} state in a one-shot regime (one Φ_3^{GHZ} state with probability $\frac{2}{3}$ from two Φ_3^{W} states). This lower bound can be compared with the upper bound in Theorem 8 given above.

Nevertheless, distillation of Φ_M^{GHZ} states is only an example of a key distillation technique [13,14,84,86,124–127]. A more general conference key agreement scenario of our interest incorporates distillation of twisted Φ_M^{GHZ} states (see Definition 3) [14,61,119,128,129]. In that case, an approach for upper bounding conference key rates that is different than the estimation of Φ_M^{W} to Φ_M^{GHZ} conversion rates is required. This approach corresponds to a possible gap between rates of Φ_M^{GHZ} (that can be distilled) and secret key distillation. Since the Φ_M^{GHZ} state is an instance of a private state, an upper bound on the conference key rate is also an upper bound on the distillation rate from any state. For plotting our numerical results, we concentrate on secret key distillation from n copies of the Φ_M^{W} state in order to compare with other limitations discussed in this section.

The upper bound in Theorem 8 has optimization over all possible biseparable states. Computation of the exact value of the bound given in Eq. (79) need not be feasible in general. As we take the infimum in Eq. (80), we can obtain non-trivial upper bounds on the upper bound given in Eq. (79) by considering optimization over suitable subsets of biseparable states. We make an educated guess for the

form of biseparable state to yield a non-trivial upper bound. We remark here that the set of biseparable states is not closed under a tensor product, so we have to find different states for any tensor power n of Φ_M^W or Φ_M^{GHZ} states. We devise two families of biseparable states, $\pi_W^{n,M}$ and $\pi_{\text{GHZ}}^{n,M}$, adjusted to both number of copies, n , and number of parties, M ,

$$\pi_{\text{GHZ}}^{n,M} := \frac{1}{M} \sum_{i=1}^M \left[\mathcal{S}_{1,i} \left(\frac{I}{2} \otimes \Phi_{M-1}^{\text{GHZ}} \right) \right]^{\otimes n}, \quad (83)$$

$$\pi_W^{n,M} := \frac{1}{M} \sum_{i=1}^M (\mathcal{S}_{1,i} (|0\rangle\langle 0| \otimes \Phi_{M-1}^W))^{\otimes n}, \quad (84)$$

where the operator $\mathcal{S}_{1,i}$ swaps the qubit of the first party with the qubit of the i th party. The choice of $\pi_{\text{GHZ}}^{n,M}$ and $\pi_W^{n,M}$ states is motivated by keeping the correlation between $M-1$ parties most similar to those in Φ_M^{GHZ} or Φ_M^W states while keeping one party explicitly separated. Additionally, $\pi_{\text{GHZ}}^{n,M}$ and $\pi_W^{n,M}$ states, by definition, are symmetric with respect to permutation of parties because of permutations with $\mathcal{S}_{1,i}$.

We would like to point out that the $\pi_W^{1,3}$ presented here is closer to the Φ_3^W state in the Hilbert-Schmidt norm than the state (let us call it Υ) in Ref. [130], even though the state constructed there was supposed to be the biseparable state closest to Φ_3^W in the Hilbert-Schmidt norm. This result is due to different definitions of biseparability; the state in Ref. [130] is a tensor product with respect to one of the cuts, whereas we make use of the convexity of the set of biseparable states. Indeed, our states are biseparable by construction (see Sec. IV A).

The upper bound on the asymptotic secret key rate can be compared with the lower bound on asymptotic Φ_3^{GHZ} states from Φ_3^W state distillation [80] in the following way. First, we notice that if two parties unite, then the $M-1$ -partite key is no less than the initial M -partite key because the set of operations of the M -partite LOCC protocol is a strict subset of the set of operations for the case in which two parties, i and j , are in the same laboratory. We have the following Proposition:

Proposition 3: For any M -partite state $\rho_{[M]}$, the asymptotic secret-key-agreement rate satisfies the following inequality:

$$\max_k K_D(\rho_{[M+1]_k}) \leq K_D(\rho_{[M]}) \leq \min_{i,j} K_D(\rho_{[M-1]_{ij}}), \quad (85)$$

where $[M] = [1, \dots, M]$ and $[M-1]_{ij} = [1, \dots, i-1, (i, j), i+1, \dots, j-1, j+1, \dots, M]$ indicate a state $\rho_{[M-1]}$ in which subsystems i and j are merged. Analogously, $[M+1]_k = [1, \dots, k-1, k_1, k_2, k+1, \dots, M+1]$ indicates the state in which subsystem k is split into systems k_1 and k_2 .

Proof.—It is enough to notice that the class of LOCC protocols involved in the definition of $K_D(\rho_{[M]})$ is strictly contained in the class of protocols involved in the definition of $K_D(\rho_{[M-1]_{ij}})$. Indeed, the merged parties can still simulate any operation from the former class; however, together, they can perform many more operations, including global quantum operations on all merged subsystems together. Since K_D is defined as the supremum of the key rate over such protocols, the upper bound follows. For the lower bound, it is enough to notice that by splitting subsystem(s) of ρ , we restrict the class of operations that can be used to distill the key. ■

We immediately observe that Proposition 3 provides a whole family of nonequivalent upper bounds. To see this, one can consider a state that is not invariant under permutations. What is more, one can continue merging as long as there is still two or more subsystems left.

Corollary 7: For any M -partite state $\rho_{[M]}$ defined on the Hilbert space \mathcal{H} , the asymptotic secret-key-agreement rate satisfies the following inequality:

$$\max_L K_D(\rho_{[L]}) \leq K_D(\rho_{[M]}) \leq \min_N K_D(\rho_{[N]}), \quad (86)$$

where the state $\rho_{[L]}$ is obtained from the state $\rho_{[M]}$ by splitting its subsystems so that $L \geq \log \dim(\mathcal{H})$. Analogously, the state $\rho_{[N]}$ is obtained via any merging of subsystems of $\rho_{[M]}$, such that $\rho_{[N]}$ has at least two subsystems.

Hence, in the particular case of the Φ_3^W state, we can also skip minimization with respect to i, j since the state is symmetric. Using properties of entanglement measures [131–133], we have

$$K_D(\Phi_3^W) \leq K_D(\Phi_{2+1}^W) \leq E_r^\infty(\Phi_{2+1}^W) \quad (87)$$

$$= h_2\left(\frac{1}{3}\right) \approx 0.9183 \text{ bit}, \quad (88)$$

where $h_2(x)$ is the binary entropy function.

The asymptotic key rate and bounds on it are usually noninteger real numbers. In the one-shot regime, expressing these quantities in a similar manner, instead of integers obtained with floor or ceiling functions, is no less meaningful because the amount of secret key and the value of bounds are functions of privacy test parameter ε , which can vary, yielding, in general, different values of these quantities. Therefore, dependence of the scenario on the privacy parameter ε is interesting on its own. See Appendix H and Ref. [134].

Remark 6: It is natural that the analogies of Proposition 3 and Corollary 7 hold for the multiplex quantum channel \mathcal{N} . The upper bound on the M -partite multiplex quantum channel takes the form of the $M-1$ -partite multiplex channel, where the new party's type is determined according to the following rule: If the two parties are of the same

type (say, B), then the new type is the same (B in that case). If the types are different, then the new type always becomes A because, e.g., when B and C are merged, they have the ability to both read and write.

IX. DISCUSSION

We have provided universal limits on the rates at which one can distribute the conference key over a quantum network described by a multiplex quantum channel. We have shown that multipartite private states are necessarily genuine multipartite entangled. As a consequence, it is not possible to distill multipartite private states from tensor-stable biseparable states. We have obtained an upper bound on the single-shot, classical preprocessing and postprocessing assisted secret-key-agreement capacity. The bound is in terms of the hypothesis-testing divergence with respect to biseparable states of the output state of the multiplex channel, maximized over all fully separable input states. We have further provided strong-converse bounds on the LOCC-assisted private capacity of multiplex channels that are in terms of the max-relative entropy of entanglement as well as the regularized relative entropy of entanglement. In the case of tele-covariant multiplex channels, we have also obtained bounds in terms of the relative entropy of entanglement of the resource state. We have shown the versatility of our bounds by applying it to several communication scenarios, including measurement-device-independent QKD and conference key agreement as well as quantum key repeaters. In addition to our upper bounds, we have also provided lower bounds on asymptotic conference key rates, which are asymptotically achievable in Devetak-Winter-like protocols. We also derived an upper bound on the secret key that can be distilled from finite copies of multipartite states via LOCC, and we showed some numerical examples. The task of distillation of Φ_3^{GHZ} from Φ_3^{W} was extensively studied in the literature [81,122,123]. Here, we initiate the study on the distillation of the key rather than Φ_3^{GHZ} distillation from the Φ_3^{W} state. This is the rate of the distillation of “twisted” Φ_3^{GHZ} being private states—a class to which Φ_3^{GHZ} belongs. It would be interesting to find if the distillation of the key from Φ_3^{W} is just equivalent to the distillation of Φ_3^{GHZ} (see recent result on this topic [127]).

Distillation of the secret key allows trusted parties to access private random bits. Our lower bound on an asymptotic LOCC-assisted secret-key-agreement capacity over a multiplex channel also provides an asymptotic achievable rate of private random bits for trusted parties over a multiplex channel with classical preprocessing and postprocessing.

Our work also provides frameworks for the resource theories of multipartite entanglement for quantum multipartite channels (analogous to bipartite channels as discussed in Refs. [51,53,97,98]). In this context, it is natural to extend the results of Ref. [135], where the so-called layered QKD is considered, to the noisy case of multipartite private states. It would be interesting to systematically

consider other frameworks in the resource theory of multipartite entanglement. An important future direction for application purposes is to identify new information processing tasks and determine bounds on the rate regions of classical and quantum communication protocols over a multiplex channel (e.g., see Refs. [53,136–142]).

ACKNOWLEDGMENTS

S. D. is grateful to Jonathan P. Dowling (3 April 1955–5 June 2020) for insightful discussions. The authors thank Koji Azuma, Nicolas Cerf, Marcus Huber, Liang Jiang, Sumeet Khatri, Glaucia Murta, Mark M. Wilde, and Paweł Żyliński for valuable discussions. S. D. acknowledges individual fellowships at Université Libre de Bruxelles; this project received funding from the European Union’s Horizon 2020 research and innovation program under the Marie Skłodowska-Curie Grant Agreement No. 801505. S. B. acknowledges funding from the European Union’s Horizon 2020 research and innovation program, Grant Agreement No. 820466 (project CiViQ), the postdoctoral fellowships program Beatriu de Pinós, funded by the Secretary of Universities and Research (Government of Catalonia), and by the Horizon 2020 program of research and innovation of the European Union under the Marie Skłodowska-Curie Grant Agreement No. 801370 (2019 BP 00097), as well as from the Government of Spain (FIS2020-TRANQI and Severo Ochoa CEX2019-000910-S), Fundació Cellex, Fundació Mir-Puig, Generalitat de Catalunya (CERCA, AGAUR SGR 1381, and Quantum-CAT). K. H. and M. W. acknowledge support from the grant Sonata Bis 5 (Grant No. 2015/18/E/ST2/00327) from the National Science Center. We acknowledge partial support by the Foundation for Polish Science (IRAP project, ICTQT, Contract No. MAB/2018/5, co-financed by the EU within Smart Growth Operational Programme). The “International Centre for Theory of Quantum Technologies” project (Contract No. MAB/2018/5) is carried out within the International Research Agendas Programme of the Foundation for Polish Science co-financed by the European Union from the funds of the Smart Growth Operational Programme, axis IV: Increasing the research potential (Measure 4.3).

APPENDIX A: GENERALIZED DIVERGENCES AND THEIR PROPERTIES

Any generalized divergence $\mathbf{D}(\cdot||\cdot)$ satisfies the following two properties for an isometry U and a state τ [63]:

$$\mathbf{D}(\rho||\sigma) = \mathbf{D}(U\rho U^\dagger||U\sigma U^\dagger), \quad (\text{A1})$$

$$\mathbf{D}(\rho||\sigma) = \mathbf{D}(\rho \otimes \tau||\sigma \otimes \tau). \quad (\text{A2})$$

The sandwiched Rényi relative entropy obeys the following “monotonicity in α ” inequality [64]:

$$\tilde{D}_\alpha(\rho\|\sigma) \leq \tilde{D}_\beta(\rho\|\sigma) \text{ if } \alpha \leq \beta, \text{ for } \alpha, \beta \in (0, 1) \cup (1, \infty). \quad (\text{A3})$$

The following inequality states that the sandwiched Rényi relative entropy $\tilde{D}_\alpha(\rho\|\sigma)$ between states ρ, σ is a particular generalized divergence for certain values of α [143,144]. For a quantum channel \mathcal{N} ,

$$\tilde{D}_\alpha(\rho\|\sigma) \geq \tilde{D}_\alpha(\mathcal{N}(\rho)\|\mathcal{N}(\sigma)), \quad \forall \alpha \in [1/2, 1) \cup (1, \infty). \quad (\text{A4})$$

In the limit $\alpha \rightarrow 1$, the sandwiched Rényi relative entropy $\tilde{D}_\alpha(\rho\|\sigma)$ between quantum states ρ, σ converges to the quantum relative entropy [63,64]:

$$\lim_{\alpha \rightarrow 1} \tilde{D}_\alpha(\rho\|\sigma) = D(\rho\|\sigma), \quad (\text{A5})$$

and the quantum relative entropy [92] between states is

$$D(\rho\|\sigma) := \text{Tr}[\rho \log_2(\rho - \sigma)] \quad (\text{A6})$$

for $\text{supp}(\rho) \subseteq \text{supp}(\sigma)$ and otherwise it is ∞ .

In the limit $\alpha \rightarrow 1/2$, the sandwiched Rényi relative entropy $\tilde{D}_\alpha(\rho\|\sigma)$ converges to $-\log_2 F(\rho, \sigma)$, where $F(\rho, \sigma)$ is the fidelity between ρ, σ defined as

$$F(\rho, \sigma) := \left[\text{Tr} \left[\sqrt{\sqrt{\sigma} \rho \sqrt{\sigma}} \right] \right]^2. \quad (\text{A7})$$

The following inequality relates $D_h^\varepsilon(\rho\|\sigma)$ to $\tilde{D}_\alpha(\rho\|\sigma)$ for density operators ρ, σ , $\alpha \in (1, \infty)$ and $\varepsilon \in (0, 1)$ (see Refs. [145–147] and Lemma 5 in Ref. [148]):

$$D_h^\varepsilon(\rho\|\sigma) \leq \tilde{D}_\alpha(\rho\|\sigma) + \frac{\alpha}{\alpha-1} \log \left(\frac{1}{1-\varepsilon} \right). \quad (\text{A8})$$

The following inequality also holds [94]:

$$D_h^\varepsilon(\rho\|\sigma) \leq \frac{1}{1-\varepsilon} (D(\rho\|\sigma) + h_2(\varepsilon)), \quad (\text{A9})$$

where $h_2(\varepsilon) := -\varepsilon \log_2 \varepsilon - (1-\varepsilon) \log_2 (1-\varepsilon)$ is the binary entropy function.

In a specific case, ε -hypothesis-testing relative entropy can be calculated exactly.

Lemma 3: If ρ is a pure state and it is one of the eigenvectors of σ , i.e., there exists decomposition $\sigma = p_0 \rho + \sum_{i=1} p_i \gamma_i^\perp$, with $\sum_{i=0} p_i = 1$, $0 \leq p_i \leq 1$, $p_0 \neq 0$ and states γ_i^\perp orthogonal to ρ , then for any $\varepsilon \in [0, 1]$,

$$D_h^\varepsilon(\rho\|\sigma) = -\log_2 \text{Tr}[\Omega \sigma], \quad (\text{A10})$$

with $\Omega = (1-\varepsilon)\rho$.

APPENDIX B: MULTIPLEX QUANTUM CHANNELS

All network channels that are possible in a communication setting are special cases of multiplex quantum channels $\mathcal{N}_{\vec{A}' \vec{B} \rightarrow \vec{A} \vec{C}}$ (see Fig. 1):

- (1) Point-to-point quantum channel: This is a quantum channel of the form $\mathcal{N}_{B_b \rightarrow C_c}$ with a single sender and a single receiver. When a multiplex quantum channel has the form $\mathcal{N}_{B_b \rightarrow C_c}$ then $\mathcal{A} = \emptyset$ and $|\mathcal{B}| = 1 = |\mathcal{C}|$. This is arguably the simplest form of a communication (network) channel as it involves only two parties with one party sending input to the channel and the other receiving the output from the channel.
- (2) Bidirectional quantum channel: This is a multiplex quantum channel of the form $\mathcal{N}_{A_1 A_2 \rightarrow A_1 A_2}$ with two parties who are both senders and receivers, i.e., $|\mathcal{A}| = 2$ and $B = \emptyset = C$ (cf. Refs. [53,101]).
- (3) Quantum interference channel: This is a bipartite quantum channel of the form $\mathcal{N}_{B_1 B_2 \rightarrow C_1 C_2}$ with two senders and two receivers (cf. Ref. [149]). We may also call $\mathcal{N}_{\vec{B} \rightarrow \vec{C}}$, with an equal number of senders and receivers, as the quantum interference channel.
- (4) Broadcast quantum channel: This is a multipartite quantum channel of the form $\mathcal{N}_{B_b \rightarrow \vec{C}}$ with a single sender and multiple receivers (cf. Refs. [150,151]). We may also call $\mathcal{N}_{\vec{B} \rightarrow \vec{C}}$ as a broadcast channel if the number of senders is less than the number of receivers.
- (5) Multiple access quantum channel: This is a multipartite quantum channel of the form $\mathcal{N}_{\vec{B} \rightarrow C_c}$ with multiple senders and a single receiver (cf. Ref. [152]). We may also call $\mathcal{N}_{\vec{B} \rightarrow \vec{C}}$ as a multiple access channel if the number of senders is more than the number of receivers.
- (6) Physical box: Any physical box with quantum or classical inputs and quantum or classical outputs.
- (7) Network quantum channels of types $\mathcal{N}_{\vec{A}' \vec{B} \rightarrow \vec{A} \vec{C}}$ and $\mathcal{N}_{\vec{A}' \vec{B} \rightarrow \vec{A}}$.

If inputs and outputs to a multiplex channel are classical systems and underlying processes are governed by classical physics, then the channel is called a classical multiplex channel (see Ref. [137] for examples of such network channels). If inputs and outputs to the channel are quantum and classical systems, respectively, then the channel is called a quantum-to-classical channel. If inputs and outputs to the channel are classical and quantum systems, respectively, then the channel is called a classical to a quantum channel.

APPENDIX C: PRIVACY TEST

Recall the definition of the twisting operation

$$U_{KS}^{\text{tw}} = \sum_{i_1, \dots, i_M=0}^{K-1} |i_1 \dots i_M\rangle \langle i_1 \dots i_M|_{\vec{K}} \otimes U_S^{(i_1 \dots i_M)} \quad (\text{C1})$$

and a privacy test as

$$\Pi_{KS}^{\gamma,K} = U_{KS}^{\text{tw}}(\Phi_{\bar{K}}^{\text{GHZ}} \otimes \mathbb{1}_{\bar{S}})U_{KS}^{\text{tw}\dagger} \quad (\text{C2})$$

$$= \frac{1}{K} \sum_{i,k=0}^{K-1} (|i\rangle\langle k|)_{\bar{K}}^{\otimes M} \otimes U_{\bar{S}}^{(i^M)} U_{\bar{S}}^{(k^M)\dagger}, \quad (\text{C3})$$

where we have defined the notation $i^M := \underbrace{i\dots i}_{M\text{times}}$. We now

provide the proof of Theorem 1:

Proof of Theorem 1.—We begin by showing the bound for pure biseparable states $|\varphi\rangle_{\overline{KS}}$. For such a state, there exists a bipartition of the parties, defined by nonempty index sets $I \subset \{1, \dots, M\}$ and $J = \{1, \dots, M\} \setminus I$, such that the state is a product with respect to that bipartition. Namely, $|\varphi\rangle_{\overline{KS}} = |\tilde{\varphi}\rangle_{S_I K_I} \otimes |\tilde{\varphi}\rangle_{S_J K_J}$, where we have

defined $\mathcal{H}_{S_I K_I} = \bigotimes_{i \in I} \mathcal{H}_{S_i K_i}$ and $\mathcal{H}_{S_J K_J} = \bigotimes_{j \in J} \mathcal{H}_{S_j K_j}$. Let us also define $m := |I|$ and $n := |J|$ and note that $M = m + n$. We can expand

$$|\tilde{\varphi}\rangle_{S_I K_I} = \sum_{i_1, \dots, i_m=0}^{K-1} \tilde{\alpha}_{i_1 \dots i_m} |i_1 \dots i_m\rangle_{K_I} \otimes |\tilde{\varphi}_{i_1 \dots i_m}\rangle_{S_I}, \quad (\text{C4})$$

$$|\tilde{\varphi}\rangle_{S_J K_J} = \sum_{j_1, \dots, j_n=0}^{K-1} \tilde{\alpha}_{j_1 \dots j_n} |j_1 \dots j_n\rangle_{K_J} \otimes |\tilde{\varphi}_{j_1 \dots j_n}\rangle_{S_J}. \quad (\text{C5})$$

Here, $\tilde{\alpha}_{i_1 \dots i_m} \in \mathbb{C}$ such that $\sum_{i_1, \dots, i_m=0}^{K-1} |\tilde{\alpha}_{i_1 \dots i_m}|^2 = 1$ and $\tilde{\alpha}_{j_1 \dots j_n} \in \mathbb{C}$ such that $\sum_{j_1, \dots, j_n=0}^{K-1} |\tilde{\alpha}_{j_1 \dots j_n}|^2 = 1$. Furthermore, it holds that

$$\text{Tr}[\Pi_{KS}^{\gamma,K} \varphi_{\overline{KS}}] = \text{Tr} \left[\left(\frac{1}{K} \sum_{i,k=0}^{K-1} (|i\rangle\langle k|)_{\bar{K}}^{\otimes M} \otimes U_{\bar{S}}^{(i^M)} U_{\bar{S}}^{(k^M)\dagger} \right) \tilde{\varphi}_{K_I S_I} \otimes \tilde{\varphi}_{K_J S_J} \right] \quad (\text{C6})$$

$$= \frac{1}{K} \sum_{i,k=0}^{K-1} \tilde{\alpha}_{i^m} \tilde{\alpha}_{i^n} (\tilde{\alpha}_{k^m})^* (\tilde{\alpha}_{k^n})^* \text{Tr}[U^{(i^M)\dagger} |\tilde{\varphi}_{i^m}\rangle \langle \tilde{\varphi}_{k^m}|_{S_I} \otimes |\tilde{\varphi}_{i^n}\rangle \langle \tilde{\varphi}_{k^n}|_{S_J} U^{(k^M)}] \quad (\text{C7})$$

$$= \frac{1}{K} \sum_{i,k=0}^{K-1} \tilde{\alpha}_{i^m} \tilde{\alpha}_{i^n} (\tilde{\alpha}_{k^m})^* (\tilde{\alpha}_{k^n})^* \langle \zeta_k | \zeta_i \rangle, \quad (\text{C8})$$

where we have defined the state

$$|\zeta_i\rangle_{\bar{S}} := U^{(i^M)\dagger} |\tilde{\varphi}_{i^m}\rangle_{S_I} \otimes |\tilde{\varphi}_{i^n}\rangle_{S_J}. \quad (\text{C9})$$

We note that Eq. (C8) is a probability; in particular, it is real and non-negative. Hence, it holds that

$$\frac{1}{K} \sum_{i,k=0}^{K-1} \tilde{\alpha}_{i^m} \tilde{\alpha}_{i^n} (\tilde{\alpha}_{k^m})^* (\tilde{\alpha}_{k^n})^* \langle \zeta_k | \zeta_i \rangle \quad (\text{C10})$$

$$= \left| \frac{1}{K} \sum_{i,k=0}^{K-1} \tilde{\alpha}_{i^m} \tilde{\alpha}_{i^n} (\tilde{\alpha}_{k^m})^* (\tilde{\alpha}_{k^n})^* \langle \zeta_k | \zeta_i \rangle \right| \quad (\text{C11})$$

$$\leq \frac{1}{K} \sum_{i,k=0}^{K-1} |\tilde{\alpha}_{i^m}| |\tilde{\alpha}_{i^n}| |\tilde{\alpha}_{k^m}| |\tilde{\alpha}_{k^n}| |\langle \zeta_k | \zeta_i \rangle|, \quad (\text{C12})$$

where in the first inequality, we have used the subadditivity and multiplicity of the absolute value of complex numbers. We note that for all i, k in the sum, $|\langle \zeta_k | \zeta_i \rangle| \leq 1$. Let us define $p_i = |\tilde{\alpha}_{i^m}|^2$ and note that $p_i \geq 0$ and $\sum_{i=0}^{K-1} p_i \leq 1$. Let us also define $q_i = |\tilde{\alpha}_{i^n}|^2$ and note that $q_i \geq 0$ and

$\sum_{i=0}^{K-1} q_i \leq 1$. Hence, there exist respective probability distributions $\{\hat{p}_i\}$ and $\{\hat{q}_i\}$ over $\{0, \dots, K-1\}$ such that $p_i \leq \hat{p}_i$ and $q_i \leq \hat{q}_i$ for all $i = 0, \dots, K-1$. We then obtain

$$\frac{1}{K} \sum_{i,k=0}^{K-1} |\tilde{\alpha}_{i^m}| |\tilde{\alpha}_{i^n}| |\tilde{\alpha}_{k^m}| |\tilde{\alpha}_{k^n}| |\langle \zeta_k | \zeta_i \rangle| \quad (\text{C13})$$

$$\leq \frac{1}{K} \sum_{i,k=0}^{K-1} \sqrt{p_i q_i p_k q_k} = \frac{1}{K} \left[\sum_{i=0}^{K-1} \sqrt{p_i q_i} \right]^2 \quad (\text{C14})$$

$$\leq \frac{1}{K} \left[\sum_{i=0}^{K-1} \sqrt{\hat{p}_i \hat{q}_i} \right]^2 \leq \frac{1}{K}, \quad (\text{C15})$$

where we have used the fact that the classical fidelity between two probability distributions is upper bounded by 1. This establishes the theorem for pure biseparable states with respect to arbitrary bipartitions. Noting that every mixed biseparable state $\sigma_{\overline{KS}} \in \text{BS}(\overrightarrow{KS})$ can be expressed as a convex sum of pure biseparable states finishes the proof. ■

APPENDIX D: UPPER BOUNDS ON THE CKA RATES OF MULTIPLEX CHANNELS

1. Proof of Theorem 2

Proof.—Let us consider any cppp-assisted protocol that achieves a rate $\hat{P}_{\text{cppp}}^{\mathcal{N}} \equiv \hat{P}$. Let $\rho^{(1)} \in \text{FS}(\overrightarrow{LA'} : \overrightarrow{RB} : \overrightarrow{P})$ be a fully separable state generated by the first use of LOCC among all spatially separated allies. Let

$$\tau_{\overrightarrow{LARPC}}^{(1)} := \mathcal{N}(\rho_{\overrightarrow{LARBP}}^{(1)}). \quad (\text{D1})$$

We note that $\tau^{(1)}$ is a separable state with respect to bipartition $\overrightarrow{LAR} : \overrightarrow{C} : \overrightarrow{P}$. The action of the decoder channel $\mathcal{D} := \mathcal{L}_{\overrightarrow{LARPC} \rightarrow \overrightarrow{SK}}^{(2)}$ on $\tau^{(1)}$ yields the state

$$\omega_{\overrightarrow{SK}} := \mathcal{L}^{(2)}(\tau_{\overrightarrow{LARPC}}^{(1)}). \quad (\text{D2})$$

By assumption, we have that

$$F(\gamma_{\overrightarrow{SK}}, \omega_{\overrightarrow{SK}}) \geq 1 - \varepsilon, \quad (\text{D3})$$

for some (M -partite) private state γ , which implies that there exists a projector $\Pi_{\overrightarrow{SK}}^{\gamma}$ corresponding to a γ -privacy test such that (see Proposition 1)

$$\text{Tr}[\Pi_{\overrightarrow{SK}}^{\gamma} \omega_{\overrightarrow{SK}}] \geq 1 - \varepsilon. \quad (\text{D4})$$

From Theorem 1,

$$\text{Tr}[\Pi_{\overrightarrow{SK}}^{\gamma} \sigma'_{\overrightarrow{SK}}] \leq \frac{1}{K} = 2^{-\hat{P}}, \quad (\text{D5})$$

for any $\sigma' \in \text{BS}(\overrightarrow{SK})$.

Let us suppose a state $\sigma_{\overrightarrow{LARPC}} \in \text{BS}(\overrightarrow{LA} : \overrightarrow{R} : \overrightarrow{PC})$ of the form $\sigma_{\overrightarrow{LARPC}} = \sigma_{\overrightarrow{LAR} : \overrightarrow{C}} \otimes \sigma_{\overrightarrow{P}}$, where $\sigma_{\overrightarrow{LAR} : \overrightarrow{C}}$ is arbitrary. It holds that $\sigma_{\overrightarrow{SK}} := \mathcal{L}^{(2)}(\sigma_{\overrightarrow{LARPC}}) \in \text{BS}(\overrightarrow{SK})$. Thus, the privacy test is feasible for $D_h^\varepsilon(\omega \| \sigma)$, and we find that

$$\hat{P} \leq D_h^\varepsilon(\omega_{\overrightarrow{SK}} \| \sigma_{\overrightarrow{SK}}) \quad (\text{D6})$$

$$\leq D_h^\varepsilon(\tau_{\overrightarrow{LARPC}}^{(1)} \| \sigma_{\overrightarrow{LARPC}}) \quad (\text{D7})$$

$$\leq \sup_{\psi \in \text{FS}(\overrightarrow{LA'} : \overrightarrow{RB} : \overrightarrow{P})} D_h^\varepsilon(\mathcal{N}(\psi_{\overrightarrow{LA'} : \overrightarrow{RB}}) \| \sigma_{\overrightarrow{LARPC}}) \quad (\text{D8})$$

$$= \sup_{\psi \in \text{FS}(\overrightarrow{LA'} : \overrightarrow{RB})} D_h^\varepsilon(\mathcal{N}(\psi_{\overrightarrow{LA'} : \overrightarrow{RB}}) \| \sigma_{\overrightarrow{LAR} : \overrightarrow{C}}). \quad (\text{D9})$$

The second inequality follows from the data-processing inequality. The third inequality follows from the quasiconvexity of D_h^ε . The equality follows from Eq. (A2) and a suitable choice of $\sigma_{\overrightarrow{P}}$ that always exists because, for any pure state $\psi \in \text{FS}(\overrightarrow{LA'} : \overrightarrow{RB} : \overrightarrow{P})$, the output state $\mathcal{N}(\psi)$ is separable with respect to the bipartition $\overrightarrow{LAR} : \overrightarrow{C} : \overrightarrow{P}$.

Since inequality (D9) also holds for an arbitrary $\sigma \in \text{BS}(\overrightarrow{LA} : \overrightarrow{R} : \overrightarrow{C})$, we can conclude that

$$\hat{P} \leq E_{h, \text{GE}}^\varepsilon(\mathcal{N}). \quad (\text{D10})$$

■

2. Proof of Theorem 3

Proof.—The following inequality holds for an (n, K, ε) LOCC-assisted secret-key-agreement protocol over a multiplex channel \mathcal{N} :

$$F(\omega_{\overrightarrow{SK}}, \gamma_{\overrightarrow{SK}}) \geq 1 - \varepsilon. \quad (\text{D11})$$

For any $\sigma_{\overrightarrow{SK}} \in \text{FS}(\overrightarrow{SK})$, we have the following bound due to inequality (D11) and Theorem 1:

$$\log_2 K \leq D_h^\varepsilon(\omega_{\overrightarrow{SK}} \| \sigma_{\overrightarrow{SK}}). \quad (\text{D12})$$

Employing inequality (A8) in the limit $\alpha \rightarrow +\infty$, we obtain

$$\log_2 K \leq D_h^\varepsilon(\omega_{\overrightarrow{SK}} \| \sigma_{\overrightarrow{SK}}) \quad (\text{D13})$$

$$\leq D_{\max}(\omega_{\overrightarrow{SK}} \| \sigma_{\overrightarrow{SK}}) + \log_2 \left(\frac{1}{1 - \varepsilon} \right). \quad (\text{D14})$$

The above inequality holds for arbitrary $\sigma \in \text{FS}(\overrightarrow{SK})$; therefore,

$$\log_2 K \leq E_{\max, E}(\overrightarrow{SK})_\omega + \log_2 \left(\frac{1}{1 - \varepsilon} \right), \quad (\text{D15})$$

where $E_{\max, E}(\overrightarrow{SK})_\omega$ is the max-relative entropy of entanglement of the state $\omega_{\overrightarrow{SK}}$.

The max-relative entropy of entanglement $E_{\max, E}$ of a state is monotonically nonincreasing under the action of LOCC channels, and it is zero for states that are fully separable. Using these facts, we get that

$$E_{\max,E}(\overrightarrow{SK})_{\omega} \leq E_{\max,E}(\overrightarrow{L^{(n)}A^{(n)}}:\overrightarrow{R^{(n)}}:\overrightarrow{P^{(n)}C^{(n)}})_{\tau_n} \quad (\text{D16})$$

$$= E_{\max,E}(\overrightarrow{L^{(n)}A^{(n)}}:\overrightarrow{R^{(n)}}:\overrightarrow{P^{(n)}C^{(n)}})_{\tau_n} - E_{\max,E}(\overrightarrow{L^{(1)}A^{(1)'}}:\overrightarrow{R^{(1)}B^{(1)'}}:\overrightarrow{P^{(1)}})_{\rho_1} \quad (\text{D17})$$

$$= E_{\max,E}(\overrightarrow{L^{(n)}A^{(n)}}:\overrightarrow{R^{(n)}}:\overrightarrow{P^{(n)}C^{(n)}})_{\tau_n} + \left[\sum_{i=2}^n E_{\max,E}(\overrightarrow{L^{(i)}A^{(i)'}}:\overrightarrow{R^{(i)}B^{(i)'}}:\overrightarrow{P^{(i)}})_{\rho_i} - \sum_{i=2}^n E_{\max,E}(\overrightarrow{L^{(i)}A^{(i)'}}:\overrightarrow{R^{(i)}B^{(i)'}}:\overrightarrow{P^{(i)}})_{\rho_i} \right] - E_{\max,E}(\overrightarrow{L^{(1)}A^{(1)'}}:\overrightarrow{R^{(1)}B^{(1)'}}:\overrightarrow{P^{(1)}})_{\rho_1} \quad (\text{D18})$$

$$\leq \sum_{i=1}^n [E_{\max,E}(\overrightarrow{L^{(i)}A^{(i)'}}:\overrightarrow{R^{(i)}}:\overrightarrow{P^{(i)}C^{(i)}})_{\tau_i} - E_{\max,E}(\overrightarrow{L^{(i)}A^{(i)'}}:\overrightarrow{R^{(i)}B^{(i)'}}:\overrightarrow{P^{(i)}})_{\rho_i}] \quad (\text{D19})$$

$$\leq nE_{\max,E}(\mathcal{N}). \quad (\text{D20})$$

The first equality follows because $E_{\max,E}(\overrightarrow{L^{(1)}A^{(1)'}}:\overrightarrow{R^{(1)}B^{(1)'}}:\overrightarrow{P^{(1)}})_{\rho_1} = 0$. The second inequality follows because $E_{\max,GE}$ is monotone under LOCC channels and $\rho_i = \mathcal{L}^i(\tau_{i-1})$ for all $i \in \{2, 3, \dots, n\}$. The final inequality follows from Lemma 1.

From inequalities (D15) and (D20), we conclude that

$$\log_2 K \leq nE_{\max,E}(\mathcal{N}) + \log_2 \left(\frac{1}{1-\varepsilon} \right). \quad (\text{D21})$$

3. Proof of Theorem 4

Proof.—For an (n, K, ε) LOCC-assisted secret-key-agreement protocol over a multiplex channel \mathcal{N} , such that $F(\omega_{SK}^{\rightarrow}, \gamma_{SK}^{\rightarrow}) \geq 1 - \varepsilon$, due to inequality (D11) and Theorem 1, it holds for any $\sigma_{SK}^{\rightarrow} \in \text{FS}(\overrightarrow{SK})$: that

$$\log_2 K \leq D_h^\varepsilon(\omega_{SK}^{\rightarrow} \| \sigma_{SK}^{\rightarrow}). \quad (\text{D22})$$

Using the fact that [94]

$$D_h^\varepsilon(\omega_{SK}^{\rightarrow} \| \sigma_{SK}^{\rightarrow}) \leq \frac{1}{1-\varepsilon} \left[D(\omega_{SK}^{\rightarrow} \| \sigma_{SK}^{\rightarrow}) + h(\varepsilon) \right], \quad (\text{D23})$$

where h is the binary entropy function, and that the bound (D22) holds for arbitrary $\sigma \in \text{FS}(\overrightarrow{SK})$, we obtain

$$\log_2 K \leq \frac{1}{1-\varepsilon} [E_E(\overrightarrow{SK})_{\omega} + h(\varepsilon)]. \quad (\text{D24})$$

As the relative entropy of entanglement of a state is monotonically nonincreasing under the action of LOCC channels and vanishes for states that are fully separable, we can repeat the argument in inequalities (D16)–(D20) and obtain

$$E_E(\overrightarrow{SK})_{\omega} \leq nE_E^p(\mathcal{N}) \leq nE_E^\infty(\mathcal{N}), \quad (\text{D25})$$

where the second inequality follows from Lemma 2. Taking the limits $\varepsilon \rightarrow 0$ and $n \rightarrow \infty$, we obtain

$$\hat{\mathcal{P}}_{\text{LOCC}}(\mathcal{N}) \leq E_E^\infty(\mathcal{N}), \quad (\text{D26})$$

showing the converse. As for the strong converse, we follow the argument used in Ref. [49]: From inequalities (D22) and (A8), we obtain

$$\log_2 K \leq \tilde{E}_{\alpha,E}(\overrightarrow{SK})_{\omega} + \frac{\alpha}{\alpha-1} \log_2 \left(\frac{1}{1-\varepsilon} \right), \quad (\text{D27})$$

where $\alpha \in (1, \infty)$ and $\tilde{E}_{\alpha,E}(\overrightarrow{SK})_{\omega}$ is the sandwiched Rényi relative entropy of entanglement of the state $\omega_{SK}^{\rightarrow}$. Rewriting inequality (D27), we obtain

$$\varepsilon \geq 1 - 2^{-n \frac{\alpha-1}{\alpha} \left(\frac{\log_2 K}{n} - \frac{1}{n} \tilde{E}_{\alpha,E}(\overrightarrow{SK})_{\omega} \right)}. \quad (\text{D28})$$

Assuming that the rate $\log_2 K/n$ exceeds $E_E^\infty(\mathcal{N})$, by inequality (D25), it will be larger than $(1/n)E_E(\overrightarrow{SK})_{\omega}$. Hence, there exists an $\alpha > 1$, such that $(\log_2 K/n) - (1/n)\tilde{E}_{\alpha,E}(\overrightarrow{SK})_{\omega} > 0$, and the error increases to 1 exponentially. ■

4. Proof of Theorem 5

Let $\mathcal{N}_{\vec{A} \vec{B} \rightarrow \vec{A} \vec{C}}$ be a multipartite quantum channel that is tele-covariant with respect to groups $\{\mathcal{G}_a\}_{a \in \mathcal{A}}$ and $\{\mathcal{G}_b\}_{b \in \mathcal{B}}$ as defined in Sec. V C. By definition, for all $a \in \mathcal{A}$ and $b \in \mathcal{B}$, we have

$$\frac{1}{G_a} \sum_{g_a} \mathcal{U}_{A_a''}^{g_a}(\Phi_{A_a'' L_a}^+) = \frac{\mathbb{1}_{A_a''}}{|A_a''|} \otimes \frac{\mathbb{1}_{L_a}}{|L_a|}, \quad (\text{D29})$$

$$\frac{1}{G_b} \sum_{g_b} \mathcal{U}_{B_b''}^{g_b}(\Phi_{B_b'' R_b}^+) = \frac{\mathbb{1}_{B_b''}}{|B_b''|} \otimes \frac{\mathbb{1}_{R_b}}{|R_b|}, \quad (\text{D30})$$

respectively, where $A'_a \simeq L_a$, $B'_b \simeq R_b$, and Φ^+ denotes an EPR state. Note that in order for each $\{U_{A'_a}^{g_a}\}$ and $\{U_{B'_b}^{g_b}\}$ to be one-designs, it is necessary that $|A'_a|^2 \leq G_a$ and $|B'_b|^2 \leq G_b$ [153].

For every $a \in \mathcal{A}$ and every $b \in \mathcal{B}$, we can now define $\{E_{A'_a L_a}^{g_a}\}$ and $\{E_{B'_b R_b}^{g_b}\}$, with respective elements defined as

$$E_{A'_a L_a}^{g_a} := \frac{|A'_a|^2}{G_a} U_{A'_a}^{g_a} \Phi_{A'_a L_a}^+ (U_{A'_a}^{g_a})^\dagger, \quad (\text{D31})$$

$$E_{B'_b R_b}^{g_b} := \frac{|B'_b|^2}{G_b} U_{B'_b}^{g_b} \Phi_{B'_b R_b}^+ (U_{B'_b}^{g_b})^\dagger, \quad (\text{D32})$$

where $A'_a \simeq A''_a$ and $B'_b \simeq B''_b$. It follows from the fact that $|A'_a|^2 \leq G_a$ and $|B'_b|^2 \leq G_b$ as well as Eqs. (D29) and (D30)

that $\{E_{A'_a L_a}^{g_a}\}_{g_a}$ and $\{E_{B'_b R_b}^{g_b}\}_{g_b}$ are valid POVMs for all $a \in \mathcal{A}$ and $b \in \mathcal{B}$.

The simulation of the channel $\mathcal{N}_{\vec{A}' \vec{B}' \rightarrow \vec{A} \vec{C}}$ via teleportation begins with a state $\rho_{\vec{A}' \vec{B}'}$ and a shared resource $\theta_{\vec{L} \vec{R} \vec{C}} = \mathcal{N}_{\vec{A}' \vec{B}' \rightarrow \vec{A} \vec{C}}(\Phi_{\vec{L} \vec{R} |\vec{A}' \vec{B}'})$. The desired outcome is for the receivers to receive the state $\mathcal{N}(\rho_{\vec{A}' \vec{B}'})$ and for the protocol to work independently of the input state $\rho_{\vec{A}' \vec{B}'}$. The first step is for senders \mathbf{A}_a and \mathbf{B}_b to locally perform the measurement $\{\bigotimes_{a \in \mathcal{A}} E_{A'_a L_a}^{g_a} \otimes \bigotimes_{b \in \mathcal{B}} E_{B'_b R_b}^{g_b}\}_{\vec{g}}$ and then send the outcomes \vec{g} to the receivers. Based on the outcomes \vec{g} , the receivers \mathbf{A}_a and \mathbf{C}_c then perform $W_{A_a}^{\vec{g}}$ and $W_{C_c}^{\vec{g}}$, respectively. The following analysis demonstrates that this protocol works by simplifying the form of the postmeasurement state:

$$\begin{aligned} & \left(\prod_{a \in \mathcal{A}} G_a \prod_{b \in \mathcal{B}} G_b \right) \text{Tr}_{\vec{A}' \vec{L} \vec{B}' \vec{R}} \left[\left(\bigotimes_{a \in \mathcal{A}} E_{A'_a L_a}^{g_a} \otimes \bigotimes_{b \in \mathcal{B}} E_{B'_b R_b}^{g_b} \right) \left(\rho_{\vec{A}' \vec{B}'} \otimes \theta_{\vec{L} \vec{R} \vec{C}} \right) \right] \\ &= \left(\prod_{a \in \mathcal{A}} |A'_a|^2 \prod_{b \in \mathcal{B}} |B'_b|^2 \right) \text{Tr}_{\vec{A}' \vec{L} \vec{B}' \vec{R}} \left\{ \left[\bigotimes_{a \in \mathcal{A}} U_{A'_a}^{g_a} \Phi_{A'_a L_a}^+ U_{A'_a}^{g_a \dagger} \otimes \bigotimes_{b \in \mathcal{B}} U_{B'_b}^{g_b} \Phi_{B'_b R_b}^+ U_{B'_b}^{g_b \dagger} \right] \left(\rho_{\vec{A}' \vec{B}'} \otimes \theta_{\vec{L} \vec{R} \vec{C}} \right) \right\} \end{aligned} \quad (\text{D33})$$

$$= \left(\prod_{a \in \mathcal{A}} |A'_a|^2 \prod_{b \in \mathcal{B}} |B'_b|^2 \right) \langle \Phi^+ |_{\vec{A}' \vec{B}' | \vec{L} \vec{R}} \left(\bigotimes_{a \in \mathcal{A}} U_{A'_a}^{g_a} \otimes \bigotimes_{b \in \mathcal{B}} U_{B'_b}^{g_b} \right)^\dagger \rho_{\vec{A}' \vec{B}'} \otimes \theta_{\vec{L} \vec{R} \vec{C}} \left(\bigotimes_{a \in \mathcal{A}} U_{A'_a}^{g_a} \otimes \bigotimes_{b \in \mathcal{B}} U_{B'_b}^{g_b} \right) | \Phi^+ \rangle_{\vec{A}' \vec{B}' | \vec{L} \vec{R}} \quad (\text{D34})$$

$$= \left(\prod_{a \in \mathcal{A}} |A'_a|^2 \prod_{b \in \mathcal{B}} |B'_b|^2 \right) \langle \Phi^+ |_{\vec{A}' \vec{B}' | \vec{L} \vec{R}} \left(\bigotimes_{a \in \mathcal{A}} U_{A'_a}^{g_a} \otimes \bigotimes_{b \in \mathcal{B}} U_{B'_b}^{g_b} \right)^\dagger \rho_{\vec{A}' \vec{B}'} \left(\bigotimes_{a \in \mathcal{A}} U_{A'_a}^{g_a} \otimes \bigotimes_{b \in \mathcal{B}} U_{B'_b}^{g_b} \right) \otimes \theta_{\vec{L} \vec{R} \vec{C}} | \Phi^+ \rangle_{\vec{A}' \vec{B}' | \vec{L} \vec{R}} \quad (\text{D35})$$

$$= \left(\prod_{a \in \mathcal{A}} |A'_a|^2 \prod_{b \in \mathcal{B}} |B'_b|^2 \right) \langle \Phi^+ |_{\vec{A}' \vec{B}' | \vec{L} \vec{R}} \left[\left(\bigotimes_{a \in \mathcal{A}} U_{A'_a}^{g_a} \otimes \bigotimes_{b \in \mathcal{B}} U_{B'_b}^{g_b} \right)^\dagger \rho_{\vec{L} \vec{R}} \left(\bigotimes_{a \in \mathcal{A}} U_{A'_a}^{g_a} \otimes \bigotimes_{b \in \mathcal{B}} U_{B'_b}^{g_b} \right) \right]^* \theta_{\vec{L} \vec{R} \vec{C}} | \Phi^+ \rangle_{\vec{A}' \vec{B}' | \vec{L} \vec{R}}. \quad (\text{D36})$$

The first three equalities follow by substitution and some rewriting. The fourth equality follows from the fact that

$$\langle \Phi |_{A'A} M_{A'} = \langle \Phi |_{A'A} M_{A'}^* \quad (\text{D37})$$

for any operator M , where $*$ denotes the complex conjugate, taken with respect to the basis in which $|\Phi\rangle_{A'A}$ is defined. Continuing, we have that

$$\text{Eq. (D36)} = \left(\prod_{a \in \mathcal{A}} |A'_a| \prod_{b \in \mathcal{B}} |B'_b| \right) \text{Tr}_{\vec{L} \vec{R}} \left\{ \left[\left(\bigotimes_{a \in \mathcal{A}} U_{A'_a}^{g_a} \otimes \bigotimes_{b \in \mathcal{B}} U_{B'_b}^{g_b} \right)^\dagger \rho_{\vec{L} \vec{R}} \left(\bigotimes_{a \in \mathcal{A}} U_{A'_a}^{g_a} \otimes \bigotimes_{b \in \mathcal{B}} U_{B'_b}^{g_b} \right) \right]^* \mathcal{N}_{\vec{A}' \vec{B}' \rightarrow \vec{A} \vec{C}} \left(\Phi_{\vec{L} \vec{R} | \vec{A}' \vec{B}'}^+ \right) \right\} \quad (\text{D38})$$

$$= \left(\prod_{a \in \mathcal{A}} |A'_a| \prod_{b \in \mathcal{B}} |B'_b| \right) \text{Tr}_{\vec{L} \vec{R}} \left[\mathcal{N}_{\vec{A}' \vec{B}' \rightarrow \vec{A} \vec{C}} \left(\left(\bigotimes_{a \in \mathcal{A}} U_{A'_a}^{g_a} \otimes \bigotimes_{b \in \mathcal{B}} U_{B'_b}^{g_b} \right)^\dagger \rho_{\vec{A}' \vec{B}'} \left(\bigotimes_{a \in \mathcal{A}} U_{A'_a}^{g_a} \otimes \bigotimes_{b \in \mathcal{B}} U_{B'_b}^{g_b} \right) \Phi_{\vec{L} \vec{R} | \vec{A}' \vec{B}'}^+ \right) \right] \quad (\text{D39})$$

$$= \mathcal{N}_{\vec{A}' \vec{B}' \rightarrow \vec{A} \vec{C}} \left(\left(\bigotimes_{a \in \mathcal{A}} U_{A'_a}^{g_a} \otimes \bigotimes_{b \in \mathcal{B}} U_{B'_b}^{g_b} \right)^\dagger \rho_{\vec{A}' \vec{B}'} \left(\bigotimes_{a \in \mathcal{A}} U_{A'_a}^{g_a} \otimes \bigotimes_{b \in \mathcal{B}} U_{B'_b}^{g_b} \right) \right) \quad (\text{D40})$$

$$= \left(\bigotimes_{a \in \mathcal{A}} W_{A_a}^{\vec{g}} \otimes \bigotimes_{c \in \mathcal{C}} W_{C_c}^{\vec{g}} \right)^\dagger \mathcal{N}_{\vec{A}' \vec{B}' \rightarrow \vec{A} \vec{C}} \left(\rho_{\vec{A}' \vec{B}'} \right) \left(\bigotimes_{a \in \mathcal{A}} W_{A_a}^{\vec{g}} \otimes \bigotimes_{c \in \mathcal{C}} W_{C_c}^{\vec{g}} \right). \quad (\text{D41})$$

The first equality follows because $|A| \langle \Phi |_{A'A} (\mathbb{1}_{A'} \otimes M_{AB}) | \Phi \rangle_{A'A} = \text{Tr}_A \{ M_{AB} \}$ for any operator M_{AB} . The second equality follows by applying the conjugate transpose of Eq. (D37). The final equality follows from the covariance property of the channel.

Thus, if the receivers finally perform the unitaries $\bigotimes_{a \in A} W_{A_a}^{\vec{g}} \otimes \bigotimes_{c \in C} W_{C_c}^{\vec{g}}$ upon receiving \vec{g} via a classical channel from the senders, then the output of the protocol is $\mathcal{N}_{\vec{A} \vec{B} \rightarrow \vec{A} \vec{C}}(\rho_{\vec{A} \vec{B}})$, so this protocol simulates the action of the multipartite channel \mathcal{N} on the state ρ . ■

5. Proof of Theorem 6

Before proving Theorem 6, we need the following lemma, which generalizes Lemma 7 in Ref. [62]:

Lemma 4: Let $\mathcal{T} = \{U^{tw\dagger} \rho_{SK} U^{tw} : \rho_{SK} \in \text{BS}(\cdot : SK :)\}$ be the set of twisted biseparable states. Then, for any $\sigma_{SK} \in \mathcal{T}$, it holds that

$$D(\Phi_{\vec{K}} || \sigma_{\vec{K}}) \geq \log K. \quad (\text{D42})$$

Proof.—Let $\sigma_{SK} \in \mathcal{T}$, i.e., $\sigma_{SK} = U^{tw\dagger} \rho_{SK} U^{tw}$ for some twisting unitary U^{tw} and biseparable ρ_{SK} . Here, U^{tw} defines a privacy test $\Pi_{SK}^{\gamma} = U^{tw}(\Phi_{\vec{K}} \otimes \mathbb{1}_{\vec{S}})U^{tw\dagger}$. By Theorem 6, it then holds that

$$\text{Tr}[\Phi_{\vec{K}} \sigma_{\vec{K}}] = \text{Tr}[\Pi_{SK}^{\gamma} \rho_{SK}] \leq \frac{1}{K}. \quad (\text{D43})$$

By the concavity of the logarithm, it then holds that

$$D(\Phi_{\vec{K}} || \sigma_{\vec{K}}) = -S(\Phi_{\vec{K}}) - \text{Tr}[\Phi_{\vec{K}} \log \sigma_{\vec{K}}] \quad (\text{D44})$$

$$\geq -\log \text{Tr}[\Phi_{\vec{K}} \sigma_{\vec{K}}] \quad (\text{D45})$$

$$\geq \log K, \quad (\text{D46})$$

finishing the proof. ■

Now, we can follow Ref. [48] to prove Theorem 6:

Proof of Theorem 6.—Let $\epsilon > 0$ and $n \in \mathbb{N}$. We begin by noting that in the case of teleportation-simulable multiplex channels, LOCC assistance does not enhance secret-key-agreement capacity, and the original protocol can be reduced to a cppo-assisted secret-key-agreement protocol [48]. Namely, in every round $1 \leq i \leq n$, it holds that

$$\rho_i = \mathcal{L}^i(\tau_i) = \mathcal{L}^i(\mathcal{N}_{A^{(i-1)} B^{(i-1)} \rightarrow A^{(i-1)} C^{(i-1)}}(\rho_{i-1})) \quad (\text{D47})$$

$$= \mathcal{L}^i(\mathcal{T}_{A^{(i-1)} L A B^{(i-1)} R \vec{C} \rightarrow A^{(i-1)} C^{(i-1)}}(\theta_{L A R \vec{C}} \otimes \rho_{i-1})), \quad (\text{D48})$$

where \mathcal{L}^i and \mathcal{T} are LOCC. As the initial state ρ_0 is assumed to be fully separable, we find that the final state

$\omega_{SK} = \rho_n$ of an adaptive LOCC CKA protocol, involving n uses of the teleportation-simulable multiplex channel $\mathcal{N}_{\vec{A} \vec{B} \rightarrow \vec{A} \vec{C}}$, can be expressed as

$$\omega_{SK} = \mathcal{L}_{L^n A^n R^n \vec{C} \rightarrow SK}(\theta_{L A R \vec{C}}^{\otimes n}), \quad (\text{D49})$$

where \mathcal{L} is an LOCC operation with respect to the partition $:\overline{L^n A^n} : \overline{R^n} : \overline{C^n} :$. By assumption, it holds that $\|\omega_{SK} - \gamma_{KS}\|_1 \leq \epsilon$ for some m -partite private state $\gamma_{KS} = U^{tw}(\Phi_{\vec{K}} \otimes \tau_{\vec{S}})U^{tw\dagger}$, where m is the number of parties. Let $\tilde{\sigma}_{L^n A^n R^n C^n} \in \text{BS}(\cdot : \overline{L^n A^n} : \overline{R^n} : \overline{C^n} :)$. Following the proof of Theorem 9 in Ref. [62], we obtain

$$D(\theta_{L A R \vec{C}}^{\otimes n} || \tilde{\sigma}) \geq D(\omega_{SK} || \mathcal{L}(\tilde{\sigma}_{L^n A^n R^n C^n})) \quad (\text{D50})$$

$$= D(U^{tw\dagger} \omega_{SK} U^{tw} || U^{tw\dagger} \mathcal{L}(\tilde{\sigma}_{L^n A^n R^n C^n}) U^{tw}) \quad (\text{D51})$$

$$\geq \inf_{\sigma_{SK} \in \mathcal{T}} D(\text{Tr}_{\vec{S}}[U^{tw\dagger} \omega_{SK} U^{tw}] || \sigma_{\vec{K}}) \quad (\text{D52})$$

$$\geq \inf_{\sigma_{SK} \in \mathcal{T}} D(\Phi_{\vec{K}} || \sigma_{\vec{K}}) - 4m\epsilon \log K - h(\epsilon) \quad (\text{D53})$$

$$\geq (1 - 4m\epsilon) \log K - h(\epsilon), \quad (\text{D54})$$

where, in the last two inequalities, we have used the asymptotic continuity of the relative entropy and Lemma 4, respectively. Letting $n \rightarrow \infty$ and $\epsilon \rightarrow 0$, we finish the proof. ■

6. Proof of Theorem 7

As in the proof of Theorem 6, we have

$$\omega_{SK} = \mathcal{L}_{L^n A^n R^n \vec{C} \rightarrow SK}(\theta_{L A R \vec{C}}^{\otimes n}), \quad (\text{D55})$$

where \mathcal{L} is an LOCC operation with respect to the partition $:\overline{L^n A^n} : \overline{R^n} : \overline{C^n} :$. Now, following the proof of Theorem 2, we have that

$$F(\gamma_{SK}, \omega_{SK}) \geq 1 - \epsilon, \quad (\text{D56})$$

for some private state γ ; hence, there exists a projector Π_{SK}^{γ} corresponding to a γ -privacy test such that (see Proposition 1)

$$\text{Tr}[\Pi_{SK}^{\gamma} \omega_{SK}] \geq 1 - \epsilon. \quad (\text{D57})$$

On the other hand, from Theorem 1, we have

$$\text{Tr}[\Pi_{SK}^Z \sigma_{SK}^-] \leq \frac{1}{K}, \quad (\text{D58})$$

for any $\sigma \in \text{FS}(\vec{SK})$. Let us suppose a state $\sigma'_{LA\vec{R}\vec{C}} \in \text{FS}(\vec{LA}:\vec{R}:\vec{C})$, and let us define $\sigma_{SK}^- = \mathcal{L} \xrightarrow{L^n A^n} \xrightarrow{R^n} \xrightarrow{C^n} \xrightarrow{SK} (\sigma'_{LA\vec{R}\vec{C}})^{\otimes n}$, which is in $\text{FS}(\vec{SK})$. Hence, for all $\alpha > 1$, it holds that

$$\log_2 K \leq D_h^\epsilon(\omega_{SK}^- \parallel \sigma_{SK}^-) \quad (\text{D59})$$

$$\leq D_h^\epsilon(\theta_{LA\vec{R}\vec{C}}^{\otimes n} \parallel \sigma'_{LA\vec{R}\vec{C}}^{\otimes n}) \quad (\text{D60})$$

$$\leq \tilde{D}_\alpha(\theta_{LA\vec{R}\vec{C}}^{\otimes n} \parallel \sigma'_{LA\vec{R}\vec{C}}^{\otimes n}) + \frac{\alpha}{\alpha-1} \log_2 \left(\frac{1}{1-\epsilon} \right) \quad (\text{D61})$$

$$= n \tilde{D}_\alpha(\theta_{LA\vec{R}\vec{C}} \parallel \sigma'_{LA\vec{R}\vec{C}}) + \frac{\alpha}{\alpha-1} \log_2 \left(\frac{1}{1-\epsilon} \right). \quad (\text{D62})$$

The first inequality holds for any $\sigma \in \text{FS}(\vec{SK})$. The second inequality follows from the data-processing inequality. The third inequality follows from Eq. (A8). The equality is due to the additivity of \tilde{D}_α [64]. As the above holds for any $\sigma'_{LA\vec{R}\vec{C}} \in \text{FS}(\vec{LA}:\vec{R}:\vec{C})$, we obtain Theorem 7.

APPENDIX E: REPEATER AS A MULTIPARTITE CHANNEL

In order to provide bounds for more repeater protocols that involve two-way communication between Alice and Charlie or between Bob and Charlie before Charlie's measurement, we have to slightly generalize our results in Sec. V. Namely, in addition to trusted parties $\{\mathbf{X}_i\}_{i=1}^M = \{\mathbf{A}_a\}_a \cup \{\mathbf{B}_b\}_b \cup \{\mathbf{C}_c\}_c$, we can add a number of cooperative but untrusted parties $\{\tilde{\mathbf{X}}_i\}_{i=1}^{\tilde{M}} := \{\tilde{\mathbf{A}}_{\tilde{a}}\}_{\tilde{a} \in \tilde{\mathcal{A}}} \cup \{\tilde{\mathbf{B}}_{\tilde{b}}\}_{\tilde{b} \in \tilde{\mathcal{B}}} \cup \{\tilde{\mathbf{C}}_{\tilde{c}}\}_{\tilde{c} \in \tilde{\mathcal{C}}}$. Let us denote the quantum systems held by respective untrusted parties as $\tilde{A}'_{\tilde{a}}, \tilde{L}_{\tilde{a}}, \tilde{A}_{\tilde{a}}, \tilde{B}_{\tilde{b}}, \tilde{R}_{\tilde{b}}, \tilde{C}_{\tilde{c}}, \tilde{P}_{\tilde{c}}$ and redefine

$$\vec{A}' := \{A'_a\}_{a \in \mathcal{A}} \cup \{\tilde{A}'_{\tilde{a}}\}_{\tilde{a} \in \tilde{\mathcal{A}}}, \vec{A} := \{A_a\}_{a \in \mathcal{A}} \cup \{\tilde{A}_{\tilde{a}}\}_{\tilde{a} \in \tilde{\mathcal{A}}},$$

$$\vec{L} := \{L_a\}_{a \in \mathcal{A}} \cup \{\tilde{L}_{\tilde{a}}\}_{\tilde{a} \in \tilde{\mathcal{A}}},$$

$$\vec{B} := \{B_b\}_{b \in \mathcal{B}} \cup \{\tilde{B}_{\tilde{b}}\}_{\tilde{b} \in \tilde{\mathcal{B}}}, \vec{R} := \{R_b\}_{b \in \mathcal{B}} \cup \{\tilde{R}_{\tilde{b}}\}_{\tilde{b} \in \tilde{\mathcal{B}}},$$

$$\vec{C} := \{C_c\}_{c \in \mathcal{C}} \cup \{\tilde{C}_{\tilde{c}}\}_{\tilde{c} \in \tilde{\mathcal{C}}}, \vec{P} := \{P_c\}_{c \in \mathcal{C}} \cup \{\tilde{P}_{\tilde{c}}\}_{\tilde{c} \in \tilde{\mathcal{C}}},$$

while keeping the old definitions for \vec{K} and \vec{S} . We then assume that we have a multiplex channel $\mathcal{N}_{\vec{A}'\vec{B} \rightarrow \vec{A}\vec{C}}$ and LOCC operations \mathcal{L}^i , for $i = 1, \dots, n$, among trusted and

untrusted parties. However, we assume that as part of the last round of LOCC, \mathcal{L}^{n+1} , all subsystems belonging to untrusted parties are traced out, resulting in a state ω_{SK}^- among the trusted parties only. It is now easy to show that the proofs of Theorems 3 and 4 also go through in this slightly generalized scenario. Namely, tracing out parties in a fully separable state results in a fully separable state on the remaining parties, and by the monotonicity of the generalized divergences, inequalities (D16) and (D25) also hold if we trace out the untrusted parties in order to obtain ω . Note that the same does not hold true in the case of Theorem 2, where we have the distance to the set of biseparable states, which is not preserved under the trace-out.

Returning to the quantum key repeater, we can now identify Alice and Bob as two trusted parties and Charlie as an untrusted party and define a multiplex channel as the tensor product of the two channels from Alice to Charlie and Bob to Charlie, namely, $\mathcal{N}_{AB \rightarrow C}^{\text{repeater}} := \mathcal{N}_{A \rightarrow C_A}^1 \otimes \mathcal{N}_{B \rightarrow C_B}^2$, with $C := C_A C_B$. We include the local state preparation by Alice and Bob; the LOCC performed by Alice, Charlie, and Bob during key distillation protocols; and Bob's entanglement-swapping measurement and subsequent classical communication into the LOCC operations that interleave the uses of $\mathcal{N}_{AB \rightarrow C}^{\text{repeater}}$. Crucially, the final LOCC operation has to include the trace-out of Charlie's system, as he is an untrusted party. Application of the generalized versions of Theorem 3 or Theorem 4 then provides us with an upper bound on the achievable key rate in terms of $\min\{E_{\max, E}(\mathcal{N}_{AB \rightarrow C}^{\text{repeater}}), E_E^\infty(\mathcal{N}_{AB \rightarrow C}^{\text{repeater}})\}$. As has been shown in Ref. [74], there are examples of channels acting on finite-dimensional systems where the regularized relative entropy of entanglement is strictly less than max-relative entropy of entanglement, in which case, Theorem 4 provides tighter bounds than the ones provided in Ref. [50]. For tele-covariant channels, we can invoke Remark 5 and Theorem 5 to obtain bounds in terms of the relative entropy of entanglement.

Let us now consider repeater chains with more than a single repeater station. We assume a protocol where each channel has to be used the same number of times to get the desired fidelity. We consider Alice and Bob as trusted parties and the repeater stations C_1, \dots, C_l as cooperative but untrusted parties. Defining a multiplex channel $\mathcal{N}_{AC_1 \dots C_l \rightarrow C_1 \dots C_l B}^{\text{repeater chain}} := \mathcal{N}_{A \rightarrow C_1}^1 \otimes \mathcal{N}_{C_1 \rightarrow C_2}^2 \otimes \dots \otimes \mathcal{N}_{C_{l-1} \rightarrow C_l}^l \otimes \mathcal{N}_{C_l \rightarrow B}^{l+1}$ and including entanglement purification and swapping operations of all nesting levels into the LOCC operations, we then apply Theorem 3 or Theorem 4 to bound the achievable key rate between Alice and Bob by $\min\{E_{\max, E}(\mathcal{N}^{\text{repeater chain}}), E_E^\infty(\mathcal{N}^{\text{repeater chain}})\}$. If involved channels are tele-covariant, then we obtain bounds in terms of the relative entropy of entanglement.

APPENDIX F: LIMITATIONS ON SOME MDI-QKD PROTOTYPES

Following the discussion in Sec. VID, let us now consider MDI-QKD settings with the noise model for transmission of qubit systems from both \mathbf{A}_{a_1} and \mathbf{A}_{a_2} to Charlie through qubit channels given by either the depolarizing channel $\mathcal{D}_{A_i \rightarrow C_i}^l$ or the dephasing channel $\mathcal{D}_{A_i \rightarrow C_i}^s$:

$$\mathcal{D}_{A_i \rightarrow C_i}^l(\rho_{A_i}) = \lambda_l \rho_{C_i} + \frac{1 - \lambda_l}{2} \mathbb{1}_{C_i}, \quad (\text{F1})$$

$$\mathcal{D}_{A_i \rightarrow C_i}^s(\rho_{A_i}) = \lambda_s \rho_{C_i} + (1 - \lambda_s) \hat{Z} \rho_{C_i} \hat{Z}^\dagger, \quad (\text{F2})$$

where

$$-\frac{1}{3} \leq \lambda_l \leq 1, \quad 0 \leq \lambda_s \leq 1, \quad (\text{F3})$$

\hat{Z} is a Pauli-Z operator, and ρ is an arbitrary input state. Like the MDI-QKD setup with erasure channels discussed earlier, we assume that Charlie can perform a perfect Bell measurement $\mathcal{M}_{\bar{C} \rightarrow X}$ with probability q and failure probability $1 - q$. We notice that the multiplex channels $\mathcal{N}_{\bar{A} \rightarrow \bar{Z}}^{\text{MDI}, \mathcal{D}^l}$, $\mathcal{N}_{\bar{A} \rightarrow \bar{Z}}^{\text{MDI}, \mathcal{D}^s}$ for these MDI-QKD prototypes are also tele-covariant, which implies that the MDI-QKD capacities for respective MDI-QKD settings, i.e., with depolarizing channels and dephasing channels, are upper bounded as (see following subsections for proofs and plots (Figs. 7 and 8) for some values of q):

- (1) MDI-QKD with depolarizing channels \mathcal{D}^l [Eq. (F2)], where $-\frac{1}{3} \leq \lambda_l \leq 1$,

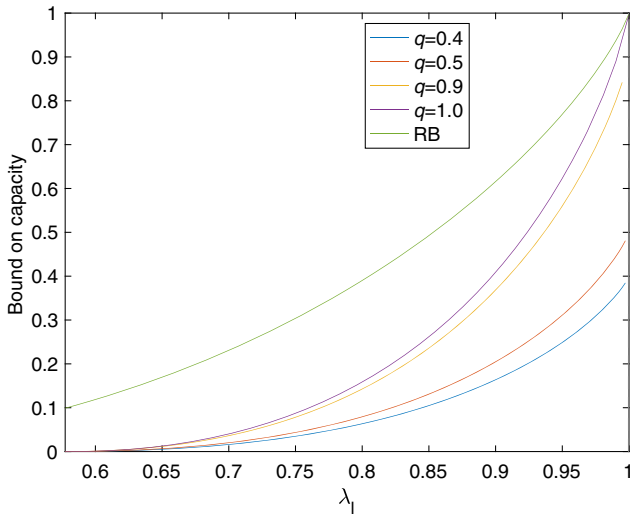


FIG. 7. Upper bounds [Eq. (F4)] on the secret key capacities for the MDI-QKD protocol with depolarizing channels for different values of parameters q and λ_l , in comparison to the RB bound [48].

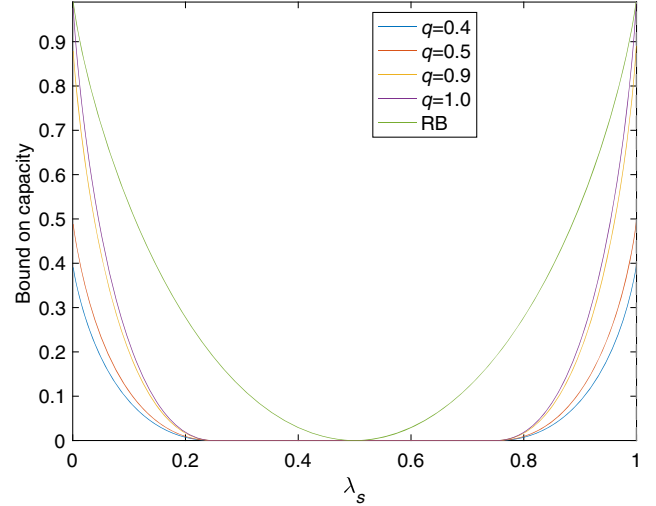


FIG. 8. Upper bounds [Eq. (F5)] on the secret key capacities for the MDI-QKD protocol with dephasing channels for different values of parameters q and λ_s , in comparison to the RB bound [48].

$$\tilde{P}_{\text{LOCC}}(\mathcal{N}^{\text{MDI}, \mathcal{D}^l}) \leq q \left[1 - h_2 \left(\frac{3}{4} \lambda_l^2 + \frac{1}{4} \right) \right] \quad (\text{F4})$$

for $\frac{1}{\sqrt{3}} < \lambda_l \leq 1$, and 0 otherwise.

- (2) MDI-QKD with dephasing channels \mathcal{D}^s [Eq. (F1)], where $0 \leq \lambda_s \leq 1$,

$$\tilde{P}_{\text{LOCC}}(\mathcal{N}^{\text{MDI}, \mathcal{D}^s}) \leq \begin{cases} q(1 - h_2(\frac{1}{2} p_-(\lambda_s))) & \text{for } \lambda_s > \frac{3}{4} \\ 0 & \text{for } \frac{1}{4} \leq \lambda_s \leq \frac{3}{4} \\ q(1 - h_2(\frac{1}{2} p_-(1 - \lambda_s))) & \text{for } \lambda_s < \frac{1}{4}, \end{cases} \quad (\text{F5})$$

where $p_-(x) := 4x^2 - 3x + 1$.

1. MDI-QKD via depolarizing channels

In this section, we show a bound on MDI-QKD (or, equivalently, on a particular type of quantum repeater). In the latter setup, there are three stations: A , B , and an intermediate one $C \equiv C_A C_B$. We consider the links AC_A and CB_B to be depolarizing channels \mathcal{D}^s , both with the same parameter λ_l [see Eq. (F2)]. We also consider that the Bell measurement, followed by communication of the results to both the parties, happens only with probability q . With probability $(1 - q)$, the state of C is just traced out. We call the multiplex channel for a given MDI-QKD setup composed of depolarizing channels \mathcal{D}^l with Bell measurement, which happens with probability q in total, a q -depolarizing-MDIQKD channel.

The upper bound that we derive below quantitatively demonstrates that the operation of distillation of entanglement along the links does not commute with the operation of entanglement swapping. Indeed, even for $q = 1$, if one does the Bell measurement first, the output key is zero for $\lambda_l \leq (1/\sqrt{3})$.

We are interested in the Choi-Jamiolkowski state of the q -depolarizing-MDIQKD channel, which we obtain from the Choi states (up to local unitary as the input state is Ψ^-) of the two depolarizing channels. The latter two states read $\lambda_l \Psi^- + (1 - \lambda_l) \frac{\mathbb{1}_{AB}}{4}$. The Choi state ρ_{AB}^{out} reads

$$\begin{aligned} \rho_{AB}^{\text{out}} := & \frac{\lambda_l^2 q}{4} [\Psi_{AB}^- \otimes |00\rangle\langle 00|_{I_A I_B} + \Psi_{AB}^+ \otimes |11\rangle\langle 11|_{I_A I_B} \\ & + \Phi_{AB}^- \otimes |22\rangle\langle 22|_{I_A I_B} + \Phi_{AB}^+ \otimes |33\rangle\langle 33|_{I_A I_B}] \\ & \otimes |00\rangle\langle 00|_{I'_A I'_B} \\ & + (1 - \lambda_l^2) q \frac{\mathbb{1}_{AB}}{4} \otimes \frac{1}{4} \sum_{i=0}^3 |ii\rangle\langle ii|_{I_A I_B} \otimes |00\rangle\langle 00|_{I'_A I'_B} \\ & + (1 - q) \frac{\mathbb{1}_{AB}}{4} \otimes |\perp\rangle\langle \perp|_{I_A I_B} \otimes |11\rangle\langle 11|_{I'_A I'_B}. \end{aligned} \quad (\text{F6})$$

Let us examine this case. First, with probability $(1 - q)$, the parties are left with the initial state on AB , which is $\frac{\mathbb{1}_{AB}}{4}$, and the “flag” $|11\rangle\langle 11|_{I'_A I'_B}$ reporting error in the Bell measurement. With probability q , they obtain a flag $|00\rangle\langle 00|_{I'_A I'_B}$, which informs us that the Bell measurement was successful. They also receive the classical result of the Bell measurement: $\{|ii\rangle\langle ii|_{I_A I_B}\}_{i=0}^3$. Only with probability λ_l^2 does this measurement result in the output of the appropriate Bell state on AB . With probability $(1 - \lambda_l^2) = (1 - \lambda_l)\lambda_l + \lambda_l(1 - \lambda_l) + (1 - \lambda_l)^2$, we have one of three possibilities with respective probabilities: (i) teleportation of $\mathbb{1}_{C_B}/2$ from C_A to A with probability $\lambda_l(1 - \lambda_l)$, (ii) teleportation of $\mathbb{1}_{C_A}/2$ from C_B to B with probability $(1 - \lambda_l)\lambda_l$, and (iii) a Bell measurement on systems $C_A C_B$ of the state $(\mathbb{1}_{C_A}/4) \otimes (\mathbb{1}_{C_B}/4)$ followed by communication of the outcomes [with probability $(1 - \lambda_l)^2$]. As one can check by inspection, all three operations result in the state $\frac{\mathbb{1}_{AB}}{4}$ on system AB .

The relative entropy of ρ_{AB}^{out} reads

$$E_R(\rho_{AB}^{\text{out}}) \leq q E_R(\rho_{AB|00}^{\text{out}}) + (1 - q) E_R(\rho_{AB|11}^{\text{out}}) \quad (\text{F7})$$

$$= q E_R(\rho_{AB|00}^{\text{out}}), \quad (\text{F8})$$

where $\rho_{AB|11}^{\text{out}} = (\mathbb{1}_{AB}/4) \otimes |\perp\rangle\langle \perp|_{I_A I_B} \otimes |11\rangle\langle 11|_{I'_A I'_B}$ and $\rho_{AB|00}^{\text{out}}$ is such that $(1 - q)\rho_{AB|11} + q\rho_{AB|00} = \rho_{AB}$. We have used the convexity of the relative entropy and the

fact that it is zero for a maximally mixed state. We then observe that

$$\begin{aligned} E_R(\rho_{AB|00}^{\text{out}}) = & E_R\left(\left(\sum_{i=0}^3 \lambda_l^2 |\psi_i\rangle\langle \psi_i|_{AB} + (1 - \lambda_l^2) \frac{\mathbb{1}_{AB}}{4}\right)\right. \\ & \left. \otimes |ii\rangle\langle ii|_{I_A I_B} \otimes |00\rangle\langle 00|_{I'_A I'_B}\right), \end{aligned} \quad (\text{F9})$$

where $|\psi_i\rangle\langle \psi_i|$ are the Bell states. Next, we use the fact that for each i , the state $\lambda_l^2 |\psi_i\rangle\langle \psi_i|_{AB} + (1 - \lambda_l^2) \frac{\mathbb{1}_{AB}}{4}$ is a Bell diagonal state. A Bell diagonal state of the form $\sum_j p_j |\psi_j\rangle\langle \psi_j|$ has E_R equal to $1 - h(p_{\max})$, where $p_{\max} = \max_j p_j$ is the maximal of the weights of the Bell state $|\psi_j\rangle\langle \psi_j|$ in the mixture, or 0 if $p_{\max} \leq \frac{1}{2}$. In our case, $p_{\max} = \lambda_l^2 + (1 - \lambda_l^2)/4$. Thus, via convexity and Eq. (F8), we obtain that

$$E_R(\rho_{AB}^{\text{out}}) \leq q \left[1 - h_2\left(\lambda_l^2 + \frac{(1 - \lambda_l^2)}{4}\right) \right] \quad (\text{F10})$$

for $\lambda_l^2 + (1 - \lambda_l^2)/4 > 1/2$, and 0 otherwise. The condition $\lambda_l^2 + (1 - \lambda_l^2)/4 > 1/2$ on λ_l is equivalent to $\lambda_l > (1/\sqrt{3})$. This implies that for $q = 1$, the bound is zero for $\lambda_l \in (\frac{1}{3}, (1/\sqrt{3})]$, for which the depolarizing channel is nonzero, and hence, its private capacity is nonzero as well. We interpret this as the noncommutativity of the independent and identically distributed (i.i.d.) Bell measurement and entanglement distillation. Indeed, for this range of λ_l , given access to an isotropic state $\rho(\lambda_l)$, one can distill $E_D(\rho(\lambda_l)) = (1 - h_2(\lambda_l))$ of entanglement, and hence, the quantum capacity $\mathcal{Q}(\mathcal{D}^l) = 1 - h_2(\lambda_l)$ (or zero for $\lambda_l \leq 1/3$). On the other hand, this amount of key becomes inaccessible when the Bell measurement is done first.

2. MDI-QKD via dephasing channels

In this section, we consider two dephasing channels [Eq. (F1)] between Alice and Charlie and Bob and Charlie. We again observe that the operation of distillation and i.i.d. entanglement swapping via the Bell measurement do not commute. Altering them leads to different amounts of key in the output. We use the fact that MDI-QKD via the dephasing channel is teleportation covariant.

Note that the Choi-Jamiolkowski state (up to local unitary operations as the input state is Ψ^-) of the dephasing channel equals $\lambda_s \Psi^- + (1 - \lambda_s) \Psi^+ = (2\lambda_s - 1) \Psi^- + (2 - 2\lambda_s) \rho_{cl}$, with $\rho_{cl} = \frac{1}{2}(|01\rangle\langle 01| + |10\rangle\langle 10|)$. Hence, the Choi-Jamiolkowski state of the dephasing-MDIQKD channel reads

$$\begin{aligned}
\rho_{AB}^{\text{out}} := & (2\lambda_s - 1)^2 q \Psi_{AB}^- \otimes \sum_{i=0}^3 |ii\rangle\langle ii|_{I_A I_B} \otimes |00\rangle\langle 00|_{I'_A I'_B} \\
& + (2 - 2\lambda_s)(2\lambda_s - 1) q \rho_{cl}^{AB} \\
& \otimes \frac{1}{4} \sum_{i=0}^3 |ii\rangle\langle ii|_{I_A I_B} \otimes |00\rangle\langle 00|_{I'_A I'_B} \\
& + (2 - 2\lambda_s) q \frac{\mathbb{1}_{AB}}{4} \otimes \frac{1}{4} \sum_{i=0}^3 |ii\rangle\langle ii|_{I_A I_B} \otimes |00\rangle\langle 00|_{I'_A I'_B} \\
& + (1 - q) \frac{\mathbb{1}_{AB}}{4} \otimes |\perp\rangle\langle \perp|_{I_A I_B} \otimes |11\rangle\langle 11|_{I'_A I'_B}, \quad (\text{F11})
\end{aligned}$$

given that Alice has performed the control-Pauli operations on her systems AI_A . We can safely assume that this decoding has been done because the local unitary operation does not change the relative entropy of entanglement. The first case is a straightforward result of correct entanglement swapping. Regarding the next term, with probability $(2 - 2\lambda_s) \times (2\lambda_s - 1)$, a subsystem C_A of the state ρ_{cl} gets correctly teleported to A , and hence, finally, ρ_{cl}^{AB} is shared by Alice and Bob. However, with probability $(2 - 2\lambda_s) = (2 - 2\lambda_s)^2 +$

$(2 - 2\lambda_s)(2\lambda_s - 1)$, the resulting state is maximally mixed because, with probability $(2 - 2\lambda_s)^2$, the state on system C is traced out; hence, a product of subsystems of ρ_{cl}^{AB} is an output. On the other hand, with probability $(2 - 2\lambda_s) \times (2\lambda_s - 1)$, subsystem C_B of the state ρ_{cl} is teleported to Bob; however, Bob does not do the decoding. It is then straightforward to check that $\frac{1}{4} \sum_{i=0}^3 \sigma_i^B \otimes \mathbb{1}_A \rho_{cl}^{AB} \hat{\sigma}_i^B \otimes \mathbb{1}_A$, with $\hat{\sigma}_i$ being Pauli operators, is the maximally mixed state of two qubits.

The relative entropy of ρ_{AB}^{out} reads

$$E_R(\rho_{AB}^{\text{out}}) \leq q E_R(\rho_{AB|00}^{\text{out}}) + (1 - q) E_R(\rho_{AB|11}^{\text{out}}) \quad (\text{F12})$$

$$= q E_R(\rho_{AB|00}^{\text{out}}), \quad (\text{F13})$$

where $\rho_{AB|11}^{\text{out}} = (\mathbb{1}_{AB}/4) \otimes |\perp\rangle\langle \perp|_{I_A I_B} \otimes |11\rangle\langle 11|_{I'_A I'_B}$ and $\rho_{AB|00}^{\text{out}}$ is such that $(1 - q)\rho_{AB|11} + q\rho_{AB|00} = \rho_{AB}$. We have again used the convexity of the relative entropy and the fact that it is zero for a maximally mixed state. We then observe that

$$E_R(\rho_{AB|00}^{\text{out}}) = E_R\left((2\lambda_s - 1)^2 |\Psi^-\rangle\langle \Psi^-|_{AB} + (2 - 2\lambda_s)(2\lambda_s - 1) \rho_{cl}^{AB} + (2 - 2\lambda_s) \frac{\mathbb{1}_{AB}}{4}\right), \quad (\text{F14})$$

where we have neglected systems $I_A I_B$ and $I'_A I'_B$ due to subadditivity of E_R and the fact that it is zero for both the states $\sum_{i=0}^3 |ii\rangle\langle ii|_{I_A I_B}$ and $|00\rangle\langle 00|_{I'_A I'_B}$. The resulting state is Bell diagonal [note that $\rho_{cl}^{AB} = \frac{1}{2}(|\Psi^-\rangle\langle \Psi^-| + |\Psi^+\rangle\langle \Psi^+|)$]; it is thus sufficient to find the largest weight of a Bell state to compute its relative entropy. Bell diagonal states are separable if the largest weight is less than or equal to half, i.e., when none of the Bell states (Φ^+ , Φ^- , Ψ^+ , Ψ^-) has weight greater than $1/2$.

For the case $\lambda_s \geq \frac{1}{2}$, the state $|\Psi^-\rangle\langle \Psi^-|$ is in the mixed state $\rho_{AB|00}^{\text{out}}$ with probability $(2\lambda_s - 1)^2 + (2 - 2\lambda_s)(2\lambda_s - 1) + (2 - 2\lambda_s)/4 = \frac{1}{2}(4\lambda_s^2 - 3\lambda_s + 1)$.

Thus, keeping the structure of the Choi state of the dephasing channel in mind, we arrive at the following bound:

$$E_R(\rho_{AB}^{\text{out}}) \leq \begin{cases} q(1 - h_2(\frac{1}{2}p_-(\lambda_s))) & \text{for } \lambda_s > \frac{3}{4} \\ 0 & \text{for } \frac{1}{4} \leq \lambda_s \leq \frac{3}{4} \\ q(1 - h_2(\frac{1}{2}p_-(1 - \lambda_s))) & \text{for } \lambda_s < \frac{1}{4}, \end{cases} \quad (\text{F15})$$

where $p_-(x) := 4x^2 - 3x + 1$.

APPENDIX G: COMPLEXITY OF FINDING LOWER BOUNDS OF THE SKA RATE FOR THE BIDIRECTIONAL NETWORK

Here, we briefly comment on the complexity of finding a subgraph, which allows us to realize the conference key agreement with the capacity indicated by the inequality (75). As we show, the complexity is a polynomial of low degree $O(n^2)$. In what follows, a minimum spanning tree is a tree with a minimal sum of the weights of its edges. A minimum bottleneck spanning tree is the one in which the edge with the highest weight has the lowest possible value for the considered graph.

The algorithm of finding the maximal of the minimal edges over all spanning trees of the graph is as follows.

- (1) Find the maximal weight of the edges of G (denoted as M).
- (2) Find the minimum spanning tree T_{MST} in the graph $G' = (V_G, E_G)$, which is the same as G but with the weights of the edges changed from $w(e)$ to $M - w(e)$, where $M \equiv \max_{e' \in E_G} w(e')$.
- (3) Find the minimal weight of the edges in T , denoted w_{min} . Return $M - w_{\text{min}}$.

The correctness of this algorithm follows from the fact that every minimum spanning tree is a minimal bottleneck spanning tree. Finding the highest weight of the edges of this tree that is as low as possible is the opposite task from

ours. Indeed, we aim at finding trees with the lowest weight over its edges to be as high as possible, which is why we search for the minimal spanning tree in the graph with converted edges to $M \equiv \max_{e' \in E_G} w(e') - w(e)$. Next, we use the fact that $\min_{T \subseteq G} \max_{e \in E_T} [M - w(e)] = M - \max_{T \subseteq G} \min_{e \in E_T} w(e)$, so $M - w_{\min}$ is the solution. The overall time complexity of this algorithm is $O(m + n \log n)$. Indeed, the first step takes $O(m)$ time. The next two take $O(m + n \log n)$, where finding the minimum spanning tree is via Prim's algorithm based on the data structure called the Fibonacci heap [117]. The final step takes $O(n \log n)$, which is the time for sorting the weights of edges (e.g., by the QuickSort algorithm). Taking into account that m scales pessimistically as n^2 , we obtain $O(n^2)$ as the worst-case complexity.

To summarize, the value of the lower bound can be found efficiently on a classical computer, given that all the capacities describing the bidirectional network are known and represented in the form of a graph.

APPENDIX H: KEY DISTILLATION FROM STATES—PLOTS

To calculate our upper bounds, we utilize the technique of semidefinite programming (SDP) with MATLAB (version) library “SDPT3 4.0” [154], see Ref. [134]. We calculate upper bounds for several cases, incorporating both Φ_M^{GHZ} states and Φ_M^{W} states. First, we vary the number of copies of the state that enter the protocol; second, we make calculations for multipartite states with the number of parties exceeding three. Finally, we extend our consideration to states subjected to dephasing or depolarizing noises characterized in Eq. (H1) (each qubit is subjected to noise separately). We investigate the effect of noise in the case of a different number of copies and different number of parties:

$$\rho_{\text{noisy}} = \mathcal{D}^{\otimes M}(\rho), \quad (\text{H1})$$

for \mathcal{D} given by

$$\mathcal{D}_{\text{deph}}^q(\omega) = q\omega + (1-q)\sigma_z\omega\sigma_z, \quad (\text{H2})$$

$$\mathcal{D}_{\text{depol}}^q(\omega) = q\omega + (1-q)\frac{\mathbb{1}}{2}, \quad (\text{H3})$$

where σ_z is the Pauli Z matrix and q is the noise parameter.

We present the plots for the upper bound on the key rate distilled from both Φ_M^{GHZ} and Φ_M^{W} states and tensor powers of them. The plots are a function of the ε parameter controlling the fidelity of the target state $\rho_{\bar{A}}$ with respect to a private state.

We compare the performance of our upper bound and choice of biseparable states for a tripartite single copy state in Figs. 9 and 10. In the control plot in Fig. 9, for the

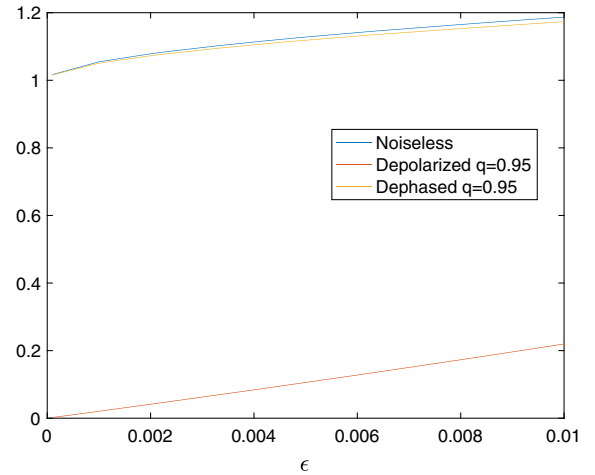


FIG. 9. Plot of ε -hypothesis-testing upper bound on conference key rate for a single copy of the Φ_3^{GHZ} state, for noiseless, dephased, and depolarized cases.

noiseless Φ_M^{GHZ} state, the upper bound, as expected, exhibits the value to be just above 1 for the chosen range of ε , which indicates that the ε -hypothesis-testing upper bound is not too loose. For the single copy tripartite Φ_M^{W} state, the value of the upper bound in Fig. 10 for $\varepsilon \approx 0$ is below 0.6, which is below the value of the rate of the optimal LOCC asymptotic protocol, approximately 0.643 per copy [80]. In the case of two copies of bipartite Φ_M^{W} in Fig. 11, we obtain an upper bound that for $\varepsilon \approx 0$ has a value of around 1.18, which is significantly above the $\frac{2}{3}$ achieved by the protocol described earlier in this Appendix, and 1.286, which is an asymptotic limit for the state being two copies of the Φ_M^{W} state (Theorem 2 in Ref. [80]). Both of these results are in agreement with the fact that single-copy and two-copy one-shot protocols constitute a very

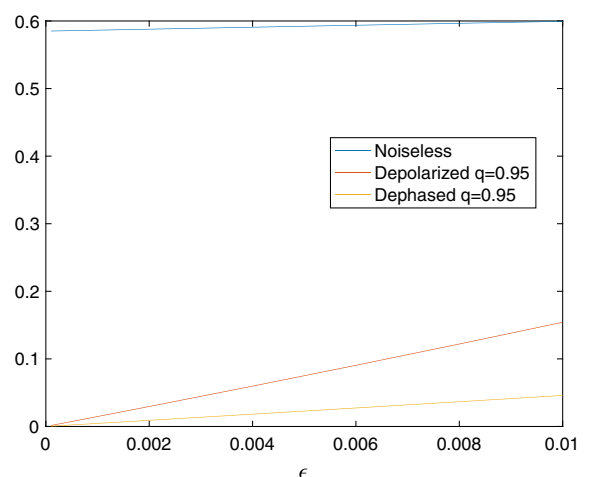


FIG. 10. Plot of ε -hypothesis-testing upper bound on the conference key rate for a single copy of the Φ_3^{W} state, for noiseless, dephased, and depolarized cases.

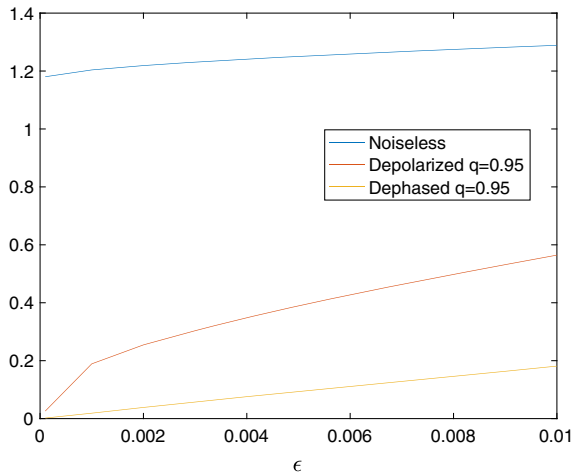


FIG. 11. Plot of ε -hypothesis-testing upper bound on the conference key rate for two copies of the Φ_3^W state, for noiseless, dephased, and depolarized cases.

limited class of protocols compared to those available for calculating the asymptotic limit. For two copies of the Φ_M^W state, the large gap between our upper bound for the conference key rate and the rate of the Φ_M^{GHZ} state distillation protocol makes us think that, indeed, the former is larger than the latter. However, a formal proof is still missing. Moreover, we notice that the optimal protocol Φ_M^W to Φ_M^{GHZ} conversion has to incorporate at least three copies of the Φ_M^W state because our ε -hypothesis-testing upper bound is smaller than the asymptotic limit for Φ_M^{GHZ} distillation.

- [1] C. H. Bennett and G. Brassard, *Quantum Cryptography: Public Key Distribution and Coin Tossing*, in *International Conference on Computer System and Signal Processing, 1984* (IEEE, New York, 1984), Vol. 175, p. 8.
- [2] A. K. Ekert, *Quantum Cryptography Based on Bell's Theorem*, *Phys. Rev. Lett.* **67**, 661 (1991).
- [3] J. P. Dowling and G. J. Milburn, *Quantum Technology: The Second Quantum Revolution*, *Phil. Trans. R. Soc. A* **361**, 1655 (2003).
- [4] R. Renner, Ph.D. thesis, ETH Zürich (2005), <https://arxiv.org/abs/quant-ph/0512258>.
- [5] T.-Y. Chen, H. Liang, Y. Liu, W.-Q. Cai, L. Ju, W.-Y. Liu, J. Wang, H. Yin, K. Chen, Z.-B. Chen, C.-Z. Peng, and J.-W. Pan, *Field Test of a Practical Secure Communication Network with Decoy-State Quantum Cryptography*, *Opt. Express* **17**, 6540 (2009).
- [6] G. Vallone, D. Bacco, D. Dequal, S. Gaiarin, V. Luceri, G. Bianco, and P. Villoresi, *Experimental Satellite Quantum Communications*, *Phys. Rev. Lett.* **115**, 040502 (2015).
- [7] Q. Zhang, F. Xu, Y.-A. Chen, C.-Z. Peng, and J.-W. Pan, *Large Scale Quantum Key Distribution: Challenges and Solutions*, *Opt. Express* **26**, 24260 (2018).
- [8] B. Hensen, H. Bernien, A. E. Dréau, A. Reiserer, N. Kalb, M. S. Blok, J. Ruitenber, R. F. L. Vermeulen, R. N. Schouten, C. Abellán, W. Amaya, V. Pruneri, M. W. Mitchell, M. Markham, D. J. Twitchen, D. Elkouss, S. Wehner, T. H. Taminiau, and R. Hanson, *Loophole-Free Bell Inequality Violation Using Electron Spins Separated by 1.3 Kilometres*, *Nature (London)* **526**, 682 (2015).
- [9] S. J. Pauka, K. Das, R. Kalra, A. Moini, Y. Yang, M. Trainer, A. Bousquet, C. Cantaloube, N. Dick, G. C. Gardner, M. J. Manfra, and D. J. Reilly, *A Cryogenic Interface for Controlling Many Qubits*, [arXiv:1912.01299](https://arxiv.org/abs/1912.01299).
- [10] C. E. Bradley, J. Randall, M. H. Abobeih, R. C. Berrevoets, M. J. Degen, M. A. Bakker, M. Markham, D. J. Twitchen, and T. H. Taminiau, *A Ten-Qubit Solid-State Spin Register with Quantum Memory up to One Minute*, *Phys. Rev. X* **9**, 031045 (2019).
- [11] Y.-A. Chen *et al.*, *An Integrated Space-to-Ground Quantum Communication Network over 4,600 Kilometres*, *Nature (London)* **589**, 214 (2021).
- [12] P. W. Shor, *Algorithms for Quantum Computation: Discrete Logarithms and Factoring*, in *Proceedings of the 35th Annual Symposium on Foundations of Computer Science* (IEEE Computer Society Press, Los Alamitos, California, 1994), pp. 124–134.
- [13] K. Chen and H.-K. Lo, *Conference Key Agreement and Quantum Sharing of Classical Secrets with Noisy GHZ States*, in *Proceedings of the International Symposium on Information Theory, 2005* (IEEE, New York, 2005), pp. 1607–1611.
- [14] R. Augusiak and P. Horodecki, *Multipartite Secret Key Distillation and Bound Entanglement*, *Phys. Rev. A* **80**, 042307 (2009).
- [15] D. M. Greenberger, M. A. Horne, and A. Zeilinger, *Going Beyond Bell's Theorem*, in *Bell's Theorem, Quantum Theory and Conceptions of the Universe*, edited by M. Kafatos, *Fundamental Theories of Physics* Vol. 37 (Springer, Dordrecht, 1989), https://doi.org/10.1007/978-94-017-0849-4_10.
- [16] H. J. Kimble, *The Quantum Internet*, *Nature (London)* **453**, 1023 (2008).
- [17] S. Wehner, D. Elkouss, and R. Hanson, *Quantum Internet: A Vision for the Road Ahead*, *Science* **362**, eaam9288 (2018).
- [18] X.-S. Ma, T. Herbst, T. Scheidl, D. Wang, S. Kropatschek, W. Naylor, B. Wittmann, A. Mech, J. Kofler, E. Anisimova, V. Makarov, T. Jennewein, R. Ursin, and A. Zeilinger, *Quantum Teleportation over 143 Kilometres Using Active Feed-Forward*, *Nature (London)* **489**, 269 (2012).
- [19] S.-K. Liao *et al.*, *Satellite-to-Ground Quantum Key Distribution*, *Nature (London)* **549**, 43 (2017).
- [20] K. Azuma, K. Tamaki, and H.-K. Lo, *All-Photonic Quantum Repeaters*, *Nat. Commun.* **6**, 6787 (2015).
- [21] H.-J. Briegel, W. Dür, J. I. Cirac, and P. Zoller, *Quantum Repeaters: The Role of Imperfect Local Operations in Quantum Communication*, *Phys. Rev. Lett.* **81**, 5932 (1998).
- [22] W. Dür, H.-J. Briegel, J. I. Cirac, and P. Zoller, *Quantum Repeaters Based on Entanglement Purification*, *Phys. Rev. A* **59**, 169 (1999).

- [23] W. J. Munro, K. Azuma, K. Tamaki, and K. Nemoto, *Inside Quantum Repeaters*, *IEEE J. Sel. Top. Quantum Electron.* **21**, 78 (2015).
- [24] V. Makarov, A. Anisimov, and J. Skaar, *Effects of Detector Efficiency Mismatch on Security of Quantum Cryptosystems*, *Phys. Rev. A* **74**, 022313 (2006).
- [25] B. Qi, C.-H. F. Fung, H.-K. Lo, and X. Ma, *Time-Shift Attack in Practical Quantum Cryptosystems*, *Quantum Inf. Comput.* **7**, 073 (2007).
- [26] E. Diamanti, H.-K. Lo, B. Qi, and Z. Yuan, *Practical Challenges in Quantum Key Distribution*, *npj Quantum Inf.* **2**, 16025 (2016).
- [27] H.-K. Lo, M. Curty, and B. Qi, *Measurement-Device-Independent Quantum Key Distribution*, *Phys. Rev. Lett.* **108**, 130503 (2012).
- [28] S. L. Braunstein and S. Pirandola, *Side-Channel-Free Quantum Key Distribution*, *Phys. Rev. Lett.* **108**, 130502 (2012).
- [29] S. Pirandola, C. Ottaviani, G. Spedalieri, C. Weedbrook, S. L. Braunstein, S. Lloyd, T. Gehring, C. S. Jacobsen, and U. L. Andersen, *High-Rate Measurement-Device-Independent Quantum Cryptography*, *Nat. Photonics* **9**, 397 (2015).
- [30] Y. Fu, H.-L. Yin, T.-Y. Chen, and Z.-B. Chen, *Long-Distance Measurement-Device-Independent Multiparty Quantum Communication*, *Phys. Rev. Lett.* **114**, 090501 (2015).
- [31] M. Lucamarini, Z. L. Yuan, J. F. Dynes, and A. J. Shields, *Overcoming the Rate–Distance Limit of Quantum Key Distribution without Quantum Repeaters*, *Nature (London)* **557**, 400 (2018).
- [32] X. Ma, P. Zeng, and H. Zhou, *Phase-Matching Quantum Key Distribution*, *Phys. Rev. X* **8**, 031043 (2018).
- [33] K. Tamaki, H.-K. Lo, W. Wang, and M. Lucamarini, *Information Theoretic Security of Quantum Key Distribution Overcoming the Repeaterless Secret Key Capacity Bound*, [arXiv:1805.05511](https://arxiv.org/abs/1805.05511).
- [34] J. Lin and N. Lütkenhaus, *Simple Security Analysis of Phase-Matching Measurement-Device-Independent Quantum Key Distribution*, *Phys. Rev. A* **98**, 042332 (2018).
- [35] C. Cui, Z.-Q. Yin, R. Wang, W. Chen, S. Wang, G.-C. Guo, and Z.-F. Han, *Twin-Field Quantum Key Distribution without Phase Postselection*, *Phys. Rev. Applied* **11**, 034053 (2019).
- [36] M. Curty, K. Azuma, and H.-K. Lo, *Simple Security Proof of Twin-Field Type Quantum Key Distribution Protocol*, *npj Quantum Inf.* **5**, 1 (2019).
- [37] H. Liu, W. Wang, K. Wei, X.-T. Fang, L. Li, N.-L. Liu, H. Liang, S.-J. Zhang, W. Zhang, H. Li, L. You, Z. Wang, H.-K. Lo, T.-Y. Chen, F. Xu, and J.-W. Pan, *Experimental Demonstration of High-Rate Measurement-Device-Independent Quantum Key Distribution over Asymmetric Channels*, *Phys. Rev. Lett.* **122**, 160501 (2019).
- [38] M. Minder, M. Pittaluga, G. L. Roberts, M. Lucamarini, J. F. Dynes, Z. L. Yuan, and A. J. Shields, *Experimental Quantum Key Distribution Beyond the Repeaterless Secret Key Capacity*, *Nat. Photonics* **13**, 334 (2019).
- [39] I. W. Primaatmaja, E. Lavie, K. T. Goh, C. Wang, and C. C. W. Lim, *Versatile Security Analysis of Measurement-Device-Independent Quantum Key Distribution*, *Phys. Rev. A* **99**, 062332 (2019).
- [40] K. Horodecki, M. Horodecki, P. Horodecki, and J. Oppenheim, *Secure Key from Bound Entanglement*, *Phys. Rev. Lett.* **94**, 160502 (2005).
- [41] M. Christandl and A. Winter, *Squashed Entanglement: An Additive Entanglement Measure*, *J. Math. Phys. (N.Y.)* **45**, 829 (2004).
- [42] C. H. Bennett, D. P. DiVincenzo, J. A. Smolin, and W. K. Wootters, *Mixed-State Entanglement and Quantum Error Correction*, *Phys. Rev. A* **54**, 3824 (1996).
- [43] V. Vedral, M. B. Plenio, M. A. Rippin, and P. L. Knight, *Quantifying Entanglement*, *Phys. Rev. Lett.* **78**, 2275 (1997).
- [44] V. Vedral and M. B. Plenio, *Entanglement Measures and Purification Procedures*, *Phys. Rev. A* **57**, 1619 (1998).
- [45] M. Horodecki, P. Horodecki, and R. Horodecki, *General Teleportation Channel, Singlet Fraction, and Quasidistillation*, *Phys. Rev. A* **60**, 1888 (1999).
- [46] N. Datta, *Max-Relative Entropy of Entanglement, Alias Log Robustness*, *Int. J. Quantum. Inform.* **07**, 475 (2009).
- [47] M. Takeoka, S. Guha, and M. Wilde, *Fundamental Rate-Loss Tradeoff for Optical Quantum Key Distribution*, *Nat. Commun.* **5**, 5235 (2014).
- [48] S. Pirandola, R. Laurenza, C. Ottaviani, and L. Banchi, *Fundamental Limits of Repeaterless Quantum Communications*, *Nat. Commun.* **8**, 15043 (2017); see also, [arXiv:1512.04945](https://arxiv.org/abs/1512.04945).
- [49] M. M. Wilde, M. Tomamichel, and M. Berta, *Converse Bounds for Private Communication over Quantum Channels*, *IEEE Trans. Inf. Theory* **63**, 1792 (2017).
- [50] M. Christandl and A. Müller-Hermes, *Relative Entropy Bounds on Quantum, Private and Repeater Capacities*, *Commun. Math. Phys.* **353**, 821 (2017).
- [51] S. Das, S. Bäuml, and M. M. Wilde, *Entanglement and Secret-Key-Agreement Capacities of Bipartite Quantum Interactions and Read-Only Memory Devices*, *Phys. Rev. A* **101**, 012344 (2020).
- [52] S. Bäuml, S. Das, and M. M. Wilde, *Fundamental Limits on the Capacities of Bipartite Quantum Interactions*, *Phys. Rev. Lett.* **121**, 250504 (2018).
- [53] S. Das, Ph.D. thesis, Louisiana State University (2018), <https://arxiv.org/abs/1901.05895>.
- [54] R. Laurenza and S. Pirandola, *General Bounds for Sender-Receiver Capacities in Multipoint Quantum Communications*, *Phys. Rev. A* **96**, 032318 (2017).
- [55] K. P. Seshadreesan, M. Takeoka, and M. M. Wilde, *Bounds on Entanglement Distillation and Secret Key Agreement for Quantum Broadcast Channels*, *IEEE Trans. Inf. Theory* **62**, 2849 (2016).
- [56] M. Takeoka, K. P. Seshadreesan, and M. M. Wilde, *Unconstrained Capacities of Quantum Key Distribution and Entanglement Distillation for Pure-Loss Bosonic Broadcast Channels*, *Phys. Rev. Lett.* **119**, 150501 (2017).
- [57] S. Bäuml, M. Christandl, K. Horodecki, and A. Winter, *Limitations on Quantum Key Repeaters*, *Nat. Commun.* **6**, 6908 (2015).

- [58] K. Azuma, A. Mizutani, and H.-K. Lo, *Fundamental Rate-Loss Tradeoff for the Quantum Internet*, *Nat. Commun.* **7**, 13523 (2016).
- [59] L. Rigovacca, G. Kato, S. Bäuml, M. Kim, W. J. Munro, and K. Azuma, *Versatile Relative Entropy Bounds for Quantum Networks*, *New J. Phys.* **20**, 013033 (2018).
- [60] S. Pirandola, *End-to-End Capacities of a Quantum Communication Network*, *Commun. Phys.* **2**, 51 (2019)
- [61] S. Bäuml and K. Azuma, *Fundamental Limitation on Quantum Broadcast Networks*, *Quantum Sci. Technol.* **2**, 024004 (2017).
- [62] K. Horodecki, M. Horodecki, P. Horodecki, and J. Oppenheim, *General Paradigm for Distilling Classical Key from Quantum States*, *IEEE Trans. Inf. Theory* **55**, 1898 (2009).
- [63] M. M. Wilde, A. Winter, and D. Yang, *Strong Converse for the Classical Capacity of Entanglement-Breaking and Hadamard Channels via a Sandwiched Rényi Relative Entropy*, *Commun. Math. Phys.* **331**, 593 (2014).
- [64] M. Müller-Lennert, F. Dupuis, O. Szehr, S. Fehr, and M. Tomamichel, *On Quantum Rényi Entropies: A New Definition and Some Properties*, *J. Math. Phys. (N.Y.)* **54**, 122203 (2013).
- [65] F. Buscemi and N. Datta, *The Quantum Capacity of Channels with Arbitrarily Correlated Noise*, *IEEE Trans. Inf. Theory* **56**, 1447 (2010).
- [66] C. H. Bennett, H. J. Bernstein, S. Popescu, and B. Schumacher, *Concentrating Partial Entanglement by Local Operations*, *Phys. Rev. A* **53**, 2046 (1996).
- [67] S. Das, S. Khatri, and J. P. Dowling, *Robust Quantum Network Architectures and Topologies for Entanglement Distribution*, *Phys. Rev. A* **97**, 012335 (2018).
- [68] I. Devetak and A. Winter, *Distillation of Secret Key and Entanglement from Quantum States*, *Proc. R. Soc. A* **461**, 207 (2005).
- [69] S. Pirandola, R. García-Patrón, S. L. Braunstein, and S. Lloyd, *Direct and Reverse Secret-Key Capacities of a Quantum Channel*, *Phys. Rev. Lett.* **102**, 050503 (2009).
- [70] M. Christandl and A. Winter, *Squashed Entanglement: An Additive Entanglement Measure*, *J. Math. Phys. (N.Y.)* **45**, 829 (2004).
- [71] R. R. Tucci, *Quantum Entanglement and Conditional Information Transmission*, [arXiv:quant-ph/9909041](https://arxiv.org/abs/quant-ph/9909041).
- [72] R. R. Tucci, *Entanglement of Distillation and Conditional Mutual Information*, [arXiv:quant-ph/0202144](https://arxiv.org/abs/quant-ph/0202144).
- [73] D. Yang, K. Horodecki, M. Horodecki, P. Horodecki, J. Oppenheim, and W. Song, *Squashed Entanglement for Multipartite States and Entanglement Measures Based on the Mixed Convex Roof*, *IEEE Trans. Inf. Theory* **55**, 3375 (2009).
- [74] K. Fang and H. Fawzi, *Geometric Rényi Divergence and Its Applications in Quantum Channel Capacities*, [arXiv:1909.05758](https://arxiv.org/abs/1909.05758).
- [75] K. Azuma and G. Kato, *Aggregating Quantum Repeaters for the Quantum Internet*, *Phys. Rev. A* **96**, 032332 (2017).
- [76] S. Bäuml, K. Azuma, G. Kato, and D. Elkouss, *Linear Programs for Entanglement and Key Distribution in the Quantum Internet*, *Commun. Phys.* **3**, 55 (2020).
- [77] M. Christandl and R. Ferrara, *Private States, Quantum Data Hiding, and the Swapping of Perfect Secrecy*, *Phys. Rev. Lett.* **119**, 220506 (2017).
- [78] P. van Loock, W. Alt, C. Becher, O. Benson, H. Boche, C. Deppe, J. Eschner, S. Höfling, D. Meschede, P. Michler *et al.*, *Extending Quantum Links: Modules for Fiber- and Memory-Based Quantum Repeaters*, *Adv. Quantum Technol.* **3**, 1900141 (2020).
- [79] C. H. Bennett, S. Popescu, D. Rohrlich, J. A. Smolin, and A. V. Thapliyal, *Exact and Asymptotic Measures of Multipartite Pure-State Entanglement*, *Phys. Rev. A* **63**, 012307 (2000).
- [80] J. A. Smolin, F. Verstraete, and A. Winter, *Entanglement of Assistance and Multipartite State Distillation*, *Phys. Rev. A* **72**, 052317 (2005).
- [81] B. Fortescue and H.-K. Lo, *Random Bipartite Entanglement from W and W -Like States*, *Phys. Rev. Lett.* **98**, 260501 (2007).
- [82] S. Kıntaş and S. Turgut, *Transformations of W -Type Entangled States*, *J. Math. Phys. (N.Y.)* **51**, 092202 (2010).
- [83] W. Cui, E. Chitambar, and H. K. Lo, *Randomly Distilling W -Class States into General Configurations of Two-Party Entanglement*, *Phys. Rev. A* **84**, 052301 (2011).
- [84] P. Vrana and M. Christandl, *Asymptotic Entanglement Transformation between W and GHZ States*, *J. Math. Phys. (N.Y.)* **56**, 022204 (2015).
- [85] C. Spee, J. I. de Vicente, D. Sauerwein, and B. Kraus, *Entangled Pure State Transformations via Local Operations Assisted by Finitely Many Rounds of Classical Communication*, *Phys. Rev. Lett.* **118**, 040503 (2017).
- [86] P. Vrana and M. Christandl, *Distillation of Greenberger–Horne–Zeilinger States by Combinatorial Methods*, *IEEE Trans. Inf. Theory* **65**, 5945 (2019).
- [87] A. Streltsov, C. Meignant, and J. Eisert, *Rates of Multipartite Entanglement Transformations*, *Phys. Rev. Lett.* **125**, 080502 (2020).
- [88] A. Einstein, B. Podolsky, and N. Rosen, *Can Quantum-Mechanical Description of Physical Reality Be Considered Complete?*, *Phys. Rev.* **47**, 777 (1935).
- [89] E. Chitambar, D. Leung, L. Mančinska, M. Ozols, and A. Winter, *Everything You Always Wanted to Know about LOCC (but Were Afraid to Ask)*, *Commun. Math. Phys.* **328**, 303 (2014).
- [90] Y. Polyanskiy and S. Verdú, *Arimoto Channel Coding Converse and Rényi Divergence*, in *Proceedings of the 48th Annual Allerton Conference on Communication, Control, and Computation* (2010), pp. 1327–1333, <https://ieeexplore.ieee.org/document/5707067>.
- [91] N. Sharma and N. A. Warsi, *On the Strong Converse for the Quantum Channel Capacity Theorems*, *Phys. Rev. Lett.* **110**, 080501 (2013).
- [92] H. Umegaki, *Conditional Expectations in an Operator Algebra, IV (Entropy and Information)*, *Kodai Math. Sem. Rep.* **14**, 59 (1962).
- [93] N. Datta, *Min- and Max-Relative Entropies and a New Entanglement Monotone*, *IEEE Trans. Inf. Theory* **55**, 2816 (2009).

- [94] L. Wang and R. Renner, *One-Shot Classical-Quantum Capacity and Hypothesis Testing*, *Phys. Rev. Lett.* **108**, 200501 (2012).
- [95] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, England, 2000).
- [96] R. Augusiak, D. Cavalcanti, G. Pretico, and A. Acín, *Perfect Quantum Privacy Implies Nonlocality*, *Phys. Rev. Lett.* **104**, 230401 (2010).
- [97] S. Bäuml, S. Das, X. Wang, and M. M. Wilde, *Resource Theory of Entanglement for Bipartite Quantum Channels*, [arXiv:1907.04181](https://arxiv.org/abs/1907.04181).
- [98] G. Gour and C. M. Scandolo, *The Entanglement of a Bipartite Channel*, *Phys. Rev. A* **103**, 062422 (2021).
- [99] N. Friis, G. Vitagliano, M. Malik, and M. Huber, *Entanglement Certification from Theory to Experiment*, *Nat. Rev. Phys.* **1**, 72 (2019).
- [100] P. Contreras-Tejada, C. Palazuelos, and J. I. de Vicente, *Resource Theory of Entanglement with a Unique Multipartite Maximally Entangled State*, *Phys. Rev. Lett.* **122**, 120503 (2019).
- [101] C. H. Bennett, A. W. Harrow, D. W. Leung, and J. A. Smolin, *On the Capacities of Bipartite Hamiltonians and Unitary Gates*, *IEEE Trans. Inf. Theory* **49**, 1895 (2003).
- [102] E. Kaur and M. M. Wilde, *Amortized Entanglement of a Quantum Channel and Approximately Teleportation-Simulable Channels*, *J. Phys. A* **51**, 035303 (2018).
- [103] K. Fang, O. Fawzi, R. Renner, and D. Sutter, *A Chain Rule for the Quantum Relative Entropy*, *Phys. Rev. Lett.* **124**, 100501 (2020).
- [104] M. Tomamichel and V. Y. F. Tan, *Second-Order Asymptotics for the Classical Capacity of Image-Additive Quantum Channels*, *Commun. Math. Phys.* **338**, 103 (2015).
- [105] K. Goodenough, D. Elkouss, and S. Wehner, *Assessing the Performance of Quantum Repeaters for All Phase-Insensitive Gaussian Bosonic Channels*, *New J. Phys.* **18**, 063005 (2016).
- [106] M. M. Wilde and H. Qi, *Energy-Constrained Private and Quantum Capacities of Quantum Channels*, *IEEE Trans. Inf. Theory* **64**, 7802 (2018).
- [107] D. Gottesman and I. L. Chuang, *Demonstrating the Viability of Universal Quantum Computation Using Teleportation and Single-Qubit Operations*, *Nature (London)* **402**, 390 (1999).
- [108] J. I. Cirac, W. Dür, B. Kraus, and M. Lewenstein, *Entangling Operations and Their Implementation Using a Small Amount of Entanglement*, *Phys. Rev. Lett.* **86**, 544 (2001).
- [109] W. Dür, M. J. Bremner, and H. J. Briegel, *Quantum Simulation of Interacting High-Dimensional Systems: The Influence of Noise*, *Phys. Rev. A* **78**, 052325 (2008).
- [110] Y. Wu, J. Zhou, X. Gong, Y. Guo, Z.-M. Zhang, and G. He, *Continuous-Variable Measurement-Device-Independent Multipartite Quantum Communication*, *Phys. Rev. A* **93**, 022325 (2016).
- [111] C. Ottaviani, C. Lupo, R. Laurenza, and S. Pirandola, *Modular Network for High-Rate Quantum Conferencing*, *Commun. Phys.* **2**, 118 (2019).
- [112] M. Żukowski, A. Zeilinger, M. A. Horne, and A. K. Ekert, *Event-Ready-Detectors Bell Experiment via Entanglement Swapping*, *Phys. Rev. Lett.* **71**, 4287 (1993).
- [113] T. C. Ralph and G. J. Pryde, *Optical Quantum Computation*, in *Progress in Optics* (Elsevier, New York, 2010), pp. 209–269.
- [114] M. Grassl, T. Beth, and T. Pellizzari, *Codes for the Quantum Erasure Channel*, *Phys. Rev. A* **56**, 33 (1997).
- [115] F.-Y. Lu, Z.-Q. Yin, R. Wang, G.-J. Fan-Yuan, S. Wang, D.-Y. He, W. Chen, W. Huang, B.-J. Xu, G.-C. Guo, and Z.-F. Han, *Practical Issues of Twin-Field Quantum Key Distribution*, *New J. Phys.* **21**, 123030 (2019).
- [116] R. J. Wilson, *Introduction to Graph Theory* (Longman, England, 1996).
- [117] C. E. Leiserson, R. L. Rivest, T. H. Cormen, and C. Stein, *Introduction to Algorithms* (MIT Press, Cambridge, MA, 2001), Vol. 6.
- [118] W. Dür, G. Vidal, and J. I. Cirac, *Three Qubits Can Be Entangled in Two Inequivalent Ways*, *Phys. Rev. A* **62**, 062314 (2000).
- [119] R. Horodecki, P. Horodecki, M. Horodecki, and K. Horodecki, *Quantum Entanglement*, *Rev. Mod. Phys.* **81**, 865 (2009).
- [120] L. Amico, R. Fazio, A. Osterloh, and V. Vedral, *Entanglement in Many-Body Systems*, *Rev. Mod. Phys.* **80**, 517 (2008).
- [121] D. Home, D. Saha, and S. Das, *Multipartite Bell-Type Inequality by Generalizing Wigner's Argument*, *Phys. Rev. A* **91**, 012102 (2015).
- [122] B. Fortescue and H.-K. Lo, *Random-Party Entanglement Distillation in Multiparty States*, *Phys. Rev. A* **78**, 012348 (2008).
- [123] W. Cui, W. Helwig, and H.-K. Lo, *Bounds on Probability of Transformations between Multipartite Pure States*, *Phys. Rev. A* **81**, 012111 (2010).
- [124] A. Cabello, *Multiparty Key Distribution and Secret Sharing Based on Entanglement Swapping*, [arXiv:quant-ph/0009025](https://arxiv.org/abs/quant-ph/0009025).
- [125] V. Scarani and N. Gisin, *Quantum Key Distribution Between N Partners: Optimal Eavesdropping and Bell's Inequalities*, *Phys. Rev. A* **65**, 012311 (2001).
- [126] R. Augusiak and P. Horodecki, *W-Like Bound Entangled States and Secure Key Distillation*, *Europhys. Lett.* **85**, 50001 (2009).
- [127] F. Grasselli, H. Kampermann, and D. Bruß, *Conference Key Agreement with Single-Photon Interference*, *New J. Phys.* **21**, 123002 (2019).
- [128] K. Horodecki, L. Pankowski, M. Horodecki, and P. Horodecki, *Low-Dimensional Bound Entanglement with One-Way Distillable Cryptographic Key*, *IEEE Trans. Inf. Theory* **54**, 2621 (2008).
- [129] Ł. Pankowski and M. Horodecki, *Low-Dimensional Quite Noisy Bound Entanglement with a Cryptographic Key*, *J. Phys. A* **44**, 035301 (2010).
- [130] F. Verstraete, J. Dehaene, and B. de Moor, *On the Geometry of Entangled States*, *J. Mod. Opt.* **49**, 1277 (2002).
- [131] M. Horodecki, P. Horodecki, and R. Horodecki, *Limits for Entanglement Measures*, *Phys. Rev. Lett.* **84**, 2014 (2000).

- [132] S. Ishizaka and M. B. Plenio, *Publishers Note: Multi-particle Entanglement under Asymptotic Positive-Partial-Transpose-Preserving Operations*, *Phys. Rev. A* **72**, 059907 (2005).
- [133] M. B. Plenio and S. Virmani, *An Introduction to Entanglement Measures*, *Quantum Inf. Comput.* **7**, 1 (2007).
- [134] See Supplemental Material at <http://link.aps.org/supplemental/10.1103/PhysRevX.11.041016> for codes to get plots.
- [135] M. Pivoluska, M. Huber, and M. Malik, *Layered Quantum Key Distribution*, *Phys. Rev. A* **97**, 032312 (2018).
- [136] C. E. Shannon, *Two-Way Communication Channels*, in *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability, Volume 1: Contributions to the Theory of Statistics* (University of California Press, Berkeley, California, 1961), pp. 611–644.
- [137] A. El Gamal and Y.-H. Kim, *Network Information Theory* (Cambridge University Press, Cambridge, England, 2012), p. 709.
- [138] K. Brádler, P. Hayden, D. Touchette, and M. M. Wilde, *Trade-off Capacities of the Quantum Hadamard Channels*, *Phys. Rev. A* **81**, 062312 (2010).
- [139] Q. Wang, S. Das, and M. M. Wilde, *Hadamard Quantum Broadcast Channels*, *Quantum Inf. Process.* **16**, 248 (2017).
- [140] F. Leditzky, M. A. Alhejji, J. Levin, and G. Smith, *Playing Games with Multiple Access Channels*, *Nat. Commun.* **11**, 1497 (2020).
- [141] S.-H. Tan and P. P. Rohde, *The Resurgence of the Linear Optics Quantum Interferometer—Recent Advances & Applications*, *Rev. Phys.* **4**, 100030 (2019).
- [142] S. Das and M. M. Wilde, *Quantum Rebound Capacity*, *Phys. Rev. A* **100**, 030302(R) (2019).
- [143] R. L. Frank and E. H. Lieb, *Monotonicity of a Relative Rényi Entropy*, *J. Math. Phys. (N.Y.)* **54**, 122201 (2013).
- [144] S. Beigi, *Sandwiched Rényi Divergence Satisfies Data Processing Inequality*, *J. Math. Phys. (N.Y.)* **54**, 122202 (2013).
- [145] F. Hiai and D. Petz, *The Proper Formula for Relative Entropy and Its Asymptotics in Quantum Probability*, *Commun. Math. Phys.* **143**, 99 (1991).
- [146] H. Nagaoka, *Strong Converse Theorems in Quantum Information Theory*, in *Proceedings of ERATO Workshop on Quantum Information Science* (2001), p. 33; also appeared in *Asymptotic Theory of Quantum Statistical Inference*, edited by M. Hayashi (World Scientific, Singapore, 2005).
- [147] T. Ogawa and H. Nagaoka, *Strong Converse and Stein’s Lemma in Quantum Hypothesis Testing*, *IEEE Trans. Inf. Theory* **46**, 2428 (2000).
- [148] T. Cooney, M. Mosonyi, and M. M. Wilde, *Strong Converse Exponents for a Quantum Channel Discrimination Problem and Quantum-Feedback-Assisted Communication*, *Commun. Math. Phys.* **344**, 797 (2016).
- [149] O. Fawzi, P. Hayden, I. Savov, P. Sen, and M. M. Wilde, *Classical Communication over a Quantum Interference Channel*, *IEEE Trans. Inf. Theory* **58**, 3670 (2012).
- [150] I. Devetak, *The Private Classical Capacity and Quantum Capacity of a Quantum Channel*, *IEEE Trans. Inf. Theory* **51**, 44 (2005).
- [151] J. Yard, P. Hayden, and I. Devetak, *Quantum Broadcast Channels*, *IEEE Trans. Inf. Theory* **57**, 7147 (2011).
- [152] J. Yard, P. Hayden, and I. Devetak, *Capacity Theorems for Quantum Multiple-Access Channels: Classical-Quantum and Quantum-Quantum Capacity Regions*, *IEEE Trans. Inf. Theory* **54**, 3091 (2008).
- [153] A. Ambainis, M. Mosca, A. Tapp, and R. de Wolf, *Private Quantum Channels*, *Proceedings of the IEEE 41st Annual Symposium on Foundations of Computer Science* (2000), pp. 547–553, <https://ieeexplore.ieee.org/document/892142>.
- [154] K. C. Toh, M. J. Todd, and R. H. Tütüncü, *SDPT3—A Matlab Software Package for Semidefinite Programming, Version 1.3*, *Optim. Methods Software* **11**, 545 (1999).

Fundamental limitations on the device-independent quantum conference key agreementKarol Horodecki ^{1,2}, Marek Winzewski ^{1,3} and Siddhartha Das ^{4,5}¹*International Centre for Theory of Quantum Technologies, University of Gdańsk, 80-308 Gdańsk, Poland*²*Institute of Informatics, National Quantum Information Centre, Faculty of Mathematics, Physics and Informatics, University of Gdańsk, Wita Stwosza 57, 80-308 Gdańsk, Poland*³*Institute of Theoretical Physics and Astrophysics, National Quantum Information Centre, Faculty of Mathematics, Physics and Informatics, University of Gdańsk, Wita Stwosza 57, 80-308 Gdańsk, Poland*⁴*Centre for Quantum Information and Communication, École polytechnique de Bruxelles, Université libre de Bruxelles, Brussels 1050, Belgium*⁵*Center for Security, Theory and Algorithmic Research, International Institute of Information Technology, Hyderabad, Gachibowli, Telangana 500032, India*

(Received 21 October 2021; accepted 23 December 2021; published 7 February 2022)

We provide several general upper bounds on the rate of a key secure against a quantum adversary in the device-independent conference key agreement (DI-CKA) scenario. They include bounds by reduced entanglement measures and those based on multipartite secrecy monotones such as a multipartite squashed entanglement-based measure, which we refer to as reduced c-squashed entanglement. We compare the latter bound with the known lower bound for the protocol of conference key distillation based on the parity Clauser-Horne-Shimony-Holt game. We also show that the gap between the DI-CKA rate and the device-dependent rate is inherited from the bipartite gap between device-independent and device-dependent key rates, giving examples that exhibit the strict gap.

DOI: [10.1103/PhysRevA.105.022604](https://doi.org/10.1103/PhysRevA.105.022604)**I. INTRODUCTION**

Building a quantum secure internet is one of the most important challenges in the field of quantum technologies [1,2]. It would ensure worldwide information-theoretically secure communication. The idea of quantum repeaters [3–5] gives hope that this dream will come true. However, the level of quantum security proposed originally in a seminal article by Bennett and Brassard [6] seems to be insufficient due to the fact that on the way between an honest manufacturer and an honest user, an active hacker can change with the inner workings of a quantum device, making it totally insecure [7]. Indeed, the hardware Trojan-horse attacks on random number generators are known [8], and the active hacking on quantum devices became a standard testing approach since the seminal attack by Makarov [9]. The idea of device-independent (DI) security overcomes this obstacle [7,10] (see also [11] and references therein). Although difficult to be done in practice, it has been demonstrated quite recently in several recent experiments [12–14].

In parallel, the study of the limitations of this approach in terms of upper bounds on the distillable key has been put forward [15–18]. However, these approaches focus on point-to-point quantum device-independent secure communication. In this paper we introduce the upper bounds on the performance of the device-independent conference key agreement (DI-CKA) [19,20]. The task of the conference agreement is to distribute to $N > 2$ honest parties the same secure key for one-time-pad encryption. A protocol achieving this task in a device-independent manner has been shown in Ref. [20]. We

set an upper bound on the performance of such protocols in a network setting.

We focus on physical behaviors with N users (for arbitrary $N > 2$), where each user is both the sender and receiver of the behavior treated as a black box. This situation is a special case of a network describable with a multiplex quantum channel where inputs and outputs are classical with quantum phenomena going inside the physical behavior [21]. All N trusted parties have the role of both the sender and receiver from the channel and their goal is to obtain a secret key in a device-independent way against a quantum adversary. Aiming at upper bounds on the device-independent key, we narrow the consideration to the independent and identically distributed case. In this scenario, the honest parties share n identical devices. All the N parties set (classical) inputs $\mathbf{x} = (x_1, \dots, x_N)$ to each of the n shared devices $P(\mathbf{a}|\mathbf{x})$ and receive (classical) outputs $\mathbf{a} = (a_1, \dots, a_N)$ from each of them. We restrict our consideration to quantum devices. Such devices are realized by certain measurements $\mathcal{M} \equiv \otimes_{i=1}^N \mathbf{M}_{a_i}^{x_i}$ on quantum states $\rho_{A_1, \dots, A_N} \equiv \rho_{N(A)}$. We define these devices $(\rho_{N(A)}, \mathcal{M}) = \text{Tr}[\rho_{N(A)} \otimes_{i=1}^N \mathbf{M}_{a_i}^{x_i}]$.

In this work we provide upper bounds on the device-independent conference key distillation rates for arbitrary multipartite states. As the first main result, we introduce a multipartite generalization of the cc-squashed entanglement provided in Ref. [22] and developed in Ref. [18]. With a little abuse of notation with respect to that used in Refs. [16,18], for the sake of the reader, we will omit the fact that the measure is multipartite as well as reduce the abbreviation *cc* in its name and here call it just reduced c-squashed entanglement,

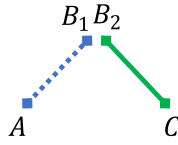


FIG. 1. Depiction of a construction of a tripartite state $\rho_{A(B_1B_2)C}$ with a gap $K_{\text{DI}}(\rho_{A(B_1B_2)C}) < K_{\text{DD}}(\rho_{A(B_1B_2)C})$ from a state ρ_{AB_1} (blue dotted line) satisfying $K^\downarrow(\rho_{AB_1}) < K_{\text{DD}}(\rho_{AB_1})$ (provided in Ref. [16]) and any state ρ_{B_2C} (green solid line) satisfying $K_{\text{DD}}(\rho_{B_2C}) \geq K_{\text{DD}}(\rho_{AB_1})$, e.g., the singlet state.

denoting it by $E_{\text{sq,dev}}^c$. We show that $E_{\text{sq,dev}}^c$ upper bounds the device-independent key rate in the independent and identically distributed setting, achieved by protocols which use a single input to generate the key $\hat{\mathbf{x}}$, denoted by $K_{\text{DI,dev}}^{\text{id},\hat{\mathbf{x}}}$. The subscript dev in the notation refers to the fact that the adversary has to mimic the statistics of the honestly implemented device. We then generalize this to the case when only some parameters of the device have to be reproduced by the attack and refer to quantities with the subscript par. Typical parameters are the level of violation of a Bell inequality and the quantum bit error rate. In the above finding, we use the notion of multipartite squashed entanglement given in Ref. [23] as well as the dual one also studied in Ref. [24]. Therein, the abbreviation *c* stands for classical, as the systems of the honest parties are classical due to the measurement accordingly to the definition. The bound reads

$$K_{\text{DI,dev}}^{\text{id},\hat{\mathbf{x}}}(\rho_{N(A)}, \mathcal{M}) \leq E_{\text{sq,dev}}^c(\rho_{N(A)}, \mathcal{M}) \\ \equiv \inf_{(\sigma, \mathcal{N})=(\rho_{N(A)}, \mathcal{M})} \frac{1}{N-1} I(A_1 : \dots : A_N \downarrow E)_{\mathcal{N}(\hat{\mathbf{x}}) \otimes \mathbb{1} \sigma}. \quad (1)$$

In the above \downarrow denotes the action of any channel transforming E to some system E' and $I(A_1 : \dots : A_N | E)_\sigma = \sum_{i=1}^N S(A_i | E)_\sigma - S(A_1, \dots, A_N | E)_\sigma$, with $S(X|Y)_\sigma$ the conditional von Neumann entropy of the state $\sigma_{X,Y}$. The quantity $I(A_1 : \dots : A_N | E')$ (after the action of the channel on E) is evaluated on the classical-quantum state emerging from the measurement $\mathbf{N}_a^{\hat{\mathbf{x}}}$ corresponding to input $\hat{\mathbf{x}}$ of the device (σ, \mathcal{N}) , on systems A_1, \dots, A_N . Let us note here that, due to the findings of Ref. [18], for $N = 2$ and the case when the system E' , as the output of a channel, is classical, the above bound is equal to the intrinsic information given in Ref. [17] for the case of a single measurement generating the key.

Our technique is based on the approach of Ref. [25], where the upper bounds on a key distillable against a quantum adversary via local operations and public communication (LOPC) were studied. To achieve this, we generalize the upper bound via (quantum) intrinsic information to the case in which the adversary's system can be of infinite dimension. This technical contribution is necessary, as in the case of a device-independent attack, the dimension of the attacking state can be infinite. Indeed, while measurements \mathbf{x} produce from the attacking state σ finite-dimensional results \mathbf{a} yielding a quantum behavior $P(\mathbf{a} | \mathbf{x}) = \text{Tr}[\sigma \mathbf{M}_a^{\mathbf{x}}]$, the system

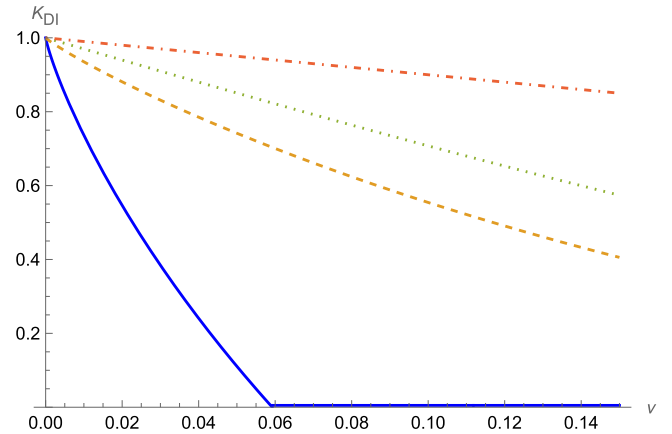


FIG. 2. Plot of upper and lower bounds on the DI-CKA of Ref. [20]. The yellow dashed line represents an upper bound (not optimized) on the upper bound $\frac{1}{N-1} I(N(A) \downarrow E)$ from Eq. (40) with the attack strategy in Eq. (68). The green dotted line represents an upper bound (not optimized) on $D_N(N(A) \downarrow E)$ from Corollary 3. The red dash-dotted curve is the trivial upper bound obtained in Corollary 6 via the relative entropy of entanglement bound $(1 - \nu)$. The blue solid line represents the lower bound from Ref. [20].

of the adversary, which may hold purification of the state σ , can still be of infinite dimension.¹

We then compare the obtained upper bounds with the lower bound on the DI-CKA provided in Ref. [20] (see Fig. 2). We obtain the plots by considering simplification of the reduced *c*-squashed entanglement and its dual measure. Namely, the extension to the adversarial system of the state attacking the honest parties device is classical, i.e., diagonal in the computational basis. For that reason, the bound which we use is in fact a secrecy monotone, called (multipartite) intrinsic information [26], and its dual.

As the second main result, we show how to construct multipartite states with a strict gap between the rate of the quantum device-independent conference key rates K_{DI} and quantum device-dependent conference key rates K_{DD} . As a proxy, we use a bipartite state that satisfies $K_{\text{DD}} > K^\downarrow$, where K^\downarrow is the reduced distillable key introduced in Ref. [16]. The reduced distillable key of $\rho_{N(A)}$ is the maximum value of the choice of measurements \mathcal{M} of the distillable key of the adversarial state σ minimized over the choices of the state σ and measurements \mathcal{N} so that the device $(\sigma, \mathcal{N}) \equiv \text{Tr}[\sigma \otimes_{i=1}^N \mathbf{N}_{a_i}^{x_i}]$ is equal to the honestly implemented device (ρ, \mathcal{M}) . The mentioned gap between K_{DI} and K_{DD} means that also in a multipartite case for some states ρ_{A_1, \dots, A_N} (and any $N > 2$), there is neither a Bell-like inequality that can be used for testing nor a distillation protocol based on LOPC that can achieve $K_{\text{DI}}(\rho_{A_1, \dots, A_N}) = K_{\text{DD}}(\rho_{A_1, \dots, A_N})$. See Fig. 1.

Finally, we discuss the issue of genuine nonlocality [11] and genuine entanglement in the context of the DI-CKA [27]. As the third main result, we provide a nontrivial bound

¹In that we have filled in the gap in the proofs of Corollaries 3 and 4 of Ref. [18], where implicit assumption of the adversary holding finite-dimensional state has been made.

on the device-independent key achievable in a parallel measurement scenario, when all the parties set all values of the inputs x_i in parallel. Furthermore, generalizing reduced bipartite entanglement measures in Ref. [18] to multipartite entanglement measures, we show that the reduced regularized relative entropy of genuine entanglement [21] upper bounds the DI-CKA rate of multipartite quantum states. We further focus on the performance of protocols using a single input for key generation, as using such protocols is standard practice (see, e.g., [28]).

The remainder of this paper is organized as follows. Section II is devoted to basic facts and provides bounds on the DI conference key via entanglement measures. In Sec. III, we develop an upper bound on the DI-CKA via reduced c-squashed entanglement. In Sec. IV, we develop an upper bound via a dual version of reduced c-squashed entanglement. In Sec. V, we provide particular examples for the performance of upper bounds considered in the paper. In Sec. VI, we provide examples of multipartite states which exhibit a fundamental gap between the device-dependent and -independent secure key rates. Section VII discusses the connection between genuine nonlocality, entanglement, and DI-CKA. We conclude in Sec. VIII with a summary and some directions for future study.

II. BOUNDS ON DEVICE-INDEPENDENT KEY DISTILLATION RATE OF STATES

In this section, we introduce the scenario of device-independent conference key distillation from n identical devices, each shared by N honest users. We then introduce definitions and facts used in subsequent sections.

Consider a setup wherein N multiple trusted spatially separated users (allies) have to extract a secret key, i.e., conference key, against the quantum adversary. Since we aim at upper bounds on the device-independent conference key, we assume that the parties share n identical devices. The device has its honest implementation, which is reflected by the state and measurement, denoted by (ρ, \mathcal{M}) , that were intended to be delivered by a provider. The adversary may replace this honest implementation with a different device (σ, \mathcal{N}) , however, such that it yields the same input-output statistics as the honest one. Typically, the statistics tested by the allies are the level of violation of some Bell inequality and the quantum bit error rate, i.e., the probability that the outputs of the honest parties are not equal to each other given the raw key has been generated. In some cases, we will also consider the full statistics reflected by the pair (ρ, \mathcal{M}) . We note here that the state σ can be finite or infinite dimensional, as we do not restrict the strategies of the adversary in that respect.

The honest device is given by $\mathcal{M} \equiv \{M_{a_1}^{x_1} \otimes M_{a_2}^{x_2} \otimes \dots \otimes M_{a_N}^{x_N}\}_{\mathbf{a}|\mathbf{x}}$, where $\mathbf{x} := (x_1, x_2, \dots, x_N)$ and $\mathbf{a} := (a_1, a_2, \dots, a_N)$ for some $N \in \mathbb{N}$. For each $i \in [N] := \{1, 2, \dots, N\}$, the set $\{a_i\}$ denotes the finite set of measurement outcomes for measurement choices x_i . The measurement outcomes, i.e., outputs of the device, are secure from the adversary and assumed to be in the possession of the receivers (allies). The joint probability distribution is

given as

$$p(\mathbf{a}|\mathbf{x}) = \text{Tr} [M_{a_1}^{x_1} \otimes M_{a_2}^{x_2} \otimes \dots \otimes M_{a_N}^{x_N} \rho_{A_1 A_2 \dots A_N}] \quad (2)$$

for measurement \mathcal{M} on N -partite state $\rho_{N(A)}$ defined on the separable Hilbert space $\mathcal{H}_{A_1} \otimes \mathcal{H}_{A_2} \otimes \dots \otimes \mathcal{H}_{A_N}$; in what follows we will use $N(A) \equiv A_1 \dots A_N$ for the ease of notation. The tuple $\{\rho, \mathcal{M}\}$, where $\mathcal{M} := (\{M_{a_1}^{x_1}\}_{x_1}, \{M_{a_2}^{x_2}\}_{x_2}, \dots, \{M_{a_N}^{x_N}\}_{x_N})$, is called the quantum strategy of the distribution. The number of inputs $\{x_i\}$ and corresponding possible outputs $\{a_i\}$ of the local measurement at A_i are arbitrarily finite in general. We denote the identity superoperator by id and the identity operator by $\mathbb{1}$.

Let $\omega(\rho, \mathcal{M})$ denote the violation of the given multipartite Bell-type inequality \mathcal{B} by state ρ when the measurement settings are given by \mathcal{M} . We note that by multipartite Bell-type inequality we mean any inequality derived using locally realistic hidden variable (LRHV) theories (see, e.g., [29–37]) such that any violation of a given inequality by a density operator implies the nonexistence of an LRHV model for the device represented by this state and some measurements. There are families of Bell-type inequalities directly based on the joint probability distribution of local measurements that get violated by all pure multipartite (genuinely) entangled states [35,36]. On the other hand, there are Bell-type inequalities based on correlation functions of local measurements for which some families of pure multipartite (genuinely) entangled states satisfy the inequalities [38].

Let $P_{\text{err}}(\rho, \mathcal{M})$ denote the expected quantum bit error rate (QBER). Both the Bell violation and the QBER are functions of the probability distribution of the behavior. In addition, $\Phi_N^{\text{GHZ}} := |\Phi_N^{\text{GHZ}}\rangle\langle\Phi_N^{\text{GHZ}}|$ defines the N -partite Greenberger-Horne-Zeilinger (GHZ)w state.

$$|\Phi_N^{\text{GHZ}}\rangle = \frac{1}{\sqrt{d}} \sum_{i=0}^{d-1} |i\rangle_{A_1} \otimes |i\rangle_{A_2} \otimes \dots \otimes |i\rangle_{A_N} \quad (3)$$

for $d = \min_i \dim(\mathcal{H}_{A_i})$. For $N = 2$, Φ_2^{GHZ} is a maximally entangled (Bell) state Φ_2 .

If $\{p(\mathbf{a}|\mathbf{x})\}_{\mathbf{a}|\mathbf{x}}$ obtained from (ρ, \mathcal{M}) and another pair of states and measurements (σ, \mathcal{N}) are the same, we write $(\sigma, \mathcal{N}) = (\rho, \mathcal{M})$. In most DI-CKA protocols, instead of using the statistics of the full correlation, we use the Bell violation and the QBER to test the level of security of the observed statistics. In this way, for practical reasons, the protocols coarse grain the statistics and we only use partial information of the full statistics to extract the device-independent key. In this context, the notation $(\sigma, \mathcal{N}) = (\rho, \mathcal{M})$ also implies that $\omega(\sigma, \mathcal{N}) = \omega(\rho, \mathcal{M})$ and $P_{\text{err}}(\sigma, \mathcal{N}) = P_{\text{err}}(\rho, \mathcal{M})$. When conditional probabilities associated with (ρ, \mathcal{M}) and (σ, \mathcal{N}) are ε -close to each other, then we write $(\rho, \mathcal{M}) \approx_\varepsilon (\sigma, \mathcal{N})$. For our purpose, it suffices to consider the distance

$$d(p, p') = \sup_{\mathbf{x}} \|p(\cdot|\mathbf{x}) - p'(\cdot|\mathbf{x})\|_1 \leq \varepsilon. \quad (4)$$

The device-independent distillable key rate of a device is informally defined as the supremum of the finite key rates κ achievable by the best protocol on any device compatible with (ρ, \mathcal{M}) , within an appropriate asymptotic block length limit and security parameter. Another approach taken is to

minimize the key rate of the statistics compatible with the Bell parameter and a QBER (see, e.g., [22]). For our purpose, we constrain ourselves to the situation when the compatible devices are supposedly independent and identically distributed. This constraint is because, as noted in Ref. [16], the upper bound on the key in the independent and identically distributed scenario is automatically the upper bound on the device-independent conference key in the general scenario since the independent and identically distributed attack is just one of the possible attacks in the general device-independent scenario.

An ideal conference key state $\tau^{(K)}$, with $\log_2 K$ secret key bits for N allies, is

$$\tau_{N(A)E}^{(K)} := \frac{1}{K} \sum_{k=0}^{K-1} |k\rangle\langle k|_{A_1} \otimes |k\rangle\langle k|_{A_2} \otimes \cdots \otimes |k\rangle\langle k|_{A_N} \otimes \sigma_E, \quad (5)$$

where σ_E is a state of the only system E accessible to an adversary, i.e., the adversary is uncorrelated with trusted users and gets no information about their secret bits. Consider the relations

$$(\rho, \mathcal{M}) \approx_\varepsilon (\sigma, \mathcal{N}), \quad (6)$$

$$\omega(\rho, \mathcal{M}) \approx_\varepsilon \omega(\sigma, \mathcal{N}), \quad (7)$$

$$P_{\text{err}}(\rho, \mathcal{M}) \approx_\varepsilon P_{\text{err}}(\sigma, \mathcal{M}), \quad (8)$$

where Eq. (6) implies Eqs. (7) and (8). Formally, the definition of device-independent quantum key distillation rate in the independent and identically distributed scenario is given as follows.

Definition 1 (cf. [16]). The maximum device-independent quantum key distillation rate of a device (ρ, \mathcal{M}) with independent and identically distributed behavior is defined as

$$K_{\text{DI,dev}}^{\text{iid}}(\rho, \mathcal{M}) := \inf_{\varepsilon > 0} \limsup_{n \rightarrow \infty} \sup_{\hat{\mathcal{P}}} \inf_{(6)} \kappa_n^\varepsilon(\hat{\mathcal{P}}((\sigma, \mathcal{N})^{\otimes n})), \quad (9)$$

where κ_n^ε is the quantum key rate achieved for any security parameter ε , block length or number of copies n , and measurements \mathcal{N} . Here $\hat{\mathcal{P}}$ is a protocol composed of classical local operations and public (classical) communication (CLOPC) acting on n identical copies of (σ, \mathcal{N}) which, composed with the measurement, results in a quantum local operations and public (classical) communication (QLOPC) protocol.

The following lemma follows from the definition of $K_{\text{DI}}^{\text{iid}}$ (generalizing statements from bipartite DI quantum key distillation in Refs. [16,18] to the DI-CKA).

Lemma 1. The maximum device-independent quantum key distillation rate $K_{\text{DI}}^{\text{iid}}$ of a device (ρ, \mathcal{M}) is equal to the maximum device-independent quantum key distillation rate of a device (σ, \mathcal{N}) when $(\rho, \mathcal{M}) = (\sigma, \mathcal{N})$:

$$(\rho, \mathcal{M}) = (\sigma, \mathcal{N}) \Rightarrow K_{\text{DI,dev}}^{\text{iid}}(\rho, \mathcal{M}) = K_{\text{DI,dev}}^{\text{iid}}(\sigma, \mathcal{N}). \quad (10)$$

The maximal DI-CKA rate $K_{\text{DI,dev}}(\rho, \mathcal{M})$ for the device (ρ, \mathcal{M}) is upper bounded by the maximal device-dependent conference key agreement (DD-CKA) rate $K_{\text{DD}}(\sigma)$ for all (σ, \mathcal{N}) such that $(\sigma, \mathcal{N}) = (\rho, \mathcal{M})$ (cf. [16]), i.e.,

$$K_{\text{DI,dev}}(\rho, \mathcal{M}) \leq \inf_{(\sigma, \mathcal{N})=(\rho, \mathcal{M})} K_{\text{DD}}(\sigma). \quad (11)$$

The device-dependent quantum key distillation rate $K_{\text{DD}}(\rho)$ (cf. [25]) is the maximum secret key (against quantum eavesdropper) that can be distilled between allies using local operations and classical communication (LOCC) (see, e.g., [21]),

$$K_{\text{DD}}(\rho) := \inf_{\varepsilon > 0} \lim_{n \rightarrow \infty} \sup_{\Lambda \in \text{QLOPC}} \left\{ \frac{\log_2 d_n}{n} \left| \Lambda(\rho^{\otimes n}) \approx_\varepsilon \tau^{(d_n)} \right. \right\}, \quad (12)$$

where $\rho \approx_\varepsilon \sigma \iff \frac{1}{2} \|\rho - \sigma\|_1 \leq \varepsilon$.

Corollary 1. For entanglement measures Ent which upper bound the maximum device-dependent key distillation rate, i.e., $K_{\text{DD}}(\rho) \leq \text{Ent}(\rho)$ for a density operator ρ , we have

$$K_{\text{DI,dev}}(\rho, \mathcal{M}) \leq \inf_{(\sigma, \mathcal{N})=(\rho, \mathcal{M})} K_{\text{DD}}(\sigma) \quad (13)$$

$$\leq \inf_{(\sigma, \mathcal{N})=(\rho, \mathcal{M})} \text{Ent}(\sigma). \quad (14)$$

Remark 1. We do not make any assumption about the dimension of the Hilbert space on which the state σ (Definitions 1–3) is defined as the systems can be finite dimensional or infinite dimensional. We only assume that the systems A_i accessible by the allies for key distillation upon measurement are finite dimensional in an honest setting. The systems E accessible to an adversary can be finite dimensional or infinite dimensional, depending on the cheating strategy (see Lemma 7, Appendix B).

As discussed above, a large class of device-independent quantum key distillation protocols relies on the Bell violation and the QBER of the device $p(\mathbf{a}|\mathbf{x})$. For such protocols, we can define the device-independent key distillation protocol as follows.

Definition 2 (cf. [22]). The maximal device-independent quantum key distillation rate of a device (ρ, \mathcal{M}) with independent and identically distributed behavior, Bell violation $\omega(\rho, \mathcal{M})$, and QBER $P_{\text{err}}(\rho, \mathcal{M})$ is defined as

$$K_{\text{DI,par}}^{\text{iid}}(\rho, \mathcal{M}) := \inf_{\varepsilon > 0} \limsup_{n \rightarrow \infty} \sup_{\hat{\mathcal{P}}} \inf_{(7),(8)} \kappa_n^\varepsilon(\hat{\mathcal{P}}((\sigma, \mathcal{N})^{\otimes n})). \quad (15)$$

Remark 2. As Eq. (6) implies Eqs. (7) and (8), it follows from the definitions of $K_{\text{DI}}^{\text{iid}}(\rho, \mathcal{M})$ and $K_{\text{DI,par}}^{\text{iid}}(\rho, \mathcal{M})$ that $K_{\text{DI,par}}^{\text{iid}}(\rho, \mathcal{M}) \leq K_{\text{DI,dev}}^{\text{iid}}(\rho, \mathcal{M})$.

Definition 3. The maximum device-independent quantum key distillation rate $K_{\text{DI}}^{\text{iid}}$ of a bipartite state ρ_{AB} is given by

$$K_{\text{DI,dev(par)}}^{\text{iid}}(\rho) = \sup_{\mathcal{M}} K_{\text{DI,dev(par)}}^{\text{iid}}(\rho, \mathcal{M}). \quad (16)$$

Remark 3 (cf. [11]). We note that there may exist states ρ for which $K_{\text{DI}}^{\text{iid}}(\rho) = 0$ but $K_{\text{DI}}^{\text{iid}}(\rho^{\otimes k}) > 0$ for some $k \in \mathbb{N}$.

III. REDUCED C-SQUASHED ENTANGLEMENT BOUND

In this section we generalize the notion of the cc-squashed entanglement [18,22] to the multipartite form. Next we prove that the properly scaled reduced c-squashed entanglement serves as an upper bound on the device-dependent conference key of the classical-quantum state. Furthermore, via Lemma 2 and Proposition 1 we prove that the reduced c-squashed entanglement is convex. This result may be further applied

to generate numeric upper bounds with the convexification technique [18,39]. Then we prove the main result of this section. Namely, we prove that the independent and identically distributed quantum device-independent conference key is upper bounded by the reduced c-squashed entanglement. Finally, we show that similar results hold when the honest parties broadcast the inputs to their devices so that the adversary can learn them.

In what follows, we first prove that the “measured” version of the multipartite squashed entanglement E_{sq}^q defined in Ref. [23], if properly scaled, upper bounds the conference key secure against the quantum adversary. Let us first recall facts and definitions. The multipartite conditional mutual information of a state $\rho_{A_1, \dots, A_N E}$ reads [40]

$$I(A_1 : \dots : A_N | E)_\rho = \sum_{i=1}^N S(A_i | E)_\rho - S(A_1, \dots, A_N | E)_\rho. \quad (17)$$

Here, the conditional entropy $S(A_i | E)_\rho = S(A_i E)_\rho - S(E)_\rho$, with $S(AB) := \text{Tr}[\rho_{AB} \log_2 \rho_{AB}]$ and $S(A) := \text{Tr}[\rho_A \log_2 \rho_A]$ being von Neumann entropies. The von Neumann entropy reduces to the Shannon entropy $H(X)$ for classical register X , $H(X) = -\sum_x p(x) \log_2 p(x)$, where $\{p(x)\}_x$ is the probability distribution associated with the random variable X . It will be crucial to note that the following identity holds [23]:

$$\begin{aligned} I(A_1 : \dots : A_N | E)_\rho &= I(A_1 : A_2 | E)_\rho + I(A_3 : A_1 A_2 | E)_\rho \\ &\quad + I(A_4 : A_1 A_2 A_3 | E)_\rho + \dots + I(A_N : A_1 \dots A_{N-1} | E)_\rho. \end{aligned} \quad (18)$$

Here $I(A : B | C)_\rho = S(AC)_\rho + S(BC)_\rho - S(C)_\rho - S(ABC)_\rho$ is the conditional mutual information.

Remark 4. Let Σ be any permutation of indices $1, \dots, N$. Then

$$\begin{aligned} I(A_1 : \dots : A_N | E)_\rho &= I(A_{\Sigma(1)} : \dots : A_{\Sigma(N)} | E)_\rho \\ &= I(A_{\Sigma(1)} : A_{\Sigma(2)} | E)_\rho + I(A_{\Sigma(3)} : A_{\Sigma(1)} A_{\Sigma(2)} | E)_\rho \\ &\quad + I(A_{\Sigma(4)} : A_{\Sigma(1)} A_{\Sigma(2)} A_{\Sigma(3)} | E)_\rho \\ &\quad + \dots + I(A_{\Sigma(N)} : A_{\Sigma(1)} \dots A_{\Sigma(N-1)} | E)_\rho. \end{aligned} \quad (19)$$

Further, the multipartite squashed entanglement of a quantum state ρ_{A_1, \dots, A_N} is defined as follows (Definition 3 of Ref. [23]).

Definition 4 (from [23]). For an N -partite state ρ_{A_1, \dots, A_N} ,

$$E_{\text{sq}}^q(\rho_{A_1, \dots, A_N}) := \inf_{\sigma} I(A_1 : A_2 : \dots : A_N | E)_\sigma, \quad (20)$$

where the infimum is taken over states $\sigma_{A_1, \dots, A_N E}$ that are extensions of ρ_{A_1, \dots, A_N} , i.e., $\text{Tr}_E[\sigma_{A_1, \dots, A_N E}] = \rho_{A_1, \dots, A_N}$.

We will need to generalize the notion of the cc-squashed entanglement [18,22] to the multipartite form.

Definition 5. A reduced c-squashed entanglement of a state ρ_{A_1, \dots, A_N} is defined as

$$E_{\text{sq}}^c(\rho, \mathbf{M}) := \inf_{\Lambda: E \rightarrow E'} I(A_1 : \dots : A_N | E')_{\mathbf{M}_{N(A)} \otimes \Lambda \psi_{N(A)E}^\rho}, \quad (21)$$

where $\mathbf{M}_{N(A)}$ is an N -tuple of positive-operator-valued measures (POVMs) M_{A_1}, \dots, M_{A_N} and state $|\psi_{N(A)E}^\rho\rangle$ is a purification of $\rho_{N(A)}$.

The first theorem comes with the following fact, which is a multipartite generalization of Theorem 5 from Ref. [18], where $N = 2$. Namely, the device-dependent conference key of the classical-quantum state is upper bounded by the properly scaled reduced c-squashed entanglement. We are ready to state a theorem that shows that the c-squashed entanglement, when properly scaled, upper bounds the device-dependent key.

Theorem 1. For an N -partite state $\rho_{N(A)}$, its purification ψ^ρ , and an N -tuple of POVMs $\mathbf{M}_{N(A)}$, there is

$$K_{\text{DD}}(\mathbf{M}_{N(A)} \otimes \text{id}_E \psi^\rho) \leq \frac{1}{N-1} E_{\text{sq}}^c(\rho_{N(A)}, \mathbf{M}_{N(A)}). \quad (22)$$

Proof. We closely follow the proof of Theorem 3.5 of Ref. [25], however, based not on Theorem 3.1 of Ref. [25], but its generalization to the case where system E need not be finite (see Lemma 7, Appendix B). Namely, any function which satisfies (i) monotonicity under LOPC, (ii) asymptotic continuity, (iii) normalization, and (iv) subadditivity is, after regularization, an upper bound on the distillable key secure against the quantum adversary (K_{DD}).

We first show the monotonicity. The LOPC consist of local operation and public communication. A local operation consists of adding a local ancilla, performing a unitary transformation, and a partial trace. It is easy to see that adding a local ancilla on one system does not alter this quantity. The same holds for the unitary transformation. The partial trace does not increase it as it can be rewritten in terms of the conditional mutual information terms as in Eq. (18). Then the same argument as in the proof of Theorem 3.5 of Ref. [25] [see Eq. (57) therein] applies.

Finally, for classical communication, we use the form given in Eq. (18) to verify the inequality stated below for the case when A_i produces locally the variable C_i and then broadcasts it to all the parties in the form of C_j for $j \neq i$ and to the adversary in the form of C_{N+1} (note that broadcasting followed by a partial trace, if needed, can simulate any classical communication among N parties):

$$I(A_1 : \dots : A_i C_i : \dots : A_N | E)_\rho \quad (23)$$

$$\stackrel{\text{(I)}}{=} I(A_i C_i : A_2 : \dots : A_{i-1} : A_1 : A_{i+1} : \dots : A_N | E)_\rho \quad (24)$$

$$\begin{aligned} &= I(A_i C_i : A_2 | E)_\rho + I(A_3 : A_i C_i A_2 | E)_\rho \\ &\quad + I(A_4 : A_i C_i A_2 A_3 | E)_\rho + \dots \\ &\quad + I(A_N : A_i C_i \dots A_{N-1} | E)_\rho \end{aligned} \quad (25)$$

$$\begin{aligned} &\stackrel{\text{(II)}}{\geq} I(A_i C_i : A_2 C_2 | EC_{N+1})_\rho \\ &\quad + I(A_3 C_3 : A_i C_i A_2 C_2 | EC_{N+1})_\rho \\ &\quad + I(A_4 C_4 : A_i C_i A_2 C_2 A_3 C_3 | EC_{N+1})_\rho + \dots \\ &\quad + I(A_N C_N : A_i C_i \dots A_{N-1} C_{N-1} | EC_{N+1})_\rho \end{aligned} \quad (26)$$

$$= I(A_1 C_1 : \dots : A_i C_i : \dots : A_N C_N | EC_{N+1})_\rho. \quad (27)$$

In Eq. (23) [step (I)], we transposed the labels i and 1 (see Remark 4), and step (II) [inequality (25)] follows (termwise) from the monotonicity of the (three-partite) mutual information function proved in Ref. [25] (see the proof of Theorem 3.5 therein).

Regarding asymptotic continuity, we consider two states $\rho_{N(A)}$ and $\sigma_{N(A)}$ such that $\|\rho_{N(A)} - \sigma_{N(A)}\|_1 \leq \epsilon$. Then, as in the proof of Theorem 3.5 of Ref. [25], for any map $\Lambda : E \rightarrow E'$ there is $\|\rho'_{N(A)} - \sigma'_{N(A)}\|_1 \leq \epsilon$, where $\rho'_{N(A)} := \text{id}_{N(A)} \otimes \Lambda(\rho_{N(A)})$ and $\sigma'_{N(A)} := \text{id}_{N(A)} \otimes \Lambda(\sigma_{N(A)})$. Then, by the expansion Eq. (18) we obtain

$$\begin{aligned} & |I(A_i : A_1 \dots A_{i-1} | E')_\rho - I(A_i : A_1 \dots A_{i-1} | E')_\sigma| \\ & \leq 2\epsilon \log_2 d_{A_i} + 2g(\epsilon), \end{aligned} \quad (28)$$

with $d_{A_i} := \dim(\mathcal{H}_{A_i})$ and $g(\epsilon) := (1 + \epsilon) \log_2(1 + \epsilon) - \epsilon \log_2 \epsilon$, where we use Lemma 5 from Appendix A, provided in Ref. [41]. Hence, in total we get

$$\begin{aligned} & |I(A_1 : \dots : A_N | E')_\rho - I(A_1 : \dots : A_N | E')_\sigma| \\ & \leq 2\epsilon \sum_{i=1}^{N-1} \log_2 d_{A_i} + (N-1)2g(\epsilon) \\ & \leq (N-1)[2\epsilon \max_{i \in \{1, \dots, N\}} \log_2 d_{A_i} + 2g(\epsilon)]. \end{aligned} \quad (29)$$

For a finite natural N , the right-hand side of the above approaches 0, with $\epsilon \rightarrow 0$.

The subadditivity follows again from the fact that we can split the term $I(A_1 : \dots : A_N | E)$ into $N-1$ terms of the form $I(A_i : A_1 \dots A_{i-1} | E)$. Further treating $A_1 \dots A_{i-1}$ together as B_i (equivalent of B in the proof of Theorem 3.5 in Ref. [25]), we can prove the additivity of the form

$$I(A_i A'_i : B_i B'_i | EE') = I(A_i : B_i | E) + I(A'_i : B'_i | E') \quad (30)$$

for each term and notice the subadditivity from the fact that in the infimum in the definition of $E_{\text{sq}}^c(\rho, \mathcal{M})$ there are product channels; hence in general the formula can be lower than the above.

Finally, we consider normalization. It is straightforward to see that, assuming $d_{A_i} = d_A$ for each $i \in \{1, \dots, N\}$, on the state representing the ideal key $\tau_{N(A)E} = \frac{1}{d_A} \sum_{i=0}^{d_A-1} |ii \dots i\rangle \langle ii \dots i| \otimes \rho_E$ [Eq. (5)] there is $E_{\text{sq}}^c(\tau, \mathbf{M}) = (N-1) \log_2 d_A$, by noticing that on the product state $I(A_i : A_1 \dots A_{i-1} | E) = I(A_i : A_1 \dots A_{i-1}) = \log_2 d_A$, and there are $N-1$ of such terms in the definition of E_{sq}^c . We assume here also that the measurement \mathbf{M} is generating the key in the computational basis.

We have further an analog of Observation 4 of Ref. [18]. Its proof goes along similar lines. Indeed, it does not depend on either the type of objective function that is minimized or the number of parties; hence we omit it here.

Observation 1. For an N -partite state $\rho_{N(A)}$ and a POVM $\mathbf{M}_{N(A)} = \mathbf{M}_{A_1}, \dots, \mathbf{M}_{A_N}$ there is

$$E_{\text{sq}}^c(\rho, \mathbf{M}) = \inf_{\rho_{N(A)E} \in \text{Ext}(\rho_{N(A)})} I(A_1 : \dots : A_N | E)_{\mathbf{M}_{N(A)} \otimes \mathbb{1}_E(\rho_{N(A)E})}, \quad (31)$$

where $\text{Ext}(\rho_{N(A)})$ stands for the state extension of $\rho_{N(A)}$, i.e., $\rho_{N(A)E}$ is a density operator such that $\text{Tr}_E[\rho_{N(A)E}] = \rho_{N(A)}$.

Owing to Observation 1, we can obtain the analog of Lemma 6 of Ref. [18], which states that E_{sq}^c is convex.

Lemma 2. For a tuple of POMVs $\mathbf{M}_{N(A)}$, two states $\rho_{N(A)}^{(1)}$ and $\rho_{N(A)}^{(2)}$, and $0 < p < 1$, there is

$$E_{\text{sq}}^c(\bar{\rho}_{N(A)}) \leq p E_{\text{sq}}^c(\rho_{N(A)}^{(1)}) + (1-p) E_{\text{sq}}^c(\rho_{N(A)}^{(2)}), \quad (32)$$

where $\bar{\rho}_{N(A)} = p\rho_{N(A)}^{(1)} + (1-p)\rho_{N(A)}^{(2)}$.

Proof. The proof is due to the fact that the function $E_{\text{sq}}^c(\bar{\rho}_{N(A)})$ is upper bounded by $I(A_1 : \dots : A_N | EF)$ evaluated on a state $\rho_{ABEF} = \mathbf{M}_{N(A)} \otimes \text{id}_E(p\rho^{(1)} \otimes |0\rangle\langle 0|_F + (1-p)\rho^{(2)} \otimes |1\rangle\langle 1|_F)$. Further, by Eq. (18) there is

$$\begin{aligned} & I(A_1 : \dots : A_N | EF)_\rho \\ & = p I(A_1 : \dots : A_N | E)_{\mathbf{M}_{N(A)} \otimes \text{id}_E \rho_{N(A)}^{(1)}} \\ & \quad + (1-p) I(A_1 : \dots : A_N | E)_{\mathbf{M}_{N(A)} \otimes \text{id}_E \rho_{N(A)}^{(2)}}. \end{aligned} \quad (33)$$

Since the states $\rho^{(1)}$ and $\rho^{(2)}$ were arbitrary, we get the thesis.

We further note that switching from a bipartite key distillation task to the conference key distillation does not alter the formulation or the proof of Lemma 7 of Ref. [18]. We state it below for the sake of the completeness of the further proofs.

Lemma 3. The independent and identically distributed quantum device-independent key achieved by protocols using (for generating the key) a single tuple of measurements $(\hat{x}_1, \dots, \hat{x}_N) \equiv \hat{\mathbf{x}}$ applied to \mathcal{M} of a device $(\rho_{N(A)}, \mathcal{M})$ is upper bounded as

$$\begin{aligned} & K_{\text{DI,dev}}^{\text{iid},(\hat{\mathbf{x}})}(\rho_{N(A)}, \mathcal{M}) \\ & := \inf_{\epsilon > 0} \limsup_n \sup_{\mathcal{P} \in \text{LOPC}(\sigma_{N(A)}, \mathcal{L}) \approx_\epsilon (\rho_{N(A)}, \mathcal{M})} \inf_{\mathcal{L}} \kappa_n^{\epsilon,(\hat{\mathbf{x}})}(\hat{\mathcal{P}}(\mathbf{L}(\sigma)^{\otimes n})) \\ & \leq \inf_{(\sigma_{N(A)}, \mathcal{L}) = (\rho_{N(A)}, \mathcal{M})} K_{\text{DD}}(\mathcal{L}(\hat{\mathbf{x}}) \otimes \text{id}_E(\psi^\sigma)), \end{aligned} \quad (34)$$

where $\mathbf{L} \equiv \mathcal{L}(\hat{\mathbf{x}})$ is a single pair of measurements induced by inputs $(\hat{\mathbf{x}})$ on \mathcal{L} and $\kappa_n^{\epsilon,(\hat{\mathbf{x}})}$ is the rate of the ϵ -perfect conference key achieved and classical labels from local classical operations in $\hat{\mathcal{P}} \in \text{CLOPC}$ are possessed by the allies holding systems A_i for $i \in \{1, \dots, N\}$.

Combining Theorem 1 with Lemma 3, we obtain the main result of this section. This is a bound by the reduced reduced c-squashed entanglement.

Theorem 2. The independent and identically distributed quantum device-independent conference key achieved by protocols using a single tuple of measurements $(\hat{x}_1, \dots, \hat{x}_N) \equiv \hat{\mathbf{x}}$ applied to \mathcal{M} of a device $(\rho_{N(A)}, \mathcal{M})$ is upper bounded as

$$\begin{aligned} & K_{\text{DI,dev}}^{\text{iid},\hat{\mathbf{x}}}(\rho_{N(A)}, \mathcal{M}) \\ & \leq \frac{1}{N-1} \inf_{(\sigma_{N(A)}, \mathcal{N}) = (\rho_{N(A)}, \mathcal{M})} E_{\text{sq}}^c(\sigma_{N(A)}, \mathcal{L}(\hat{\mathbf{x}})) \end{aligned} \quad (35)$$

$$=: E_{\text{sq,dev}}^c(\rho_{N(A)}, \mathcal{M}(\hat{\mathbf{x}})). \quad (36)$$

We have an analogous result for a key which is a function only of the tested parameters, that of Bell inequality violation and the quantum bit error rate.

Theorem 3. The independent and identically distributed quantum device-independent key achieved by protocols using (for generating the key) a single tuple of measurements $\hat{\mathbf{x}}$

applied to \mathcal{M} of a device $(\rho_{N(A)}, \mathcal{M})$ is upper bounded as

$$\begin{aligned} & K_{\text{DI,par}}^{\text{id},\hat{\mathbf{x}}}(\rho_{N(A)}, \mathcal{M}) \\ & := \inf_{\epsilon > 0} \limsup_n \sup_{\mathcal{P} \in \text{LOPC}} \sup_{\omega(\sigma_{N(A)}, \mathcal{L}) \approx_{\epsilon} \omega(\rho_{N(A)}, \mathcal{M})} \inf_{P_{\text{err}(\sigma_{N(A)}, \mathcal{L}) \approx_{\epsilon} P_{\text{err}(\rho_{N(A)}, \mathcal{M})}} \kappa_n^{\epsilon, \hat{\mathbf{x}}}(\mathcal{P}(\mathbf{L}(\sigma)^{\otimes n})) \\ & \leq \frac{1}{N-1} \inf_{\omega(\sigma_{N(A)}, \mathcal{L}) = \omega(\rho_{N(A)}, \mathcal{M})} \inf_{P_{\text{err}(\sigma, \mathcal{L}) = P_{\text{err}(\rho, \mathcal{M})}} E_{\text{sq}}^c(\sigma_{N(A)}, \mathbf{L}) \\ & =: E_{\text{sq,par}}^c(\rho_{N(A)}, \mathcal{M}(\hat{\mathbf{x}})), \end{aligned} \quad (37)$$

where $\mathbf{L} = \mathcal{L}(\hat{\mathbf{x}})$ is a single tuple of measurements induced by inputs $(\hat{\mathbf{x}})$ on \mathcal{L} .

For $N = 2$, the above bound recovers the result of Ref. [18].

In the definition of $E_{\text{sq,par}(\text{dev})}^c$, one can take the infimum only over the classical extensions to Eve [23]. In that case, for a single input $\hat{\mathbf{x}}$ this bound reads $\frac{1}{N-1} I(N(A) \downarrow E)$, as given in Ref. [23] (see Ref. [25] for the bipartite case). We have the following immediate corollary.

Corollary 2. The independent and identically distributed quantum device-independent key achieved by protocols using a tuple of measurements $\hat{\mathbf{x}}$ applied to a device $(\rho_{N(A)}, \mathcal{M})$ is upper bounded as

$$\begin{aligned} & K_{\text{DI,dev}}^{\text{id},\hat{\mathbf{x}}}(\rho_{N(A)}, \mathcal{M}) \\ & \leq \inf_{(\sigma_{N(A)}, \mathcal{L}) = (\rho_{N(A)}, \mathcal{M})} \frac{1}{N-1} I(N(A) \downarrow E)_{P(A_1 \dots A_N | E)} \quad (39) \\ & \equiv \inf_{(\sigma_{N(A)}, \mathcal{L}) = (\rho_{N(A)}, \mathcal{M})} \inf_{\Lambda_{E \rightarrow F}} \frac{1}{N-1} I(N(A) | F)_{P(A_1 \dots A_N | \Lambda(E))}, \end{aligned} \quad (40)$$

where $P(A_1 : \dots : A_N | E)$ is a distribution coming from measurement $\mathcal{L}(\hat{\mathbf{x}})$ on purification of $\sigma_{N(A)}$ to system E and the infimum is taken over classical channels transforming a random variable E to a random variable F .

We will exemplify Corollary 2 for $N = 3$ parties and the scenario considered in Ref. [20]. For the results, see Fig. 2. Let us also note that when one restricts the infimum in Eq. (40), the channel $\Lambda : E \rightarrow F$ has only a classical output and the above bound is a multipartite generalization of the intrinsic

information bound given in Ref. [17]. An analogous corollary holds for the case of $K_{\text{DI,par}}^{\text{id},\hat{\mathbf{x}}}$.

We finally note that $E_{\text{sq,par}}^c$ is convex also, in the multipartite case. This may prove important when one finds upper bounds, as any convexification of two plots obtained from optimization of $E_{\text{sq,par}}^c$ is then an upper bound on $K_{\text{DI,par}}^c$, as it was used in Ref. [18]. We state it below following Lemma 8 of Ref. [18].

Proposition 1. The $E_{\text{sq,par}}^c$ is convex, i.e., for every device $(\bar{\rho}, \mathcal{M})$ and an input tuple $\hat{\mathbf{x}}$ there is

$$\begin{aligned} & E_{\text{sq,par}}^c(\bar{\rho}, \mathcal{M}(\hat{\mathbf{x}})) \\ & \leq p_1 E_{\text{sq,par}}^c(\rho_1, \mathcal{M}(\hat{\mathbf{x}})) + p_2 E_{\text{sq,par}}^c(\rho_2, \mathcal{M}(\hat{\mathbf{x}})), \end{aligned} \quad (41)$$

where $\bar{\rho} = p_1 \rho_1 + p_2 \rho_2$ and $p_1 + p_2 = 1$ with $0 \leq p_1 \leq 1$.

Proof. The proof goes the same way as that for the bipartite case of Lemma 8 in Ref. [18], with the only change that we base it on the convexity of its multipartite version E_{sq}^c , i.e., Lemma 2 here, and the fact that

$$\begin{aligned} & I(A_1 A'_1 : \dots : A_N A'_N | E) [\rho_{A_1 \dots A_N} \otimes |i \dots i\rangle \langle i \dots i|_{A'_1 \dots A'_N}] \\ & = I(A_1 : \dots : A_N | E) [\rho_{A_1 \dots A_N}], \end{aligned} \quad (42)$$

where $i \in \{0, 1\}$, ρ_{A_1, \dots, A_N} is arbitrary state of systems $A_1 \dots A_N$, and we define $I(A_1 : \dots : A_N | E) [\rho] \equiv I(A_1 : \dots : A_N | E)_{\rho}$. This is because a pure product state alters neither the entropy of marginals nor the global entropy of the state.

We note that the multipartite function E_{sq}^c can be defined for multiple measurements as in Ref. [18] and the analogous results (e.g., Corollary 6 of Ref. [18]) to the bipartite case would hold for the multipartite case.

Definition 6. The reduced c-squashed entanglement of the collection of measurements \mathcal{M} with probability distribution $p(\mathbf{x})$ of the input reads

$$E_{\text{sq}}^c(\rho_{N(A)}, \mathcal{M}, p(\mathbf{x})) := \sum_{\mathbf{x}} p(\mathbf{x}) E_{\text{sq}}^c(\rho_{N(A)}, \mathbf{M}_{\mathbf{x}}). \quad (43)$$

Usually, the parties broadcast their inputs used to generate the key during the protocol. One can therefore consider a version of the distillable device-independent key achieved by such protocols which do this broadcasting. We then consider the quantum device-independent key rate

$$\kappa_{\text{DI,dev}}^{\text{id,broad}}(\rho, \mathcal{M}, p(\mathbf{x})) := \inf_{\epsilon > 0} \limsup_n \sup_{\hat{\mathcal{P}} \in \text{LOPC}} \inf_{(\sigma, \mathcal{N}) \approx_{\epsilon} (\rho, \mathcal{M})} \kappa_n^{\epsilon} \left(\hat{\mathcal{P}} \left(\left(\sum_{\mathbf{x}} p(\mathbf{x}) \mathbf{N}_{\mathbf{x}} \otimes \text{id}_E(|\psi_{\sigma}\rangle \langle \psi_{\sigma}| \otimes |\mathbf{x}\rangle \langle \mathbf{x}|_{E_{\mathbf{x}}}) \right)^{\otimes n} \right) \right), \quad (44)$$

where by broad we mean that $\mathbf{x} := (x_1, \dots, x_N)$ are broadcasted and we make it explicit by adding classical registers $E_{\mathbf{x}} := E_{x_1}, \dots, E_{x_N}$ held by Eve. We have then a generalization of Theorem of 2 to the case of more measurements that are revealed during the protocol of key distillation.

Proposition 2. The independent and identically distributed quantum device-independent key achieved by protocols using measurements of a device $(\rho_{N(A)}, \mathcal{M})$ with probability $p(\mathbf{x})$ is upper bounded as

$$K_{\text{DI,dev}}^{\text{id,broad}}(\rho, \mathcal{M}, p(\mathbf{x})) \equiv \inf_{\epsilon > 0} \limsup_n \sup_{\hat{\mathcal{P}} \in \text{LOPC}} \inf_{(\sigma, \mathcal{N}) \approx_{\epsilon} (\rho, \mathcal{M})} \kappa_n^{\epsilon} \left(\hat{\mathcal{P}} \left(\left(\sum_{\mathbf{x}} p(\mathbf{x}) \mathbf{N}_{\mathbf{x}} \otimes \text{id}_E(|\psi_{\sigma}\rangle \langle \psi_{\sigma}| \otimes |\mathbf{x}\rangle \langle \mathbf{x}|_{E_{\mathbf{x}}}) \right)^{\otimes n} \right) \right) \quad (45)$$

$$\leq \frac{1}{N-1} \inf_{(\sigma, \mathcal{N}) = (\rho, \mathcal{M})} E_{\text{sq}}^c(\sigma, \mathcal{N}, p(\mathbf{x})) \quad (46)$$

$$=: E_{\text{sq,dev}}^c(\rho, \mathcal{M}, p(\mathbf{x})), \quad (47)$$

where $\mathbf{N}_{\mathbf{x}}$ are measurements induced by \mathbf{x} on \mathcal{N} .

Proof. The proof follows straightforwardly from generalization of Lemma 10 and Theorem 10 from Ref. [18] for the case of E_{sq}^c , taking as the argument the measurements as in Eq. (43), composed with a broadcast map which for the choice of inputs \mathbf{x} creates systems E_x in state $|\mathbf{x}\rangle\langle\mathbf{x}|$.

IV. DUAL BOUND

In this section, we develop a bound analogous to that given in Theorems 2 and 3, however based on a different multipartite squashed entanglement, denoted in Ref. [24] by \tilde{E}_{sq} . It originates from a quantifier of correlations denoted by S_N in Ref. [23], which is one of the so-called secrecy monotones introduced in Ref. [26]. Although the E_{sq} and \tilde{E}_{sq} are shown in Ref. [24] to be equal on pure states, they may lead to incomparable upper bounds for mixed states. The bounds derived from these quantities are compared in the case $N = 3$ in Fig. 2.

Let us first recall the definition of S_N . In what follows we will denote it by D_N so that it is not confused with the von Neumann entropy:

$$\begin{aligned} D_N(\rho_{N(A)E}) & \\ := I(A_1 : A_2 \cdots A_N | E)_{\rho_{N(A)E}} &+ I(A_2 : A_3 \cdots A_N | A_1 E)_{\rho_{N(A)E}} \\ &+ I(A_3 : A_4 \cdots A_N | A_1 A_2 E)_{\rho_{N(A)E}} + \cdots \\ &+ I(A_{N-1} : A_N | A_1 \cdots A_{N-2} E)_{\rho_{N(A)E}}. \end{aligned} \quad (48)$$

We similarly define $D_N(\rho_{N(A)})$ as in the above, omitting conditioning on the system E . Based on this quantity, one defines an entanglement measure, which in Ref. [24] (see also [23]) reads as follows.

Definition 7. For an N -partite state ρ_{A_1, \dots, A_N} ,

$$\tilde{E}_{\text{sq}}(\rho_{N(A)}) := \inf D_N(\sigma_{N(A)E}), \quad (49)$$

where the infimum is taken over states $\sigma_{N(A)E}$ that are extensions of $\rho_{N(A)}$, i.e., $\text{Tr}_E[\sigma_{N(A)E}] = \rho_{N(A)}$.

We will refer to it as dual squashed entanglement due to the fact noted in Ref. [23] that for any N -partite state ρ the two squashed entanglement functions are dual:

$$\begin{aligned} D_N(\rho) + I(A_1 : \dots : A_N)_\rho & \\ = \sum_{i=1}^N I(A_i : A_1 \dots A_{i-1} A_{i+1} \dots A_N)_\rho. \end{aligned} \quad (50)$$

Based on \tilde{E}_{sq} , we define a quantity, called the dual c-squashed entanglement. It is a function of a pair: state $\rho_{N(A)}$ of N parties and an N -partite measurement $\mathbf{M}_{N(A)}$. The conditional mutual information in its definition is computed on the extension of the state $\rho_{N(A)}$, measured by $\mathbf{M}_{N(A)}$. When proper we will denote $D_N(\rho)$ by $D(N(A))_\rho$. Similarly, $D(N(A)|E)_\rho$ will denote $D_N(\rho_{N(A)E})$ when convenient. In cases when $\rho_{N(A)E}$ forms a classical probability distribution, we will denote this function also by $D_N(N(A)|E)_{P(A_1: \dots: A_N|E)}$.

Definition 8. A dual c-squashed entanglement of a state ρ_{A_1, \dots, A_N} is defined as

$$\tilde{E}_{\text{sq}}^c(\rho, \mathbf{M}) := \inf_{\Lambda: E \rightarrow E'} D(N(A)|E')_{\mathbf{M}_{N(A)} \otimes \Delta \Psi_{N(A)E}^\rho}, \quad (51)$$

where $\mathbf{M}_{N(A)}$ is an N -tuple of POVMs M_{A_1}, \dots, M_{A_N} and the state $|\Psi_{N(A)E}^\rho\rangle$ is a purification of $\rho_{N(A)}$.

We note that a different formulation of the dual c-squashed entanglement is in analogy to Observation 1.

Observation 2. For an N -partite state $\rho_{N(A)}$ and a POVM $\mathbf{M}_{N(A)} = M_{A_1}, \dots, M_{A_N}$ there is

$$\tilde{E}_{\text{sq}}^c(\rho, \mathbf{M}) = \inf_{\rho_{N(A)E} = \text{Ext}(\rho_{N(A)})} D(N(A)|E)_{\mathbf{M}_{N(A)} \otimes \mathbb{1}_E \rho_{N(A)E}}. \quad (52)$$

Proof. The proof is analogous to the proof of Observation 4 in [18].

Based on Theorem 1, we can have an analogous fact, claimed in Refs. [23,24]. We state the sketch of the proof for the sake of completeness of the presentation.

Theorem 4. For an N -partite state $\rho_{N(A)}$, its purification ψ^ρ , and an N -tuple of POVMs $\mathbf{M}_{N(A)}$ there is

$$K_{\text{DD}}(\mathbf{M}_{N(A)} \otimes \text{id}_E \psi^\rho) \leq \tilde{E}_{\text{sq}}^c(\rho, \mathbf{M}). \quad (53)$$

Proof. In this proof we follow the argument made for E_{sq}^c . We will prove several properties of \tilde{E}_{sq}^c , which ensure the upper bound for the key due to a multipartite version of Theorem 3.1 of Ref. [25] (see Lemma 7 in Appendix B).

To see the monotonicity of \tilde{E}_{sq}^c under LOCC, we first note the fact that this quantity does not increase under local operations (see Proposition 2 of Ref. [23]). Regarding public communication, we note, as in the proof of Theorem 1, that first when A_N is broadcasting information C_N to all the parties $A_{i \neq N}$ and E there is

$$\begin{aligned} I(A_1 : A_2 \cdots A_N C_N | E) + I(A_2 : A_3 \cdots A_N C_N | A_1 E) + \cdots \\ + I(A_{N-1} : A_N C_N | A_1 \cdots A_{N-2} E) \\ \geq I(A_1 C_1 : A_2 C_2 \cdots A_N C_N | E C_{N+1}) \\ + I(A_2 C_2 : A_3 C_3 \cdots A_N C_N | A_1 C_1 E C_{N+1}) + \cdots \\ + I(A_{N-1} C_{N-1} : A_N C_N | A_1 C_1 \cdots A_{N-2} C_{N-2} E C_{N+1}) \end{aligned} \quad (54)$$

termwise: $I(A_i : A_{i-1} \cdots A_N C_N | A_1 \cdots A_{i-2} E) \geq I(A_i C_i : A_{i-1} C_{i-1} \cdots A_N C_N | A_1 C_1 \cdots A_{i-2} C_{i-2} E C_{N+1})$. Indeed, this inequality coincides with the one proved in Ref. [25] (see the proof of Theorem 3.5 therein), where we identify A_i with B , $A_{i-1} \cdots A_N$ with A , and E with E' . For the case when another party $A_{i \neq N}$ is broadcasting, we swap (by symmetry of the formula of D_N) the A_i with A_N and follow the same argument as explained above.

Asymptotic continuity of the \tilde{E}_{sq}^c follows from the fact that it consists of conditional information terms, each of which is asymptotically continuous due to Lemma 5 (note that we allow system E to be infinite dimensional).

The argument for the subadditivity of this function follows from the fact that D_N is a product of tensor products (see Ref. [23] for the case of $N = 3$).

Regarding normalization, for the state containing an ideal key $\tau_{N(A)E}$ [Eq. (5)] there is $\tilde{E}_{\text{sq}}(\tau, \mathbf{M}) = \log_2 d_A$, where we assume $\log_2 d_i = \log_2 d_A$ for all $i \in \{1, \dots, N\}$. Indeed, as it is argued in Ref. [23], the first term in (48) equals $\log_2 d_A$, while the remaining ones are equal to 0.

Combining Lemma 3 and Theorem 4, we obtain an analog of Theorem 2.

Theorem 5. The independent and identically distributed quantum device-independent conference key achieved by protocols using a single tuple of measurements $(\hat{x}_1, \dots, \hat{x}_N) \equiv \hat{\mathbf{x}}$

applied to \mathcal{M} of a device $(\rho_{N(A)}, \mathcal{M})$ is upper bounded as

$$K_{\text{DI,dev}}^{\text{iid},\hat{\mathbf{x}}}(\rho_{N(A)}, \mathcal{M}) \leq \inf_{(\sigma_{N(A)}, \mathcal{N}) \equiv (\rho_{N(A)}, \mathcal{M})} \tilde{E}_{\text{sq}}^c(\sigma_{N(A)}, \mathcal{L}(\hat{\mathbf{x}})) \quad (55)$$

$$=: \tilde{E}_{\text{sq,dev}}^c(\rho_{N(A)}, \mathcal{M}(\hat{\mathbf{x}})). \quad (56)$$

We also note that a bound analogous to the above holds for $K_{\text{DI,par}}^{\text{iid},\hat{\mathbf{x}}}$. Due to Theorem 5, we get an analog of Corollary 2.

Corollary 3. The independent and identically distributed quantum device-independent key achieved by protocols using a tuple of measurements $\hat{\mathbf{x}}$ applied to a device $(\rho_{N(A)}, \mathcal{M})$ is upper bounded as

$$K_{\text{DI,dev}}^{\text{iid},\hat{\mathbf{x}}}(\rho_{N(A)}, \mathcal{M}) \leq \inf_{(\sigma_{N(A)}, \mathcal{L}) \equiv (\rho_{N(A)}, \mathcal{M})} D_N(N(A) \downarrow E)_{P(A_1:\dots:A_N|E)} \equiv \inf_{(\sigma_{N(A)}, \mathcal{L}) \equiv (\rho_{N(A)}, \mathcal{M})} \inf_{\Lambda_{E \rightarrow F}} D_N(N(A)|F)_{P(A_1:\dots:A_N|\Lambda(E))}, \quad (57)$$

where $P(A_1 : \dots : A_N|E)$ is a distribution coming from measurement $\mathcal{L}(\hat{\mathbf{x}})$ on purification of $\sigma_{N(A)}$ to system E and the infimum is taken over classical channels transforming a random variable E to a random variable F .

V. BOUND ON THE RATE OF A PARITY CHSH-BASED PROTOCOL BY THE REDUCED c-SQUASHED ENTANGLEMENT

In this section we consider the scenario of $N = 3$ parties and compare the known lower bound on the conference key rate [20] with the upper bounds introduced in previous sections.

Below we exemplify the use of the bound by the reduced c-squashed entanglement E_{sq}^c in the case with classical Eve, that is, when the infimum in its definition runs over the extensions of the form $\sum_i p_i \rho_{A(N)}^i \otimes |i\rangle\langle i|_E$ (or equivalently the channels acting on system E have only classical outputs). We restrict ourselves to the standard protocols with a single pair of inputs generating the key [17]. We exemplify the bound given in Corollary 2 by means of $I(N(A) \downarrow E)$. It then is in essence a matter of checking the value of the multipartite intrinsic information measure of a distribution which is the output of a key-generating measurement on the attacking state (as it is done in the bipartite case in Ref. [17]). We also provide the dual upper bound originating from Corollary 3.

To compare the introduced upper bounds with the known lower bound, for the honest implementation, we focus on the GHZ state, on which depolarizing noise acts locally on three qubits [20]. Having this state and playing a tripartite game on it [11], called the parity Clauser-Horne-Shimony-Holt (CHSH) game, one can obtain (in the low-noise regime) a secure conference key. More precisely, we have the following.

Definition 9 (parity CHSH game [20]). The parity CHSH inequality extends the CHSH inequality to N parties as follows. Let Alice and Bob₁, ..., Bob_{N-1} be the N players of the following game (the parity CHSH game). Alice and Bob₁ are asked uniformly random binary questions $x \in \{0, 1\}$ and $y \in \{0, 1\}$, respectively. The other Bobs are each asked a fixed question, e.g., always equal to 1. Alice will answer bit a , and for all $t \in \{1, \dots, N - 1\}$, Bob_t answers bit b_t . We

denote by $\bar{b} := \bigotimes_{2 \leq i \leq N-1} b_i$ the parity of all the answers of Bob₂, ..., Bob_{N-1}. The players win if and only if

$$a + b_1 = x(y + \bar{b}) \pmod{2}.$$

As for the CHSH inequality, classical strategies for the parity CHSH game must satisfy

$$P_{\text{win}}^{\text{parity CHSH}} \leq \frac{3}{4}. \quad (58)$$

The above inequality can be violated with the Φ_3^{GHZ} state, with the maximal (quantum) value of $\frac{1}{2} + \frac{1}{2\sqrt{2}}$.

We adopt the same model of noise as in Ref. [20], which is represented by qubit depolarizing channels acting the same way on each qubit of the GHZ state:

$$\mathcal{D}_\nu(\rho) = (1 - \nu)\rho + \nu \frac{\mathbb{1}}{2}. \quad (59)$$

Below we explain the result of applying this global channel to the GHZ state $|\Phi_3^{\text{GHZ}}\rangle\langle\Phi_3^{\text{GHZ}}|$ in the case of $N = 3$.

Observation 3. The GHZ state after the action of depolarizing noise on each qubit reads

$$\begin{aligned} \mathcal{D}_\nu \otimes \mathbb{1}_{B_1 B_2} (|\Phi_3^{\text{GHZ}}\rangle\langle\Phi_3^{\text{GHZ}}|_{AB_1 B_2}) \\ = (1 - \nu) |\Phi_3^{\text{GHZ}}\rangle\langle\Phi_3^{\text{GHZ}}|_{AB_1 B_2} + \nu \frac{\mathbb{1}_A}{2} \otimes \kappa_{B_1 B_2}, \end{aligned} \quad (60)$$

where the $\kappa_{B_1 B_2} = \frac{1}{2}(|00\rangle\langle 00| + |11\rangle\langle 11|)$ state is separable.

Remark 5. The fully separable state originating from a depolarizing channel (single party), i.e., $\frac{\mathbb{1}_A}{2} \otimes \kappa_{B_1 B_2}$, cannot violate the parity CHSH inequality.

After applications of the depolarizing channel to each of three qubits we obtain the following.

Corollary 4. We have

$$\begin{aligned} \mathcal{D}_\nu^{\otimes 3} (|\Phi_3^{\text{GHZ}}\rangle\langle\Phi_3^{\text{GHZ}}|_{AB_1 B_2}) \\ = (1 - \nu)^3 |\Phi_3^{\text{GHZ}}\rangle\langle\Phi_3^{\text{GHZ}}|_{AB_1 B_2} + [1 - (1 - \nu)^3] \chi_\nu, \end{aligned} \quad (61)$$

where χ_ν is a fully separable state which reads

$$\begin{aligned} \chi_\nu := \frac{1}{1 - (1 - \nu)^3} \\ \times \left((1 - \nu)^2 \nu \kappa_{AB_1} \otimes \frac{\mathbb{1}_{B_2}}{2} + (1 - \nu)^2 \nu \kappa_{AB_2} \otimes \frac{\mathbb{1}_{B_1}}{2} \right. \\ \left. + (1 - \nu)^2 \nu \kappa_{B_1 B_2} \otimes \frac{\mathbb{1}_A}{2} + (3 - 2\nu)\nu^2 \frac{\mathbb{1}_{AB_1 B_2}}{2} \right). \end{aligned} \quad (62)$$

In Ref. [20], the expected winning probability for the parity CHSH game (with respect to the depolarizing noise parameter) is calculated:

$$P_{\text{exp}} := \left[\frac{1}{2} + \frac{(1 - \nu)^N}{2\sqrt{2}} + \frac{(1 - \nu)^2(1 - (1 - \nu)^{N-2})}{8\sqrt{2}} \right]. \quad (63)$$

From the above equality for $N = 3$, the state in Eq. (61) violates the classical bound of $\frac{3}{4}$ for $0 \leq \nu < \nu_{\text{crit}}$, where $\nu_{\text{crit}} \approx 0.1189$.

In this place, we start the construction of the eavesdropper strategy. According to the DI-CKA protocol in Ref. [20], the ranges of inputs and outputs are $x \in \{0, 1\}$, $y_1 \in \{0, 1, 2\}$, $y_2 \in \{0, 1\}$, and $a, b_1, b_2 \in \{0, 1\}$. The setting $(x, y_1, y_2) = (0, 2, 0)$

associated with measurements of σ_z observable is the key-generating round:

$$P_\nu(a, b_1, b_2|x, y_1, y_2) = \text{Tr}[M_{a|x} \otimes M_{b_1|y_1} \otimes M_{b_2|y_2} \mathcal{D}_\nu^{\otimes 3}(|\Phi_3^{\text{GHZ}}\rangle\langle\Phi_3^{\text{GHZ}}|_{AB_1B_2})] \quad (64)$$

$$= (1-\nu)^3 \text{Tr}[M_{a|x} \otimes M_{b_1|y_1} \otimes M_{b_2|y_2} |\Phi_3^{\text{GHZ}}\rangle\langle\Phi_3^{\text{GHZ}}|_{AB_1B_2}] + [1 - (1-\nu)^3] \text{Tr}[M_{a|x} \otimes M_{b_1|y_1} \otimes M_{b_2|y_2} \mathcal{X}_\nu] \quad (65)$$

$$= (1-\nu)^3 P_{\text{GHZ}}(a, b_1, b_2|x, y_1, y_2) + [1 - (1-\nu)^3] P_\nu^L(a, b_1, b_2|x, y_1, y_2). \quad (66)$$

Here the behavior P_{GHZ} arises from measurements of the GHZ state (which allows us to violate the classical bound maximally). The local behavior P_ν^L arises from the same measurements (σ_z observable) for biseparable state and therefore can be expressed as a convex combination of deterministic behaviors.

Eve prepares a convex combination attack [42,43]

$$P_\nu^{\text{CC}}(a, b_1, b_2, e|x, y_1, y_2) = (1-\nu)^3 P_{\text{GHZ}}(a, b_1, b_2|x, y_1, y_2) \delta_{e,?} + [1 - (1-\nu)^3] P_\nu^L(a, b_1, b_2|x, y_1, y_2) \delta_{e,(a,b_1,b_2)}. \quad (67)$$

This attack might not be optimal as it uses a particular decomposition of P_ν . In order to optimize the attack, Eve should find a decomposition with a maximal weight of local behavior $[1 - (1-\nu)^3]$ here].

We now consider a particular strategy of postprocessing the data which is in Eve's possession, represented by a channel $E \rightarrow F$ in Corollary 2. Following Ref. [17], we consider only the distribution coming from a key-generating measurement, which according to the protocol of Ref. [20] is $X = 0$ for Alice and $B_1 = 2$ and $B_2 = 0$ for the Bobs in the case of $N = 3$,

$$P_\nu^{\text{attack}}(a, b_1, b_2, f|020) = \Lambda_{E \rightarrow F} P_\nu^{\text{CC}}(a, b_1, b_2, e|020) = (1-\nu)^3 P_{\text{GHZ}}(a, b_1, b_2|x, y_1, y_2) \delta_{e,?} + [1 - (1-\nu)^3] P_\nu^L(a, b_1, b_2|x, y_1, y_2) \delta_{e,(a,b_1,b_2)} \times [\delta_{a,b,c} \delta_{f,a} + (1 - \delta_{a,b,c}) \delta_{f,?}], \quad (68)$$

where $\delta_{a,b,c}$ is 1 if all three indices have the same value and 0 otherwise. The above attack strategy is therefore a direct three-partite generalization of strategy proposed in Ref. [17]. The eavesdropper aims to be correlated only with the events $(a, b_1, b_2) = (0, 0, 0)$ or $(a, b_1, b_2) = (1, 1, 1)$, which mimic outputs of the honest strategy of the GHZ state. By applying the above attack strategy, we are ready to plot an upper bound on the reduced cc-squashed entanglement Corollary 2.

VI. GAP BETWEEN DI-CKA AND DD-CKA

In this section we provide a bound on the conference key agreement of N parties in terms of the bounds for groupings of these parties into groups of fewer than N users. We further show that there is a gap between the device-independent and device-dependent conference key agreement rates. This gap

implies that there are states for which there are no measurements used for testing and no CLOPC protocol that can achieve the same number of keys as in the device-dependent case. The gap is inherited from the analogous gap shown for the bipartite case [16].

In what follows, by a (nontrivial) partition \mathcal{P} of the set of systems $\{A_1, \dots, A_N\}$, we mean any grouping of the systems into at least two but no more than $N - 1$ subsets such that each A_i belongs to exactly one subset and each of them belongs to some subset.

Let us now generalize the definition of the reduced device-dependent key to the case of the conference key agreement. We will further also show the fact that the latter quantity bounds the device-independent conference key (i.e., Theorem 6 of Ref. [16]).

Definition 10. The reduced device-dependent conference key rate of an N -partite state $\rho_{N(A)}$ reads

$$K^\downarrow(\rho_{N(A)}) := \sup_{\mathcal{M}} \inf_{(\sigma_{N(A)}, \mathcal{L}) = (\rho_{N(A)}, \mathcal{M})} K_{\text{DD}}(\sigma_{N(A)}). \quad (69)$$

A direct analog of Theorem 6 of Ref. [16] (with an analogous proof which we omit here) states that the *reduced* device-dependent key upper bounds the device-independent key.

Theorem 6. For any N -partite state $\rho_{N(A)}$ and any \mathcal{M} , there is

$$K_{\text{DI}}(\rho_{N(A)}, \mathcal{M}) \leq \inf_{(\sigma_{N(A)}, \mathcal{L}) = (\rho_{N(A)}, \mathcal{M})} K_{\text{DD}}(\rho_{N(A)}) \quad (70)$$

and in particular

$$K_{\text{DI}}(\rho_{N(A)}) \equiv \sup_{\mathcal{M}} K_{\text{DI}}(\rho_{N(A)}, \mathcal{M}) \leq K^\downarrow(\rho_{N(A)}). \quad (71)$$

We first observe the following bound.

Proposition 3. For any N -partite quantum behavior $(\rho_{N(A)}, \mathcal{M})$ there is

$$K_{\text{DI,dev}}^{\text{iid}}(\rho_{N(A)}, \mathcal{M}) \leq \min \left\{ \min_{\mathcal{P}} K_{\text{DI,dev}}^{\text{iid}}(\rho_{\mathcal{P}(N(A))}) \times \min_{\mathcal{P}} \inf_{(\sigma_{\mathcal{P}(N(A))}, \mathcal{L}) = (\rho_{\mathcal{P}(N(A))}, \mathcal{M})} K_{\text{DD}}(\sigma_{\mathcal{P}(N(A))}) \right\}, \quad (72)$$

where \mathcal{P} is any nontrivial partition of the set of systems A_1, \dots, A_N .

Proof. The proof of the bound by $K_{\text{DI,dev}}^{\text{iid}}(\rho_{\mathcal{P}(N(A))})$ follows from the fact that any protocol of distillation of the DI conference key from the N -partite state is a special case of a protocol that distills the DI conference key from a nontrivial partition \mathcal{P} . This is because the class of LOPC protocols in these two scenarios is in relation to $\text{LOPC}(A_1, \dots, A_N) \subsetneq \text{LOPC}(\mathcal{P}(A_1, \dots, A_N))$. The other bound follows from the fact that for any grouping \mathcal{P} , by Theorem 6 above,

$$K_{\text{DI,dev}}^{\text{iid}}(\rho_{\mathcal{P}(N(A))}, \mathcal{M}) \leq \inf_{(\sigma_{\mathcal{P}(N(A))}, \mathcal{L}) = (\rho_{\mathcal{P}(N(A))}, \mathcal{M})} K_{\text{DD}}(\sigma_{\mathcal{P}(N(A))}). \quad (73)$$

An analogous fact to the above holds for $K_{\text{DI,par}}^{\text{iid}}$ as well.

Following Ref. [16], we show now that there is a gap between the numbers of conference keys and device-independent conference keys. We will use the fact that there it has been proven that there are states with $K^\downarrow(\rho_{AB}) < K_{\text{DD}}(\rho_{AB})$. From such state ρ_{AB} we construct a multipartite

state with the property that $K_{\text{DI}}(\rho_{N(A)}) < K_{\text{DD}}(\rho_{N(A)})$, as it is described in the proof of the following theorem.

Theorem 7. Let $\rho_{AB} \in \mathcal{B}(\mathcal{H}_A \otimes \mathcal{B})$, where $\dim(\mathcal{H}_A) = d_A$ and $\dim(\mathcal{H}_B) = d_B$, be a bipartite state which admits a gap $K_{\text{DD}}(\rho_{AB}) - K^\downarrow(\rho_{AB}) \geq c > 0$ for some constant c . Then for any N there is a multipartite state $\rho_{N(A)}$ with local dimensions at most $d_A \times d_B$ with $K_{\text{DD}}(\rho_{N(A)}) - K_{\text{DI}}(\rho_{N(A)}) \geq c$.

Proof. Consider a state $\rho_{N(A)}$ constructed as a path made of state ρ_{AB} (as, e.g., in a line of a quantum repeater):

$$\tilde{\rho}_{N(A)} := \rho_{A_1^1 A_1^2} \otimes \rho_{A_2^1 A_2^2} \otimes \rho_{A_3^1 A_3^2} \otimes \cdots \otimes \rho_{A_{N-1}^1 A_{N-1}^2}. \quad (74)$$

Here $\rho_{A_1^1 A_1^2} = \rho_{A_2^1 A_2^2} = \cdots = \rho_{A_{N-1}^1 A_{N-1}^2} = \rho_{AB}$ and by the way of notation we have $A_1^1 \equiv A_1$ and $A_1^2 A_2^1 \equiv A_2, \dots, A_{N-1}^2 \equiv A_N$. That is, the first party has only system A_1^1 and the last only system A_{N-1}^2 , while the i th party for $1 < i < N$ has systems $A_{i-1}^2 A_i^1$ at hand.

Since the states $\rho_{A_i^1 A_i^2}$ form a spanning tree of a graph of N systems (in fact a path), we can follow the lower bound given in Sec. VI A of Ref. [21] and note that

$$K_{\text{DD}}(\tilde{\rho}_{N(A)}) \geq \min_i K_{\text{DD}}(\rho_{A_i^1 A_i^2}) = K_{\text{DD}}(\rho_{AB}). \quad (75)$$

Indeed, the parties can first distill a key at rate $K_{\text{DD}}(\rho_{AB})$ along the edges of the path. Denote such distilled keys by k_{ij} between nodes i and j . Further, A_1 can XOR her key k_{12} with a locally generated private random bit string r of length $K_{\text{DD}}(\rho_{AB})$ and send $k_{12} \oplus r$ to A_2 ; further, A_2 can obtain $r = k_{12} \oplus (k_{12} \oplus r)$ and send it to the next party by XORing it with the key k_{23} . This process repeated $N - 1$ times, leaves all the parties knowing r , which remained secret due to one-time pad encryption by the keys $k_{12}, k_{23}, \dots, k_{N-1,N}$. It then suffices to note that, by Proposition 3,

$$K_{\text{DI}}(\tilde{\rho}_{N(A)}, \mathcal{M}) \leq \inf_{(\sigma_{A_1^1:(A_2^1, \dots, A_N^1)}, \mathcal{L}) = (\rho_{A_1^1:(A_2^1, \dots, A_N^1)}, \mathcal{M})} K_{\text{DD}}(\sigma_{A_1^1:(A_2^1, \dots, A_N^1)}) \quad (76)$$

$$\leq \inf_{(\sigma_{A_1^1:A_2^1}, \mathcal{L}) = (\rho_{A_1^1:A_2^1}, \mathcal{M})} K_{\text{DD}}(\sigma_{A_1^1:A_2^1}). \quad (77)$$

This is because there are no more conference keys than the number of device-dependent keys distilled in the cut $A_1^1 : (A_2^1, \dots, A_{N-1}^2)$. The latter is also upper bounded by the key distilled between systems A_1^1 and A_2^1 . This is due to the fact that any distillation protocol between A_1^1 and (A_2^1, \dots, A_N) is a particular protocol distilling key between systems A_1^1 and A_2^1 .

Taking the supremum over \mathcal{M} on both sides of the inequality (77), we obtain

$$\begin{aligned} K_{\text{DI}}(\tilde{\rho}_{N(A)}) &\equiv \sup_{\mathcal{M}} K_{\text{DI}}(\tilde{\rho}_{N(A)}, \mathcal{M}) \\ &\leq K^\downarrow(\rho_{A_1^1:A_2^1}) \equiv \sup_{\mathcal{M}} \inf_{(\sigma_{A_1^1:A_2^1}, \mathcal{L}) = (\rho_{A_1^1:A_2^1}, \mathcal{M})} K_{\text{DD}}(\sigma_{A_1^1:A_2^1}) \\ &= K^\downarrow(\rho_{AB}). \end{aligned} \quad (78)$$

Hence we get $K_{\text{DI}}(\rho_{N(A)}) \leq K^\downarrow(\rho_{AB})$. This fact, by Eq. (75), and the fact that by assumption $K_{\text{DD}}(\rho_{AB}) - K^\downarrow(\rho_{AB}) \geq c > 0$ imply the following chain of inequalities:

$$K_{\text{DD}}(\tilde{\rho}_{N(A)}) \geq K_{\text{DD}}(\rho_{AB}) > K^\downarrow(\rho_{AB}) \geq K_{\text{DI}}(\tilde{\rho}_{N(A)}). \quad (79)$$

The above implies then the desired gap $K_{\text{DD}}(\tilde{\rho}_{N(A)}) - K_{\text{DI}}(\tilde{\rho}_{N(A)}) > 0$. Moreover, this gap is as large as $c > 0$ due

to the assumption that $K_{\text{DD}}(\rho_{AB}) - K_{\text{DI}}(\rho_{AB}) \geq c > 0$. The claim about dimensions follows from the form of the state given in Eq. (74).

From Ref. [16] we have the immediate corollary that there is a gap between the DI-CKA and DD-CKA.

Corollary 5. For any N there is a state $\tilde{\rho}_{N(A)}$ for which there is

$$K_{\text{DI,dev}}^{\text{iid}}(\tilde{\rho}_{N(A)}) < K_{\text{DD}}(\tilde{\rho}_{N(A)}). \quad (80)$$

Proof. Reference [16] shows an example of a bipartite state ρ_{AB} with the gap $K^\downarrow(\rho_{AB}) < K_{\text{DD}}(\rho_{AB})$. The construction given in Eq. (74) based on this ρ_{AB} proves the thesis via Theorem 7.

We note also that a bound similar to that in the above corollary holds for $K_{\text{DI,par}}^{\text{iid}}$ and K_{DI} itself due to the fact that $K_{\text{DI}}^{\text{iid}} \geq K_{\text{DI}}$ by definition [16]. We can modify the proof technique shown above to see the following general remark.

Remark 6. In the above construction one need not use only the state $\rho_{A_1^1 A_1^2}^{\otimes k}$ having $K^\downarrow(\rho_{A_1^1 A_1^2}) < K_{\text{DD}}(\rho_{A_1^1 A_1^2})$. In fact, the state on systems $A_2^1 A_2^2 \cdots A_{N-1}^1 A_{N-1}^2$ can be an arbitrary state having $K_{\text{DD}}(\rho_{A_2^1 A_2^2 \cdots A_{N-1}^1 A_{N-1}^2}) \geq K_{\text{DD}}(\rho_{A_1^1 A_1^2})$. It can be even a Φ_N^{GHZ} state of arbitrary large local dimension. This is with no change in the above proof if only $\rho_{A_1^1 A_1^2}$ is on systems $A_1^1 A_1^2$ with the gap we have mentioned. See Fig. 1 for the tripartite example.

VII. DI-CKA VERSUS GENUINE NONLOCALITY AND ENTANGLEMENT

We now discuss the topic of genuine nonlocality and entanglement in the context of the DI-CKA, introducing the notion of quantum locality.

We say that a behavior $P(\mathbf{a}|\mathbf{x})$ is local in a cut $(A_{i_1} \dots A_{i_k}) : (A_{i_{k+1}} \dots A_N)$ if it can be written as a product of two behaviors on systems $A_{i_1} \dots A_{i_k}$ and $A_{i_{k+1}} \dots A_N$, respectively. The behavior $P(\mathbf{a}|\mathbf{x})$ is genuinely nonlocal if and only if it is not a mixture of behaviors that are a product in at least one cut.

We show that any behavior from which the parties draw the conference key in a single-shot (single run) must exhibit genuine nonlocality. The scenario of a single run was considered in the context of a nonsignaling adversary [44,45]. For that reason, we depart from the traditional definition of DI quantum key distillation rate by considering a single-shot DI quantum key distillation rate obtained by an LOPC postprocessing of a distribution obtained from some behavior $P(\mathbf{a}|\mathbf{x})$ when all the parties measure all the inputs \mathbf{x} in parallel at the same time.

For the purpose of Theorem 8 below, by a ‘‘local’’ set we will mean the set of behaviors that are convex mixtures of behaviors that are a product in some cut and both behaviors in the product have quantum realization. We will denote this set by LQ (locally quantum). Any distribution which is not in LQ can be treated as genuinely nonlocal, although other definitions are adopted in the literature [11]. Exemplary extreme behavior in this set is a product of the Tsirelson behavior $P(a_1, a_2|x_1, x_2) = \text{Tr}[\Phi_2] \langle \Phi_2 | M_{a_1}^{x_1} \otimes M_{a_2}^{x_2} | \Phi_2 \rangle$ with $M_{a_1}^0 = \sigma_z$, $M_{a_1}^1 = \sigma_x$, $M_{a_2}^0 = \frac{(\sigma_x + \sigma_z)}{\sqrt{2}}$, $M_{a_2}^1 = \frac{(\sigma_z - \sigma_x)}{\sqrt{2}}$ (σ_x and σ_z being Pauli-X and Pauli-Z operators, respectively), and a determin-

istic local behavior $P(a_3|x_3): P(a_1, a_2|x_1, x_2)P(a_3|x_3)$. The theorem which we show below applies to the scenario where all the parties share a *single* copy of a device and measure the inputs in parallel to ensure nonsignaling. We give the definition of the key rate obtained in this setup in full analogy to Definition 1 as follows.

Definition 11. The maximum single-shot device-independent quantum key distillation rate of a device (ρ, \mathcal{M}) with independent and identically distributed behavior is defined as

$$K_{\text{DI,dev}}^{\text{single-shot}}(\rho, \mathcal{M}) := \inf_{\varepsilon > 0} \sup_{\hat{P}} \inf_{(6)} \kappa_n^\varepsilon(\hat{P}(\sigma, \mathcal{N})), \quad (81)$$

where κ_n^ε is the quantum key rate achieved for any security parameter ε and measurements \mathcal{N} .

Here \hat{P} is a protocol composed of classical local operations and public (classical) communication acting on a single copy of (σ, \mathcal{N}) which, composed with the measurement, results in a quantum local operations and public (classical) communication protocol.

We are ready to state the following theorem.

Theorem 8. If a behavior $P(\mathbf{a}|\mathbf{x})$ satisfies $K_{\text{DI,dev}}^{\text{single-shot}}(P(\mathbf{a}|\mathbf{x})) > 0$ then it is not in LQ.

Proof. The proof goes by contradiction. Suppose the behavior $P(\mathbf{a}|\mathbf{x})$ is not genuinely nonlocal. That is, it can be expressed as a convex combination of behaviors which are a product in at least one cut denoted by $(A_{j_1}^{(i)} \dots A_{j_k}^{(i)} : A_{j_{k+1}}^{(i)} \dots A_{j_N}^{(i)})$ for the i th behavior in the combination. We express this as

$$\sum_i p_i P_i(\mathbf{a}|\mathbf{x})_{(A_{j_1}^{(i)} \dots A_{j_k}^{(i)} : A_{j_{k+1}}^{(i)} \dots A_{j_N}^{(i)}), \quad (82)$$

where $P_i(\mathbf{a}|\mathbf{x})$ are some quantum behaviors. Consider then a device P_i as a bipartite one, with parties $(A_{j_1}^{(i)} \dots A_{j_k}^{(i)})$ together forming A' and $(A_{j_{k+1}}^{(i)} \dots A_{j_N}^{(i)})$ forming A'' . Such a device has zero bipartite DI quantum keys, as it is a product in cut $A' : A''$. By virtue of purification, Eve can have access to the mixture (82), knowing which of the mixing terms i happened. By Theorem 3 we have that from any of such terms, one cannot draw a conference key, as Eve has a local hidden variable model for it. Indeed, the right-hand side of (72) is then 0, as Eve can adopt an attack which, e.g., makes zero reduced

c-squashed entanglements [18]. We thus obtained the desired contradiction.

Let us now recall the notion of genuine entanglement. We say that a multipartite state $\rho_{A_1 A_2 \dots A_N}$ is separable in a cut $(A_{i_1} \dots A_{i_k}) : (A_{i_{k+1}} \dots A_N)$ if it can be written as convex mixtures of product states between systems $A_{i_1} \dots A_{i_k}$ and $A_{i_{k+1}} \dots A_N$. If a multipartite state $\rho_{A_1 A_2 \dots A_N}$ can be written as a mixture of separable states that are product in at least one cut then it is called biseparable. We say that $\rho_{A_1 A_2 \dots A_N}$ is genuinely entangled if and only if it is not a mixture of separable states that are product in at least one cut. It was shown in Ref. [46] that there exist N -partite states for all $N > 2$ where some genuinely entangled states admit a fully LRHV model, i.e., where all parties are separated.

Let $\text{GE}(N(A))$, $\text{BS}(N(A))$, and $\text{FS}(N(A))$ denote the set of all N -partite states $\rho_{A_1 A_2 \dots A_N}$ that are genuinely entangled, biseparable, and fully separable, respectively (see Ref. [21]). For n copies of N -partite state, when we consider partition across designated N parties, we denote local groupings by $N(A^{\otimes n})$.

Remark 7. It is necessary to consider a single-shot DI key in Theorem 8 because the set of LQ behaviors is not closed under tensor product. This is for the same reason that the set of biseparable states is not closed under a tensor product.

The following theorem follows from Observation 1 as well as Proposition 2 of Ref. [21].

Theorem 9. The maximum device-independent conference key agreement rates of a device (ρ, \mathcal{M}) are upper bounded by

$$K_{\text{DI,dev}}(\rho_{N(A)}, \mathcal{M}) \leq \inf_{(\sigma_{N(A)}, \mathcal{L})=(\rho, \mathcal{M})} E_{\text{GE}}^\infty(\sigma), \quad (83)$$

$$K_{\text{DI,par}}(\rho, \mathcal{M}) \leq \inf_{\substack{\omega(\sigma_{N(A)}, \mathcal{L})=\omega(\rho_{N(A)}, \mathcal{M}) \\ P_{\text{err}}(\sigma, \mathcal{L})=P_{\text{err}}(\rho, \mathcal{M})}} E_{\text{GE}}^\infty(\sigma), \quad (84)$$

where $E_{\text{GE}}^\infty(\zeta)$ is the regularized relative entropy of genuine entanglement [21] of a state $\zeta_{A_1 A_2 \dots A_N}$, with

$$E_{\text{GE}}^\infty(\zeta) = \inf_{\varphi \in \text{BS}(N(A^{\otimes n}))} \lim_{n \rightarrow \infty} \frac{1}{n} D(\zeta^{\otimes n} \| \varphi), \quad (85)$$

where $D(\rho \| \sigma)$ is the relative entropy between two states ρ and σ , with $D(\rho \| \sigma) = \text{Tr}[\rho(\log_2 \rho - \log_2 \sigma)]$ if $\text{supp } \rho \subseteq \text{supp } \sigma$; otherwise it is ∞ [47].

We note here that there is a trivial bound that can be obtained from Theorem 9 above, which is encapsulated in the following corollary.

Corollary 6. For any state $\rho_{N(A)}$ with $\min_{i \in \{1, \dots, N\}} d_{A_i} =: d$ there is

$$\begin{aligned} K_{\text{DI}}(\rho_{N(A)}) &\equiv \sup_{\mathcal{M}} K_{\text{DI}}(\rho, \mathcal{M}) \\ &\leq \min\{p \log_2 d : p \in [0, 1], \rho = p\rho' + (1-p)\rho_{\text{fs}}, \rho_{\text{fs}} \in \text{FS}\}. \end{aligned} \quad (86)$$

Proof. Given any decomposition of a state $\rho_{N(A)}$ into $\rho = p\rho' + (1-p)\rho_{\text{fs}}$, where the state ρ_{fs} is a fully separable state, we have

$$K_{\text{DI}}(\rho) \leq \sup_{\mathcal{M}} K_{\text{DI,dev}}^{\text{iid}}(\rho, \mathcal{M})$$

$$\begin{aligned} &\leq \sup_{\mathcal{M}} \inf_{(\sigma, \mathcal{L})=(\rho, \mathcal{M})} E_{\text{GE}}^\infty(\sigma) \leq \sup_{\mathcal{M}} E_R(\rho) \\ &\leq p E_R(\rho') \leq p \min_i \log_2 d_{A_i}, \end{aligned} \quad (87)$$

where we have used Theorem 9 (also see Corollary 6 of [21]) and the fact that $E_R(\rho) = \inf_{\kappa \in \text{FS}} D(\rho \| \kappa)$ is (i) convex,

(ii) zero on fully separable states, and (iii) does not exceed the minimum logarithm of dimensions of the input state, which can be proved by noticing that $E_R(\rho) \leq D(\rho \parallel \rho_{A_i} \otimes \rho_{A_{\neq i}}) = I(A_i : A_{\neq i}) \leq \log_2 d$ where A_i has minimal dimension among systems A_1, \dots, A_N .

We presented this bound in Fig. 2 in Sec. V and we saw that it is indeed above the upper bounds which we derive in Secs. III and IV.

VIII. CONCLUSION

We have demonstrated a number of upper bounds on the quantum secure conference key, generalizing (i) the results of Ref. [18] regarding a relative entropy based bound and (ii) the results of Ref. [17] regarding the reduced c-squashed entanglement. More precisely, we based our demonstration on two secrecy monotones which generalize intrinsic information to the multipartite case, following Ref. [23]. Although, as shown in Ref. [24], the quantities $I(A_1 : \dots : A_2)$ and $S_N(A_1 : \dots : A_N)$ (referred to as D_N) coincide for pure states, they are not comparable for mixed states and E_{sq}^c based on $I(A_1 : \dots : A_N)$ performs better as an upper bound. It is however an open problem if one can establish an inequality between the two reduced c-squashed entanglements.

Interestingly, the approach of Ref. [17] does not result in zero keys in any noise regimes for the parity CHSH game of Ref. [20]. It would be important to see if this can be improved by changing Eve’s strategy or the bound needs to be changed.

We have also shown that the fundamental gap between device-independent and device-dependent keys also holds in the multipartite case. We have given an exemplary state which is based directly on the bipartite states given in Ref. [16]. It is interesting if such a state exists in lower dimensions or even possibly on N qubits.

Finally, our results hold for the static case of quantum states. The next step would be to generalize the results of Ref. [18] for the dynamic case of quantum channels to the multipartite scenario.

Note added. The topic of upper bounds on the DI-CKA is also studied in the parallel work of [48]. Comparison between basic approaches (i.e., for the DI quantum key distribution between two honest parties) used in Ref. [48] and in this paper to get upper bounds on DI-CKA is discussed in Ref. [18].

ACKNOWLEDGMENTS

We acknowledge partial support by the Foundation for Polish Science (IRAP Project ICTQT, Contract No. MAB/2018/5, cofinanced by EU within Smart Growth Operational Programme). The International Centre for Theory of Quantum Technologies project (Contract No. MAB/2018/5) is carried out within the International Research Agendas Programme of the Foundation for Polish Science cofinanced by the European Union from the funds of the Smart Growth Operational Programme, axis IV: Increasing the research potential (Measure 4.3). M.W. acknowledges grant Sonata Bis 5 (Grant No. 2015/18/E/ST2/00327) from the National Science Center. S.D. acknowledges Individual Fellowships at Université libre de Bruxelles; this project received funding from the European

Union’s Horizon 2020 research and innovation program under the Marie Skłodowska-Curie Grant Agreement No. 801505. S.D. also thanks Harish-Chandra Research Institute, Prayagraj (Allahabad, India) for hospitality during his visit where part of this work was done.

APPENDIX A: CONTINUITY STATEMENTS

We have the following lemmas.

Lemma 4 (Alicki-Fannes-Winter continuity bounds [49]). For states ρ_{AB} and σ_{AB} , if $\frac{1}{2} \|\rho - \sigma\|_1 \leq \epsilon \leq 1$, then

$$|S(A|B)_\rho - S(A|B)_\sigma| \leq 2\epsilon \log_2 d + g(\epsilon), \tag{A1}$$

where $d = \dim(\mathcal{H}_A) < \infty$ and $g(\epsilon) := (1 + \epsilon) \log_2(1 + \epsilon) - \epsilon \log_2 \epsilon$.

Lemma 5 (from [41]). If $d = \min\{\dim(\mathcal{H}_A), \dim(\mathcal{H}_B)\} < +\infty$, then

$$|I(A; B|C)_\rho - I(A; B|C)_\sigma| \leq 2\epsilon \log_2 d + 2g(\epsilon) \tag{A2}$$

for any states ρ_{ABC} and σ_{ABC} , where $\epsilon = \frac{1}{2} \|\rho - \sigma\|_1$.

APPENDIX B: SECRECY MONOTONES

In this Appendix we revisit Theorem 3.1 of Ref. [25] and generalize the result by relaxing the constraints on the Hilbert spaces in the following way. First, we prove an analogy to Lemma A.1 of Ref. [25].

Lemma 6 (cf. [25]). The maximization in the definition of K_{DD} (12) can be restricted to protocols that use communication at most linear in the number of copies of ρ_{ABE} . The eavesdropper system is not necessarily restricted to a finite dimension.

Proof. The proof of Lemma 6 goes along the lines of the proof of Lemma A.1 in Ref. [25]. The change that is necessary to allow the eavesdropper to hold the system of infinite dimension is the use of asymptotic continuity of the conditional mutual information of Ref. [41] (see Lemma 5 herein) instead of the Alicki-Fannes inequality. This results in

$$I(A : B)_\sigma - I(A : E)_\sigma \geq l_{n_0}(1 - 4\epsilon) - 4g(\epsilon), \tag{B1}$$

where l_{n_0} is the length of the output of a distillation protocol using n_0 copies of the input state. The state σ is the output of the latter protocol. The overall key rate of the modified protocol which has linear communication admits then a lower bound

$$\tilde{R} \geq (1 - 4\epsilon)(R - \epsilon) - \frac{4g(\epsilon)}{n_0}. \tag{B2}$$

The other parts of the proof are not altered.

Lemma 7 (cf. [25]). Let $E(\rho)$ be a function mapping a tripartite quantum state ρ_{ABE} into positive numbers such that the following hold: (a) monotonicity, i.e., $E(\Lambda(\rho)) \leq E(\rho)$ for any LOPC Λ ; (b) asymptotic continuity, i.e., for any states ρ^n and σ^n on $\mathcal{H}_A \otimes \mathcal{H}_B \otimes \mathcal{H}_E$, the condition $\|\rho^n - \sigma^n\|_1 \rightarrow 0$ implies $\frac{1}{\log_2 r_n} |E(\rho^n) - E(\sigma^n)| \rightarrow 0$, where $r_n = \dim(\mathcal{H}_A^n)$; and (c) normalization, i.e., $E(\tau^{(l)}) = l$. Then the regularization of the function E given by $E^\infty(\rho) = \limsup_{n \rightarrow \infty} \frac{M(\rho^{\otimes n})}{n}$ is an upper bound on the device-dependent key distillation rate K_{DD} , i.e., $E^\infty(\rho_{ABE}) \geq K_{DD}(\rho_{ABE})$ for all ρ_{ABE} with $\dim_A <$

∞ , if in addition E satisfies (d) subadditivity on tensor products: $E(\rho^{\otimes n}) \leq nE(\rho)$; then E is an upper bound on K_{DD} .


Proof. The proof arguments are same as those stated in Ref. [25] with relaxation on the Hilbert space of E . We ob-


serve that the proof arguments hold even when there is no restriction on the $\dim(\mathcal{H}_E)$, i.e., E can be finite dimensional or infinite dimensional. It suffices to have $\dim(\mathcal{H}_A)$ be finite dimensional.

-
- [1] J. P. Dowling and G. J. Milburn, Quantum technology: The second quantum revolution, *Philos. Trans. R. Soc. A* **361**, 1655 (2003).
- [2] S. Wehner, D. Elkouss, and R. Hanson, Quantum internet: A vision for the road ahead, *Science* **362**, eaam9288 (2018).
- [3] W. Dür, H.-J. Briegel, J. I. Cirac, and P. Zoller, Quantum repeaters based on entanglement purification, *Phys. Rev. A* **59**, 169 (1999).
- [4] S. Muralidharan, L. Li, J. Kim, N. Lütkenhaus, M. D. Lukin, and L. Jiang, Optimal architectures for long distance quantum communication, *Sci. Rep.* **6**, 20463 (2016).
- [5] Q. Zhang, F. Xu, Y.-A. Chen, C.-Z. Peng, and J.-W. Pan, Large scale quantum key distribution: Challenges and solutions [invited], *Opt. Express* **26**, 24260 (2018).
- [6] C. H. Bennett and G. Brassard, in *Proceedings of the International Conference on Computers, Systems and Signal Processing, Bangalore, 1984* (IEEE, Piscataway, 1984), pp. 175–179.
- [7] S. Pironio, A. Acín, N. Brunner, N. Gisin, S. Massar, and V. Scarani, Device-independent quantum key distribution secure against collective attacks, *New J. Phys.* **11**, 045021 (2009).
- [8] G. T. Becker, F. Regazzoni, C. Paar, and W. P. Bursell, in *Cryptographic Hardware and Embedded Systems*, edited by G. Bertoni and J. S. Coron, Lecture Notes in Computer Science Vol. 8086 (Springer, Berlin, 2013), pp. 197–214.
- [9] V. Makarov, Controlling passively quenched single photon detectors by bright light, *New J. Phys.* **11**, 065003 (2009).
- [10] A. K. Ekert, Quantum Cryptography Based on Bell’s Theorem, *Phys. Rev. Lett.* **67**, 661 (1991).
- [11] N. Brunner, D. Cavalcanti, S. Pironio, V. Scarani, and S. Wehner, Bell nonlocality, *Rev. Mod. Phys.* **86**, 839(E) (2014).
- [12] W. Zhang, T. van Leent, K. Redeker, R. Garthoff, R. Schwonnek, F. Fertig, S. Eppelt, V. Scarani, C. C.-W. Lim, and H. Weinfurter, Experimental device-independent quantum key distribution between distant users, [arXiv:2110.00575](https://arxiv.org/abs/2110.00575).
- [13] D. P. Nadlinger, P. Drmota, B. C. Nichol, G. Araneda, D. Main, R. Srinivas, D. M. Lucas, C. J. Ballance, K. Ivanov, E. Y.-Z. Tan, P. Sekatski, R. L. Urbanke, R. Renner, N. Sangouard, and J.-D. Bancal, Device-independent quantum key distribution, [arXiv:2109.14600](https://arxiv.org/abs/2109.14600).
- [14] W.-Z. Liu, Y.-Z. Zhang, Y.-Z. Zhen, M.-H. Li, Y. Liu, J. Fan, F. Xu, Q. Zhang, and J.-W. Pan, High-speed device-independent quantum key distribution against collective attacks, [arXiv:2110.01480](https://arxiv.org/abs/2110.01480).
- [15] E. Kaur, M. M. Wilde, and A. Winter, Fundamental limits on key rates in device-independent quantum key distribution, *New J. Phys.* **22**, 023039 (2020).
- [16] M. Christandl, R. Ferrara, and K. Horodecki, Upper Bounds on Device-Independent Quantum Key Distribution, *Phys. Rev. Lett.* **126**, 160501 (2021).
- [17] M. Farkas, M. Balanzó-Juandó, K. Łukanowski, J. Kołodyński, and A. Acín, Bell Nonlocality Is Not Sufficient for the Security of Standard Device-Independent Quantum Key Distribution Protocols, *Phys. Rev. Lett.* **127**, 050503 (2021).
- [18] E. Kaur, K. Horodecki, and S. Das, Upper bounds on device-independent quantum key distribution rates in static and dynamic scenarios, [arXiv:2107.06411](https://arxiv.org/abs/2107.06411).
- [19] G. Murta, F. Grasselli, H. Kampermann, and D. Bruß, Quantum conference key agreement: A review, *Adv. Quantum Technol.* **3**, 2000025 (2020).
- [20] J. Ribeiro, G. Murta, and S. Wehner, Fully device-independent conference key agreement, [arXiv:1708.00798](https://arxiv.org/abs/1708.00798) (2019).
- [21] S. Das, S. Bäuml, M. Winczewski, and K. Horodecki, Universal Limitations on Quantum Key Distribution over a Network, *Phys. Rev. X* **11**, 041016 (2021).
- [22] R. Arnon-Friedman and F. Leditzky, Upper bounds on device-independent quantum key distribution rates and a revised Peres conjecture, *IEEE Trans. Inf. Theory* **67**, 6606 (2021).
- [23] D. Yang, K. Horodecki, M. Horodecki, P. Horodecki, J. Oppenheim, and W. Song, Squashed entanglement for multipartite states and entanglement measures based on the mixed convex roof, *IEEE Trans. Inf. Theory* **55**, 3375 (2009).
- [24] M. M. Wilde, Squashed entanglement and approximate private states, *Quantum Inf. Process.* **15**, 4563 (2016).
- [25] M. Christandl, A. Ekert, M. Horodecki, P. Horodecki, J. Oppenheim, and R. Renner, in *Theory of Cryptography*, edited by S. P. Vadhan, Lecture Notes in Computer Science Vol. 4392 (Springer, Berlin, 2007), pp. 456–478.
- [26] N. J. Cerf, S. Massar, and S. Schneider, Multipartite classical and quantum secrecy monotones, *Phys. Rev. A* **66**, 042309 (2002).
- [27] R. Horodecki, P. Horodecki, M. Horodecki, and K. Horodecki, Quantum entanglement, *Rev. Mod. Phys.* **81**, 865 (2009).
- [28] R. Arnon-Friedman, F. Dupuis, O. Fawzi, R. Renner, and T. Vidick, Practical device-independent quantum cryptography via entropy accumulation, *Nat. Commun.* **9**, 459 (2018).
- [29] N. D. Mermin, Extreme Quantum Entanglement in a Superposition of Macroscopically Distinct States, *Phys. Rev. Lett.* **65**, 1838 (1990).
- [30] M. Ardehali, Bell inequalities with a magnitude of violation that grows exponentially with the number of particles, *Phys. Rev. A* **46**, 5375 (1992).
- [31] A. V. Belinskĭ and D. N. Klyshko, Interference of light and Bell’s theorem, *Phys.-Usp.* **36**, 653 (1993).
- [32] M. Seevinck and G. Svetlichny, Bell-Type Inequalities for Partial Separability in N -Particle Systems and Quantum Mechanical Violations, *Phys. Rev. Lett.* **89**, 060401 (2002).
- [33] M. Żukowski and Č. Brukner, Bell’s Theorem for General N -Qubit States, *Phys. Rev. Lett.* **88**, 210401 (2002).
- [34] R. F. Werner and M. M. Wolf, All-multipartite Bell-correlation inequalities for two dichotomic observables per site, *Phys. Rev. A* **64**, 032112 (2001).

- [35] S. Yu, Q. Chen, C. Zhang, C. H. Lai, and C. H. Oh, All Entangled Pure States Violate a Single Bell's Inequality, *Phys. Rev. Lett.* **109**, 120402 (2012).
- [36] D. Home, D. Saha, and S. Das, Multipartite Bell-type inequality by generalizing Wigner's argument, *Phys. Rev. A* **91**, 012102 (2015).
- [37] M.-X. Luo, Fully device-independent model on quantum networks, [arXiv:2106.15840](https://arxiv.org/abs/2106.15840).
- [38] M. Żukowski, Č. Brukner, W. Laskowski, and M. Wieśniak, Do All Pure Entangled States Violate Bell's Inequalities for Correlation Functions? *Phys. Rev. Lett.* **88**, 210402 (2002).
- [39] M. Winczewski, T. Das, and K. Horodecki, Limitations on device independent secure key via squashed non-locality, [arXiv:1903.12154](https://arxiv.org/abs/1903.12154).
- [40] S. Watanabe, Information theoretical analysis of multivariate correlation, *IBM J. Res. Dev.* **4**, 66 (1960).
- [41] M. E. Shirokov, Tight uniform continuity bounds for the quantum conditional mutual information, for the Holevo quantity, and for capacities of quantum channels, *J. Math. Phys.* **58**, 102202 (2017).
- [42] A. Acín, N. Gisin, and L. Masanes, From Bell's Theorem to Secure Quantum Key Distribution, *Phys. Rev. Lett.* **97**, 120405 (2006).
- [43] A. Acín, S. Massar, and S. Pironio, Efficient quantum key distribution secure against no-signalling eavesdroppers, *New J. Phys.* **8**, 126 (2006).
- [44] E. Hänggi, R. Renner, and S. Wolf, Efficient device-independent quantum key distribution, in *Annual International Conference on the Theory and Applications of Cryptographic Techniques, Advances in Cryptology EUROCRYPT 2010*, Lecture Notes in Computer Science, Vol. 6110, edited by H. Gilbert (Springer, Berlin, Heidelberg, 2010), pp. 216–234.
- [45] L. Masanes, Universally Composable Privacy Amplification from Causality Constraints, *Phys. Rev. Lett.* **102**, 140501 (2009).
- [46] J. Bowles, J. Francfort, M. Fillettaz, F. Hirsch, and N. Brunner, Genuinely Multipartite Entangled Quantum States with Fully Local Hidden Variable Models and Hidden Multipartite Nonlocality, *Phys. Rev. Lett.* **116**, 130401 (2016).
- [47] H. Umegaki, Conditional expectations in an operator algebra. IV. Entropy and information, *Kodai Math. Sem. Rep.* **14**, 59 (1962).
- [48] A. Philip, E. Kaur, P. Bierhorst, and M. M. Wilde, Intrinsic non-locality and device-independent conference key agreement, [arXiv:2111.02596](https://arxiv.org/abs/2111.02596).
- [49] A. Winter, Tight uniform continuity bounds for quantum entropies: Conditional entropy, relative entropy distance and energy constraints, *Commun. Math. Phys.* **347**, 291 (2016).

Erratum: Fundamental limitations on the device-independent quantum conference key agreement [Phys. Rev. A **105, 022604 (2022)]**

Karol Horodecki, Marek Winzewski, and Siddhartha Das 

 (Received 12 January 2023; published 16 February 2023)

DOI: [10.1103/PhysRevA.107.029902](https://doi.org/10.1103/PhysRevA.107.029902)

We recently learned that the bound presented in Sec. IV of our paper is redundant. This fact motivated us, apart from describing which parts should be disregarded (see below), to also revise some of the text. It is important to note that there is no change with respect to the results and proofs of the original version of our paper. We describe below essential changes.

We have learned that the two measures of multipartite entanglement—multipartite squashed entanglement E_{sq} and its dual \tilde{E}_{sq} in our paper are, in fact, equal to each other. The equality follows from Theorem 7 in Ref. [1]. Precisely the two definitions given in our paper below are equivalent [1].

Definition 4 [from Ref. [23] (of the original paper)]. “For an N -partite state ρ_{A_1, \dots, A_N} ,

$$E_{\text{sq}}^q(\rho_{A_1, \dots, A_N}) := \inf_{\sigma} I(A_1 : A_2 : \dots : A_N | E)_{\sigma}, \quad (20)$$

where the infimum is taken over states $\sigma_{A_1, \dots, A_N E}$ that are extensions of ρ_{A_1, \dots, A_N} , i.e., $\text{Tr}_E[\sigma_{A_1, \dots, A_N E}] = \rho_{A_1, \dots, A_N}$,” and

Definition 7. “For an N -partite state ρ_{A_1, \dots, A_N} ,

$$\tilde{E}_{\text{sq}}(\rho_{N(A)}) := \inf D_N(\sigma_{N(A)E}), \quad (49)$$

where the infimum is taken over states $\sigma_{N(A)E}$ that are extensions of $\rho_{N(A)}$, i.e., $\text{Tr}_E[\sigma_{N(A)E}] = \rho_{N(A)}$.”

Here,

$$I(A_1 : \dots : A_N | E)_{\rho} = \sum_{i=1}^N S(A_i | E)_{\rho} - S(A_1, \dots, A_N | E)_{\rho}.$$

and

$$D_N(\rho_{N(A)E}) := I(A_1 : A_2 \dots A_N | E)_{\rho_{N(A)E}} + I(A_2 : A_3 \dots A_N | A_1 E)_{\rho_{N(A)E}} + I(A_3 : A_4 \dots A_N | A_1 A_2 E)_{\rho_{N(A)E}} \\ + \dots + I(A_{N-1} : A_N | A_1 \dots A_{N-2} E)_{\rho_{N(A)E}}.$$

In turn, our upper bounds on the device-independent (DI) key based on the dual measure given in Theorem 5 and Corollary 3 in Sec. IV of the paper equals the multipartite reduced c-squashed entanglement bounds given in Theorems 2 and Corollary 2 in Sec. III of our paper, respectively. For this reason, any reference to the dual function, including these in the discussion section, should be skipped in reading. As a result, Fig. 2 also should not contain the plot of a dual bound, which appeared to be not equal to the c-squashed due to lack of optimization. Figure 2 should look like Fig. 2 here.

We have also noted typographical errors that could make unclear Eq. (68) and text around it. The new text reflects the numerics that was actually performed in our paper, hence, the modification presented below does not affect the plot given in Fig. 2 there. It was

$$\begin{aligned} \text{“}P_v^{\text{attack}}(a, b_1, b_2, f|020) &= \Lambda_{E \rightarrow F} P_v^{\text{CC}}(a, b_1, b_2, e|020) \\ &= (1 - v)^3 P_{\text{GHZ}}(a, b_1, b_2 | x, y_1, y_2) \delta_{e,?} + [1 - (1 - v)^3] P_v^{\text{L}}(a, b_1, b_2 | x, y_1, y_2) \\ &\quad \times \delta_{e, (a, b_1, b_2)} [\delta_{a, b, c} \delta_{f, a} + (1 - \delta_{a, b, c}) \delta_{f, ?}], \end{aligned} \quad (68)$$

where $\delta_{a, b, c}$ is 1 if all three indices have the same value and 0 otherwise. The above attack strategy is, therefore, a direct three-partite generalization of strategy proposed in Ref. [17]. The eavesdropper aims to be correlated only with the events $(a, b_1, b_2) = (0, 0, 0)$ or $(a, b_1, b_2) = (1, 1, 1)$, which mimic outputs of the honest strategy of the Greenberger-Horne-Zeilinger (GHZ) state. By applying the above attack strategy, we are ready to plot an upper bound on the reduced cc-squashed entanglement Corollary 2.”

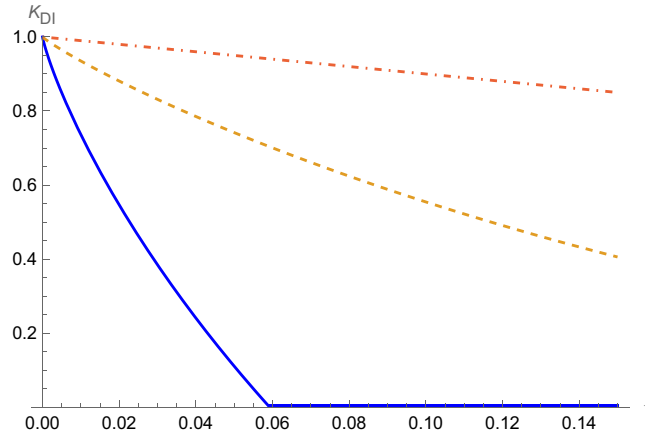


FIG. 2. Plot of upper and lower bounds on the device-independent conference key agreement (DI-CKA) of Ref. [2]. The yellow dashed line represents an upper bound (not fully optimized) on the upper bound $\frac{1}{N-1}I[N(A) \downarrow E]$ from Eq. (39) with the attack strategy in Eq. (68). The red dashed-dot curve is the trivial upper bound obtained in Corollary 5 via the relative entropy of entanglement bound $(1 - \nu)$. The blue solid line represents the lower bound from Ref. [2].

It should be now as follows:

$$\begin{aligned} \text{“}P_v^{\text{attack}}(a, b_1, b_2, f|020) &= \Lambda_{E \rightarrow F} P_v^{\text{CC}}(a, b_1, b_2, e|020) \\ &= (1 - \nu)^3 P_{\text{GHZ}}(a, b_1, b_2|020) \delta_{f,?} + [1 - (1 - \nu)^3] P_v^{\text{L}}(a, b_1, b_2|020) [\delta_{a,b_1,b_2} \delta_{f,a} + (1 - \delta_{a,b_1,b_2}) \delta_{f,?}], \end{aligned} \quad (68)$$

where δ_{a,b_1,b_2} is 1 if all three indices have the same value and 0 otherwise. The above attack strategy is, therefore, a direct three-partite generalization of strategy proposed in Ref. [3]. The eavesdropper aims to be correlated only with the events $(a, b_1, b_2) = (0, 0, 0)$ or $(a, b_1, b_2) = (1, 1, 1)$, whenever they originate from the local behavior P_v^{L} , and maps all other events to $f = ?$. By applying the above attack strategy, we are ready to plot an upper bound on the reduced c-squashed entanglement shown in Corollary 2. The latter bound is a multipartite version of the intrinsic information [4,5], used first for the bipartite case in Ref. [6] against a nonsignaling adversary (see in this context Refs. [7–9]). Here, the strategy of Eve to process her classical variable E to F is based on Ref. [3] as shown above.”

Proposition 3 originally read as follows:

Proposition 3. “For any N -partite quantum behavior $(\rho_{N(A)}, \mathcal{M})$ there is

$$K_{\text{DI,dev}}^{\text{iid}}(\rho_{N(A)}, \mathcal{M}) \leq \min \left\{ \min_{\mathcal{P}} K_{\text{DI,dev}}^{\text{iid}}(\rho_{\mathcal{P}(N(A))}), \min_{\mathcal{P}} \inf_{\{\sigma_{\mathcal{P}(N(A))}, \mathcal{L}\} = \{\rho_{\mathcal{P}(M(A))}, \mathcal{M}\}} K_{\text{DD}}(\sigma_{\mathcal{P}(N(A))}) \right\}, \quad (72)$$

where \mathcal{P} is any nontrivial partition of the set of systems A_1, \dots, A_N .”

It should now be as follows (sign \times exchanged for the comma and without a bracket]):

Proposition 3. “For any N -partite quantum behavior $(\rho_{N(A)}, \mathcal{M})$ there is

$$K_{\text{DI,dev}}^{\text{iid}}(\rho_{N(A)}, \mathcal{M}) \leq \min \left\{ \min_{\mathcal{P}} K_{\text{DI,dev}}^{\text{iid}}(\rho_{\mathcal{P}(N(A))}), \min_{\mathcal{P}} \inf_{\{\sigma_{\mathcal{P}(N(A))}, \mathcal{L}\} = \{\rho_{\mathcal{P}(M(A))}, \mathcal{M}\}} K_{\text{DD}}(\sigma_{\mathcal{P}(N(A))}) \right\}, \quad (72)$$

where \mathcal{P} is any nontrivial partition of the set of systems A_1, \dots, A_N .”

The errors listed here, and corrected typographical errors, do not affect the results and proofs in our paper. An updated version has been made available [10].




S.D. acknowledges M. E. Shirokov for pointing out that the dual bound is redundant.

- [1] N. Davis, M. E. Shirokov, and M. M. Wilde, Energy-constrained two-way assisted private and quantum capacities of quantum channels, *Phys. Rev. A* **97**, 062310 (2018).
 [2] J. Ribeiro, G. Murta, and S. Wehner, Fully device-independent conference key agreement, *Phys. Rev. A* **97**, 022307 (2018).

- [3] M. Farkas, M. Balanzó-Juandó, K. Łukanowski, J. Kołodyński, and A. Acín, Bell Nonlocality Is Not Sufficient for the Security of Standard Device-Independent Quantum Key Distribution Protocols, *Phys. Rev. Lett.* **127**, 050503 (2021).
 [4] U. Maurer and S. Wolf, The intrinsic conditional mutual information and perfect secrecy, in *Proc. 1997 IEEE Symposium on Information Theory (Abstracts)* (1997), p. 88.

- [5] U. Maurer and S. Wolf, Unconditionally secure key agreement and the intrinsic conditional information, *IEEE Trans. Info. Theor.* **45**, 499 (1999).
- [6] A. Acín, S. Massar, and S. Pironio, Efficient quantum key distribution secure against no-signalling eavesdroppers, *New J. Phys.* **8**, 126 (2006).
- [7] M. Winczewski, T. Das, and K. Horodecki, Limitations on a device-independent key secure against a nonsignaling adversary via squashed nonlocality, *Phys. Rev. A* **106**, 052612 (2022).
- [8] E. Kaur, M. M. Wilde, and A. Winter, Fundamental limits on key rates in device-independent quantum key distribution, *New J. Phys.* **22**, 023039 (2020).
- [9] A. Philip, E. Kaur, P. Bierhorst, and M. M. Wilde, Multipartite intrinsic non-locality and device-independent conference key agreement, *Quantum* **7**, 898 (2023).
- [10] K. Horodecki, M. Winczewski, and S. Das, Fundamental limitations on the device-independent quantum conference key agreement, [arXiv:2111.02467](https://arxiv.org/abs/2111.02467).

Limitations on a device-independent key secure against a nonsignaling adversary via squashed nonlocality

Marek Winczewski ^{1,2,*}, Tamoghna Das ², and Karol Horodecki ^{3,2,4}

¹*Institute of Theoretical Physics and Astrophysics, National Quantum Information Centre, Faculty of Mathematics, Physics and Informatics, University of Gdańsk, Wita Stwosza 57, 80-308 Gdańsk, Poland*

²*International Centre for Theory of Quantum Technologies, University of Gdańsk, Jana Bażyńskiego 1A, 80-309 Gdańsk, Poland*

³*Institute of Informatics, National Quantum Information Centre, Faculty of Mathematics, Physics and Informatics, University of Gdańsk, Wita Stwosza 57, 80-308 Gdańsk, Poland*

⁴*School of Electrical and Computer Engineering, Cornell University, Ithaca, New York 14850, USA*



(Received 30 September 2020; revised 2 August 2022; accepted 18 October 2022; published 29 November 2022)

We initiate a systematic study to provide upper bounds on device-independent keys, secure against a nonsignaling adversary (NSDI). We employ the idea of “squashing” on the secrecy monotones and show that squashed secrecy monotones are the upper bounds on the NSDI key. Our technique for obtaining upper bounds is based on the nonsignaling analog of quantum purification: the complete extension. As an important instance of an upper bound, we construct a measure of nonlocality called “squashed nonlocality.” Using this bound, we identify numerically a certain domain of two binary inputs and two binary outputs of nonlocal devices for which the squashed nonlocality is zero. Therefore one can not distill a secure key from these nonlocal devices via a considered (standard) class of operations. Showing a connection of our approach to the one in *New J. Phys.* **8**, 126 (2006), we provide, to our knowledge, the tightest known upper bound in the (3,2,2,2) scenario. Moreover, we formulate a security condition, equivalent to known ones, for the considered class of protocols. To achieve this, we introduce a nonsignaling norm that constitutes an analogy to the trace norm used in the security condition of the quantum key distribution.

DOI: [10.1103/PhysRevA.106.052612](https://doi.org/10.1103/PhysRevA.106.052612)

I. INTRODUCTION

Secure key distribution is a process of generation of secret key bits between two distant parties, in presence of an eavesdropper [1–3]. There are four major cryptographic security paradigms developed in the last several decades that provide a background for our investigation. These are (i) *a secret-key agreement scenario* (SKA) [1,2], (ii) *device-dependent security against a quantum adversary* (QDD) [3–7], (iii) *device-independent security against a quantum adversary* (QDI) [5,8–12], and (iv) *device-independent security against a nonsignaling adversary* (NSDI) [13–17]. We have enumerated them in order of increasing power of the eavesdropper. In what follows, we are going to use concepts of SKA paradigm to place upper bounds on the secret-key rate in the NSDI scenario in a manner that is known from the QDD paradigm. Let us then begin with a short reminder of the main ideas behind the aforementioned cryptographic setups.

In the SKA scenario, the parties share marginals of a classical probability distribution $P(ABE)$, respectively. The honest parties (often called Alice and Bob) can process their data by the so-called *local operations and public communication* (LOPC). At the same time, the eavesdropper Eve can listen to public communication and can apply any stochastic map on her data [1,2]. This paradigm is of special interest in context of security of the wireless communication.

The QDD scenario, originating conceptually from the SKA, was introduced at the early stage of quantum cryptography [4,5]. In this paradigm, the three parties share (in the worst case) a subsystem of a joined pure quantum state $|\Psi_{ABE}\rangle$. Alice and Bob can process this state by *Local quantum operations and Classical communication* (LOCC). At the same time, Eve obtains any system which is discarded by Alice and Bob and can perform any quantum operation on her subsystem [18–20]. This scenario has a drawback that Alice and Bob have to trust the inner working of their device: the dimensionality of the state and operations of measurement performed by the device. This problem has been resolved in a much more sophisticated approach of the QDI, *quantum device-independent scenario*. In this paradigm, the honest parties share an untrusted device, described by a joint conditional probability distribution $P(AB|XY)$ originating from a measurement on a quantum state ρ_{AB} : $P(AB|XY) = \text{Tr}(M_{A|X} \otimes M_{B|Y} \rho_{AB})$. Security in this scenario is based solely on statistics of the inputs X , Y and outputs A , B of the device. An eavesdropper is assumed to be restricted by the laws of quantum mechanics. She is therefore bound to hold a purifying system of a ρ_{AB} , i.e., the system E of such a pure state $|\psi_{ABE}\rangle$, that $\text{Tr}_E |\psi\rangle\langle\psi|_{ABE} = \rho_{AB}$.

A. Nonsignaling adversary scenario

In this manuscript, we focus on another branch of key distribution that has emerged in the last two decades, which is the nonsignaling device-independent (NSDI) scenario [13–17,21–23]. This scenario has even more relaxed

*Corresponding author: marek.winczewski@ug.edu.pl

assumptions than QDI. Here, the eavesdropper is restricted only by the nonsignaling condition, i.e., she can not influence statistics of the honest parties in a faster than light manner. Similarly the honest parties can share a possibly supraquantum correlation only constrained by the nonsignaling condition. The advantage of NSDI approach over SKA, QDD, and QDI scenarios is the fact that it assures security even if a new theory replacing quantum mechanics became established, as long as it is nonsignaling. The object shared by Alice, Bob, and Eve is a tripartite nonsignaling device, $P(ABE|XYZ)$, with Z and E being the input and output respectively of Eve's part of the device. On this device, the parties perform some measurements (X, Y) and post-process their output data (A, B) by some LOPC operations, to produce the secure key. This device is assumed to be (in a worst case) created by the eavesdropper who can listen to the public communication, and perform certain operations on her subsystem.

The first NSDI protocol, whose security was proven, was given by Barrett, Hardy, and Kent [13]. The protocol results in a single key bit in the noiseless scenario. Later, lower bounds on the key rate have been derived in Refs. [14,16,22], via several key distillation protocols, under the assumption that eavesdropper attacks each of the subsystems separately. In the presence of a collective eavesdropping attack, it was shown in Refs. [15,17,23], that one can obtain a nonzero key rate under the fully nonsignaling constraint. By fully nonsignaling, we mean that none of the subsystems of a device can signal to each other. More precisely, a device with $2N + 1$ inputs and $2N + 1$ outputs (N for each of the honest parties and one for the eavesdropper) is understood to have $2N + 1$ subsystems none subset k of which can signal to the remaining $2N + 1 - k$.¹ This assumption is vital, because if the device can perform signaling between its subsystems (of one party) [24], then no hash function is known to achieve privacy amplification against the nonsignaling eavesdropper. Moreover, if the device has a memory [25,26], or can signal forward (from one run² of the protocol to the next one) [27], then a wide class of hash functions can be attacked by a nonsignaling Eve. The assumption of full nonsignaling can be achieved by performing measurements in parallel on all of the $2N$ subsystems. We refer to this approach as to parallel measurement model.

The nonsignaling paradigm that allows defining the NSDI scenario became an active field of research since the seminal papers of Rastall [28], Khalfin, and Tsirelson [29] as well as Popescu and Rohrlich [30] (for a recent review on Bell nonlocality see [8]). Our findings will contribute not only to the aforementioned cryptographic scenarios (NSDI and SKA) but also to the domain of Bell nonlocality. This is because some of the functions that serve as upper bounds on the key rate that we establish in the NSDI scenario, are alternative measures of nonlocality.

¹In what follows, by "device" we mean a single-use device. A device can be used by measuring its input. A single-use device can not be measured more than once. If there is a need to perform multiple measurements on a device, then it will be assumed as a composite device consists of multiple single-use devices.

²By one single run of the protocol, we mean one use of a particular single use device.

B. Motivation

In the NSDI scenario described above, mainly the lower bounds on the key rate has been considered [13–17,21–23]. For the upper after seminal result given in Ref. [16] based on intrinsic information, upper bounds were not studied systematically until recently (an upper bound based on intrinsic information has been proposed in parallel to the approach presented in this work in Ref. [31]). In contrast, if one considers the QDD scenario, both lower bounds [18,32–34], and upper bounds on the secure key rate are well known. Indeed, the upper bounds in this scenario were studied both in the context of quantum states [19,20,34–38] and quantum channels [39–41] (see also Refs. [42–46] in this context). Similarly in the case of QDI scenario, after seminal upper bound of Refs. [31,47], a sequence of other proposals were provided recently [48–52]. Some of the upper bounds in QDD and QDI scenario [34,36–40,48–50] are based on the entanglement measure called "squashed entanglement" [34]. A welcome feature of this measure is that it is an additive function, i.e., one avoids regularization like it is the case for the relative entropy of entanglement [19,20,45,46,53]. We aim at both constructing upper bounds in the NSDI scenario and introducing alternative measures of nonlocality. Although the analog of relative entropy—the "strength of nonlocality proof" [54] (also called relative entropy of nonlocality [55]) has been constructed, no analog of squashed entanglement was known in the realm of nonlocality (for the parallel, and different approach see Ref. [31]). In our approach to the problem, we are guided by an analogy between entanglement and nonlocality. Interestingly the measure which we construct is, up to maximization over the inputs of the honest parties, equal to the one implicitly considered in Ref. [16]. It is however differently formulated, as we use the notion of a complete extension [56] to formalize it. Moreover, we prove that our measure is a convex function of the devices of the honest parties, what allows for the use of the convexification technique (that we formulate) for finding the numerical upper bound. We will see that this reformulation is fruitful for studying properties of this upper bound, which we call here "the squashed nonlocality."

II. MAIN RESULTS

In this manuscript, we construct upper bounds on the NSDI key rate, distillable via (i) direct measurement, changing device into a distribution followed by (ii) Local operations and Public communication (denoted together as MDLOPC operations). Aiming at upper bounds, we study the scenario in which the shared device consists of N independent and identically distributed (iid) copies of a nonsignaling device $P(AB|XY)$. We define a wide class of secrecy quantifiers taken from the so-called SKA (secure key agreement) model [2]. One such quantifier, we call the *squashed nonlocality*, as we define it in analogy to squashed entanglement [34], however, in the realm of nonsignaling devices. We then show that the squashed nonlocality serves as an upper bound on the key distilled by MDLOPC operations. It is important to note that almost all of the secure key distillation protocols in QDI and NSDI, proposed so far, belong to the MDLOPC class of operations (see however recent proposal [57]). Therefore our

bounds, on the amount of key, bound from above the key rate achieved by a wide class of practical protocols.

A. Family of nonlocality measures as upper bounds

One of our achievements is a construction of upper bound on the secret key in the NSDI scenario that is in an addition a (nonfaithful) measure of nonlocality. Informally, the squashed nonlocality $\mathcal{N}_{\text{sq}}(P)$, of a bipartite nonsignaling device $P := P(AB|XY)$ is given by

$$\begin{aligned} \mathcal{N}_{\text{sq}}(P) &:= \widehat{I}(A : B \downarrow E)_{\mathcal{E}(P)(ABE|XYZ)} \\ &= \max_{x,y} \min_z I(A : B \downarrow E)_{(\mathcal{M}_{x,y}^F \otimes \mathcal{M}_z^G) \mathcal{E}(P)(ABE|XYZ)}, \quad (1) \end{aligned}$$

where $\mathcal{E}(P)(ABE|XYZ)$ is the complete extension of the device P [56] and $I(A : B \downarrow E)_{P(ABE)}$ is the intrinsic information of a distribution $P(ABE)$ [58,59]. Furthermore, the honest parties choose inputs x, y (for a full direct measurement $\mathcal{M}_{x,y}^F$), while the eavesdropper is allowed to perform a more general measurement \mathcal{M}_z^G that contains in particular probabilistic mixing of input choices.

The squashed nonlocality, as we prove, possesses many properties of those desired for a measure of nonlocality such as convexity and additivity. As we show the above function is only an example of an upper bound that can be introduced using our approach. The other function that we study in this paper to be lifted from the SKA to the NSDI scenario are mutual information and conditional mutual information.

We note, however, that the above function can be equivalently formulated in a way considered implicitly in Ref. [16] by A. Acin, S. Massar, and S. Pironio (AMP). Consider a function $I_{\text{AMP},(x,y)}$:

$$\begin{aligned} I_{\text{AMP},(x,y)}(P(AB|X=x, Y=y)) \\ := \inf_{\{p(E=e), P(ABE=e|X=x, Y=y)\}} I(A : B \downarrow E)_{P(ABE|XY)}, \quad (2) \end{aligned}$$

where $P(ABE|XY) = p(E=e)P(AB, E=e|XY)$, and the infimum is taken over all ensembles $\{p(E=e), P(AB, E=e|XY)\}$ of the device $P(AB|X, Y) = \sum_e p(E=e)P(AB, E=e|X, Y)$. The equivalence can be established as follows for a device $P \equiv P(AB|XY)$:

$$\max_{(x,y)} I_{\text{AMP},(x,y)}(P) = \mathcal{N}_{\text{sq}}(P). \quad (3)$$

This fact, along with our proof of the convexity of \mathcal{N}_{sq} leads to the tightest known bound in the scenario (3,2,2,2) (see Fig. 5), i.e., with three inputs for one party, binary inputs for the other and binary outputs for both (for the proof of Eq. (3) and consequences of it see Sec. VI).

We provide a method of generating tighter (though possibly harder to compute) upper bounds. Indeed, in defining the squashed nonlocality, we used the secrecy monotone called *intrinsic information*. The nonfaithfulness³ of the squashed nonlocality is therefore due to the property inherited from the classical intrinsic information that can be zero for correlated distribution. One can, however, use some other quantifiers of

³The property of nonfaithfulness of a measure of nonlocality means that the measure is zero for some nonlocal behaviors.

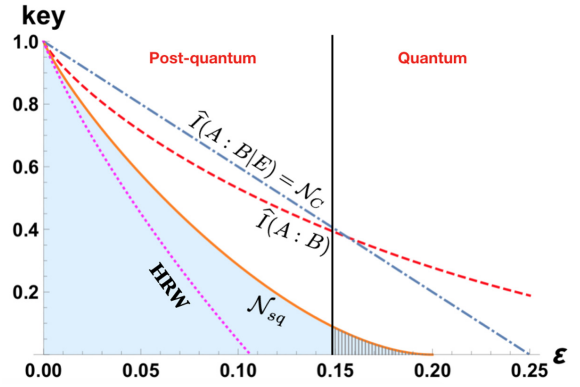


FIG. 1. Plot of several secrecy quantifiers $\widehat{M}(A : B|E)$, as an upper bound on $K_{DI}^{(\text{fid})}$, for a bipartite binary input-output device lying on the isotropic line. The dashed red line represents squashed mutual information $\widehat{I}(A : B)_{P_{\text{iso}}}$. The straight blue line represents the nonlocality cost, as well as the squashed conditional mutual entropy $\widehat{I}(A : B|E)_{\mathcal{E}(P_{\text{iso}})}$, over the complete extension $\mathcal{E}(P_{\text{iso}})$ of the given device P_{iso} . The solid orange line represents the upper bound on the squashed nonlocality \mathcal{N}_{sq} , which is the lower convex hull of the several other upper bounds on \mathcal{N}_{sq} . The dotted pink curve (HRW) corresponds to the lower bound achieved by Hänggi, Renner, and Wolf's protocol [17].

secret correlations, e.g., the so-called *reduced intrinsic mutual information*, which also leads to an upper bound. Due to an analogy between the entanglement and nonlocality, the upper bounds we provide here are also measures of nonlocality, and as such, can be studied independently.

Furthermore, we notice that our approach can be readily modified in order to construct upper bounds for a wider class of protocols in which one of the inputs of the honest parties is not announced [14]. This can be done by changing $\max_{x,y} \min_z$ to $\max_y \min_z \max_x$ in Eq. (1), what reflects the action of the parties in the latter scenario (only Bob announces his inputs).

B. MDLOPC-bound nonlocality

Using the bound, we then obtain numerically a region of nonlocal two binary input and two binary output, (2,2,2,2) devices, from which *no key* can be distilled via MDLOPC operations. These are the “isotropic” mixtures of the devices, namely, the Popescu-Rohrlich (PR) box and the box complementarity to it, the anti-PR box when the admixture of the PR box is less than 80%. Notably, this result implies that in parallel measurement model, when the same measurement on each device is performed, nonlocality does not imply secrecy. Indeed, quantum devices with mixture of PR box more than 75% exhibit nonlocality, that is, they violate the CHSH inequality [60], while as we show, all the devices below 80% have zero key distillable by MDLOPC protocols. We compare also the upper bound via nonsignaling squashed nonlocality for isotropic devices with the lower bound on the key rate taken from [17] (see Fig. 1). The lower and the upper bounds come pretty close for the state close to the PR box.

We note here that in Ref. [14] a protocol for distillation of private key from isotropic devices were given which is out

of MDLOPC class: one of the parties do not announce the input from used to generate the key. There also it was shown that the so called *intrinsic information* is zero when both the parties announce their inputs after measurements. Our bound does not extends straightforwardly to this scenario, as in our case, Eve knows that Alice and Bob draw key from *single* pair of inputs. However it indicates that keeping one of the inputs used for generating key secret, is crucial for nonzero key rate in the nonsignaling adversary scenario.

This indication is confirmed by recent result given in Ref. [50] for the case of device-independent quantum key distribution with quantum adversary. There, a broader notion of protocols is considered, also called “standard.” These are protocols during which for generation of the key each device is measured by a pair of inputs ($X = x, Y = y$) with probability $p(x, y)$ drawn in i.i.d manner, an announced before post-processing the output key rate. It is shown there, that such protocols admits an upper bound $\sum_{x=0, y=0}^{1,1} p(x, y)I(A : B \downarrow E, xy)$, i.e., the intrinsic information [58,59] averaged over choices of the inputs. Moreover it is argued, that there exist nonlocal devices (violating CHSH inequality) for which the latter upper bound is zero. This implies that no such a “standard” protocol is able to achieve nonzero key rate in the case of quantum adversary.

In a similar way, we show the MDLOPC-bound nonlocality in the (3,2,2,2) scenario [5,16]. In the latter, one party has inputs $x \in \{0, 1, 2\}$ and the other $y \in \{0, 1\}$. The inputs $x \neq 0, y$ are used for testing the value of the CHSH inequality [60], while the pair $x = 0, y = 0$ is used for generation of the raw key. The fact that distributions with isotropic parameters $p \in [0.7071, 0.8284]$ are nonlocal but no key can be distilled from them in the latter scenario was left open in Ref. [16]. Showing the equivalence given in Eq. (3) and the fact that \mathcal{N}_{sq} upper bounds the distillable key, we close the mentioned open problem, by confirming that no key can be obtained by a protocol drawing key from a single pair of settings $x = 0$ and $y = 0$. The obtained results are shown in Fig. 5.

C. Analogies between different cryptographic paradigms

We finally compare the proposed security criteria with the previously known ones [15,17,23,24,63,64], and prove their equivalence. In the case of quantum mechanics, the power of eavesdropper is fully described by system of the honest parties through the so-called *purification*. However, it is known that there is no analog of the quantum purification in the realm of devices [65,66]. To overcome this problem, we have used a recently introduced notion of *complete extension* [56], to describe eavesdropper’s power. The complete extension, $\mathcal{E}(P)(ABE|XYZ)$, of the shared device $P(AB|XY)$, is the worst-case extension that Eve can share with the honest parties. It is the worst case in the sense that it gives the eavesdropper an ultimate power as compared to quantum purification does in QDD and QDI scenarios. Indeed, the complete extension gives access to all possible ensembles of the device of the honest parties, when randomizing input and post-processing channel is applied on the extended part. It implies, as we show in detail, that this structural approach is equivalent to the one proposed in Ref. [17].

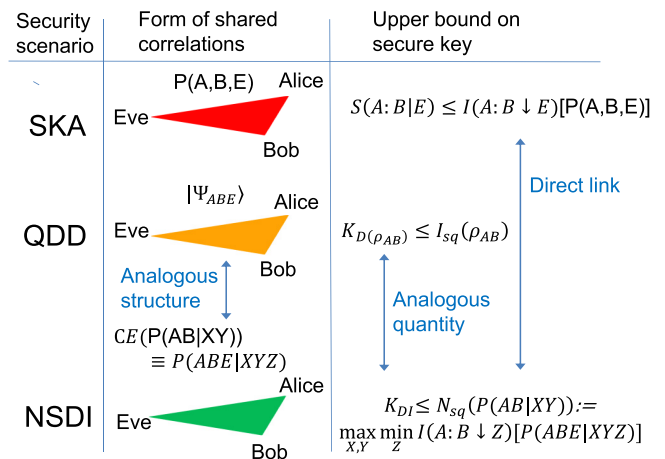


FIG. 2. Summary of part of the results which contribute to the analogy between security paradigms: SKA scenario where distributions are processed, and one of the upper bounds is the intrinsic information $I(A : B \downarrow E)$, QDD protocol, where the shared pure state is processed, and distillable key K_D is upper bounded (among others) by the measure “squashed entanglement” I_{sq} [61,62]. We reformulate NSDI paradigm so that it bases on the complete extension, $\mathcal{E}(P(AB|XY))$, of a device (conditional distribution) $P(AB|XY)$, introduce an analog of intrinsic information and squashed entanglement called “squashed nonlocality.”

We have further introduced another criterion of security, based on an operational distance measure between nonsignaling devices—the nonsignaling norm (*NS norm*) analogous to the trace norm in quantum mechanics (related to the one given in Ref. [67]). We have also proved equivalence between our criteria and the two proposed so far in Refs. [11,15,23] and [17,24,63,64], respectively. As a byproduct, we have shown that the latter two definitions are equivalent. By proving equivalence of our definition based on $\|\cdot\|_{NS}$ norm to the definition of Refs. [17,24,63,64], we have shown that the former is composable, in a sense given there.⁴ A visualization of some of the main results that contribute to developing a structural analogy between SKA, QDD, and NSDI are presented in Fig. 2.

III. SECURITY DEFINITION IN THE IID SCENARIO

In every DI secure key distillation protocol, the honest parties perform several numbers of test runs to estimate the nonlocal correlation present in the system and a (larger) number of key generation runs to generate the raw key. The raw key is further processed to yield the final key only if the device has passed the test run, i.e., the data are compatible with a sufficiently nonlocal device. Aiming at upper bounds, we study only the performance of the key generation runs. We, therefore, assume that, on the N iid (identical, independently distributed)⁵ copies of the shared device $P(AB|XY)^{\otimes N}$, the honest parties perform full direct measurement $[\mathcal{M}_{x,y}^F]^{\otimes N}$, by

⁴Naturally however, the device can not be reused in composing the protocols due to the threat of the memory attack [68].

⁵For QDI, it is known that any arbitrary device can not be expressed in terms of the IID single use device, but the security proof for

setting $X = x$ (Alice) and $Y = y$ (Bob) at their choice, followed by any composition of classical post-processing of the distribution $P(AB|xy)$, and public communication (denoted as Q). These operations result in a pair of random variables (S_A, S_B) that represents the *key*. That is, on the outputs of the measured device, the honest parties perform an LOPC protocol. An operation performed on a device, that is a composition of the direct measurements and an LOPC operations we call *Measurement on device local operation and public communication operation (MDLOPC)*.

In NSDI scenario Eve collects all the public communication Q , and then post-process her data represented by $\tilde{P}(E|Z, Q)$. She can also perform a wider class of operations than the honest parties, including the general measurement $M_z^G = \sum_{z'} p(z|z') \mathcal{M}_{z'}^F$. This is equivalent to a probabilistic choice of the inputs for direct measurements. She can do so by the general measurement M_z^G , by wiring the output of her local auxiliary device (a dice), that generates a random conditional probability distribution $p(z|z')$, to the input of her part of the device, i.e., Z of $\tilde{P}(E|Z, Q)$. However, the ultimate power of eavesdropping in this scenario is fixed by definition of the class of operations that a hypothetical agent called *distinguisher* could perform. It is assumed that distinguisher has access to *both* the output of the protocol (i.e., the keys of the honest parties) *and* the Eve's device $\tilde{P}(E|Z, Q)$. By his operations, distinguisher should be almost not able to tell apart this so-called "real" device $P^{\text{real}}(S_A, S_B, Q, E|Z)$ from an "ideal" one, i.e., containing perfectly uniform and correlated keys, product with Eve's system.

We can specify now what the key distillation protocol is. A protocol of key distillation is a sequence of MDLOPC operations $\Lambda = \{\Lambda_N\}$, performed by the honest parties on N iid copies of the shared devices. Each of this Λ_N , consists of a measurement stage $\{\mathcal{M}_N\}$, followed by post-processing $\{\mathcal{P}_N\}$, on N iid copies of $P(AB|XY)$. Moreover, for each consecutive, complete extension of N copies of shared devices $\mathcal{E}(P^{\otimes N})(ABE|XYZ)$, the protocol outputs a probability distribution in part of Alice and Bob and a device in part of Eve, which is arbitrarily close to an ideal distribution, satisfies

$$\|P_{\text{out}} - P_{\text{ideal}}^{(d_N)}\|_{\text{NS}} \leq \varepsilon_N \xrightarrow{N \rightarrow \infty} 0. \quad (4)$$

Here $P_{\text{out}} = \Lambda_N(\mathcal{E}(P^{\otimes N}))$. Moreover $\mathbf{A} = A_1 A_2 \dots A_N$, \mathbf{B} , \mathbf{X} and \mathbf{Y} are similarly defined.

The definition of the secret-key rate, based on the notion of the (i) complete extension and (ii) the key distillation protocol, satisfying the proximity in the NS norm security criterion according to the Eq. (15), is given below.

Definition 1. Given a bipartite device $P \equiv P(AB|XY)$ the secret-key rate of the protocol of key distillation Λ_N , on N iid copies of the device, denoted by $\mathcal{R}(\Lambda|P)$ is a number $\limsup_{N \rightarrow \infty} \frac{\log_2 d_N}{N}$, where $\log_2 d_N$ is the length of a secret key shared between Alice and Bob, with $d_N = \dim_{\Lambda}[\Lambda_N(\mathcal{E}(P^{\otimes N}))] \equiv |S_A|$. The device independent key rate

of the *iid* scenario is given by

$$K_{DI}^{(\text{iid})}(P) = \sup_{\Lambda} \mathcal{R}(\Lambda|P), \quad (5)$$

where the supremum is taken over all MDLOPC protocols $\{\Lambda\}$.

Later in this manuscript, we argue that the above definition is equivalent in terms of security to the one adopted earlier [11,15,17,23,24,63,64], which allows us to compare some of the existing lower bounds with the upper bounds that we provide.

IV. SQUASHING PROCEDURE

Let us suppose that $M(A : B|E)$ is a real-valued and non-negative function, with domain in the set of tripartite probability distributions $P(ABE)$, which is an upper bound on secret-key rate $S(A : B|E)$ in SKA cryptographic paradigm [2], i.e., $\forall P(ABE), M(A : B|E) \geq S(A : B|E)$. We will refer to $M(A : B|E)$ as to *secrecy quantifier*. Additionally, if $M(A : B|E)$ is monotonic with respect to LOPC and zero for product distributions, we call it a *secrecy monotone*. Squashing a secrecy monotone will not yield an MDLOPC monotonic quantifier in general. The quantifiers of secret correlations in the NSDI model can be constructed by mapping the tripartite nonsignaling device $R(ABE|XYZ)$ to a joint probability distribution, as given in the definition.

Definition 2. Corresponding to each secrecy quantifiers in SKA model $M(A : B|E)$, we associate a nonsignaling secrecy quantifier $\widehat{M}(A : B|E)$ acting on the tripartite nonsignaling devices:

$$\begin{aligned} \widehat{M}(A : B|E)_{R(ABE|XYZ)} \\ := \max_{x,y} \min_z M(A : B|E)_{(\mathcal{M}_{x,y}^F \otimes \mathcal{M}_z^G)R(ABE|XYZ)}, \end{aligned} \quad (6)$$

where

$$\begin{aligned} (\mathcal{M}_{x,y}^F \otimes \mathcal{M}_{z'}^G)R(ABE|XYZ) \\ = \sum_z p(z|z')R(ABE|X=x, Y=y, Z=z). \end{aligned} \quad (7)$$

If $R(ABE|XYZ) \equiv \mathcal{E}(P)(ABE|XYZ)$ is the complete extension of a bipartite device $P(AB|XY)$, we call $\widehat{M}(A : B|E)_{\mathcal{E}(P)(ABE|XYZ)}$ the nonsignaling squashed secrecy quantifier. If $M(A : B|E)_{R(ABE|XYZ)}$ is a secrecy monotone, we call $\widehat{M}(A : B|E)_{R(ABE|XYZ)}$ a nonsignaling secrecy monotone. Additionally if $R(ABE|XYZ)$ is a complete extension, we call it a nonsignaling squashed monotone.

Here, by $\max_{x,y}$, we mean the maximization over all possible direct measurements, $\mathcal{M}_{x,y}^F \equiv \mathcal{M}_x^F \otimes \mathcal{M}_y^F$ by the honest parties, whereas the \min_z implies that the eavesdropper will try to minimize the function over all possible choices of measurements, direct and general. Optimization over direct measurements involves a fixed input choice, whereas for general measurement, one needs to perform optimization over all possible conditional probability distributions $p(z|z')$. In our MDLOPC key distillation protocol, the eavesdropper can choose her measurement adaptively, based on the public communication variable Q . Hence the causal order of the optimization on the secrecy quantifier is that Alice and Bob first choose their optimal measurements, and then Eve performs

a broad range of cryptographic protocols can be performed via a reduction to IID [69].

her part. This gives her the maximal operational power to reduce the correlations between the honest parties.⁶

The motivation to use the term ‘‘squashed’’ in the above measures, comes from the fact that the definition of squashed entanglement, of an arbitrary quantum state ρ_{AB} , contains an optimization over all possible extensions ρ_{ABE} , where $\text{tr}_E(\rho_{ABE}) = \rho_{AB}$. This arbitrary extension ρ_{ABE} can be obtained from the purification $|\psi\rangle_{ABE}$ of the quantum state [62]. In the analogy of these, here we use the complete extension $\mathcal{E}(P)$, the nonsignaling equivalent of quantum purification, which is the key ingredient to perform an optimization over all possible nonsignaling extensions [56] of a given device P . The secrecy quantifiers, we have used for squashing, are the mutual information $I(A : B)$, the conditional mutual information $I(A : B|E)$, the intrinsic information $I(A : B \downarrow E)$ [58] and the reduced intrinsic information $I(A : B \downarrow\downarrow E)$ [70]. Among them, $I(A : B|E)$, $I(A : B \downarrow E)$ and $I(A : B \downarrow\downarrow E)$ are secrecy monotones. Hence $\hat{I}(A : B|E)$, $\hat{I}(A : B \downarrow E)$ and $\hat{I}(A : B \downarrow\downarrow E)$ are nonsignaling squashed secrecy monotones while $\hat{I}(A : B)$ is an example of a nonsignaling squashed secrecy quantifier.

The inclusions between gray, green, and orange sets in Fig. 3 follow directly from the definition of different classes of functions. Namely, all n-s secrecy measures are necessarily n-s secrecy monotones, and all n-s secrecy monotones are necessarily n-s secrecy quantifiers, but not vice versa. The strictness of the inclusions follows from a trivial example of n-s mutual information (gray area), n-s intrinsic information (orange area), and n-s intrinsic information shifted by a nonzero constant (green area). Analogous relation is true for the squashed version of the aforementioned functions. Nevertheless, the squashing procedure does not imply that the resulting function is automatically a secrecy measure or a secrecy monotone; therefore, the representatives of squashed functions are present in all three sets.

V. GENERIC UPPER BOUND AND THE SQUASHED NONLOCALITY

Below, we use the aforementioned idea of squashing for upper-bounding the secret key in the NSDI scenario with MDLOPC operations.

Theorem 1. The secret-key rate, in the nonsignaling device-independent *iid* scenario achieved with MDLOPC operations, $K_{DI}^{(iid)}$, from a device P , is upper bounded by any nonsignaling squashed secrecy quantifier evaluated for the complete extension of P :

$$\forall_P K_{DI}^{(iid)}(P) \leq \hat{M}(A : B||E)_{\mathcal{E}(P)}, \quad (8)$$

where $P \equiv P(AB|XY)$ is a single copy of a bipartite nonsignaling device shared by the honest parties, and $\mathcal{E}(P) \equiv \mathcal{E}(P)(ABE|XYZ)$ is its complete extension to the eavesdropper’s system.

Proof. For the proof, see Sec. F of Appendix.

Theorem 1, together with Definition 2, establishes a connection between the secret-key rate in the SKA and NSDI

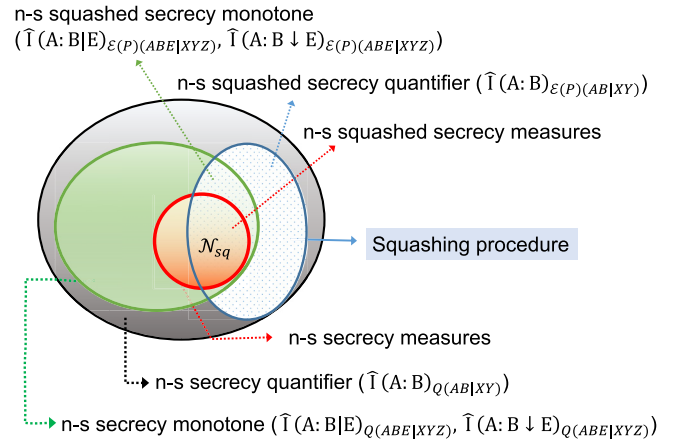


FIG. 3. The relative hierarchy of the squashed function $\hat{M}(A : B||E)$ of any bipartite device $P(AB|XY)$. Any $\hat{M}(A : B||E)$ function, which is positive semidefinite and vanishes for devices that are the product of two local devices, is called a nonsignaling squashed secrecy quantifier, and the set of all such functions is the entire region inside the black ellipse. The quantifiers that are generated from the monotones of the SK paradigm are called *nonsignaling squashed secrecy monotones* and are represented by the green region inside the green ellipse, as a subset of the secrecy quantifiers. $\hat{I}(A : B|E)$, $\hat{I}(A : B \downarrow E)$, and $\hat{I}(A : B \downarrow\downarrow E)$ are the monotones belonging to this category. If a function is additionally monotonic under MDLOPC operations for devices, vanishes for local ones, then we call it the *nonsignaling squashed secrecy measure*, and the set represented by the red region in the figure. Any secrecy quantifier will be called squashed secrecy quantifier if the extension of the device $P(AB|XY)$ has been taken to be the CE $\mathcal{E}(P)(ABE|XYZ)$. The set of such functions are denoted by the dashed blue region. The intersection of the dashed blue region with the green region includes all squashed secrecy monotones, whereas its intersection with the red contains all squashed secrecy measures. The squashed nonlocality \mathcal{N}_{sq} is a particular function from that region, which is depicted as the black dot.

scenario. The novelty of our approach is that not only it connects at least two major security paradigms, but it also opens up a new area of research—to study more tighter upper bounds on the key rate in the NSDI scenario (for parallel, different approach see Ref. [31]). In this paper, we focus on the secrecy monotone called intrinsic information $I(A : B \downarrow E)$. From this secrecy monotone via squashing we construct the so called *squashed nonlocality*, as an upper bound on the NSDI key. We then prove several important properties of squashed nonlocality, which promotes it as a measure of nonlocality. Secrecy monotone called the reduced intrinsic information $I(A : B \downarrow\downarrow E)$, provide a tighter bound on the key rate in the SKA scenario, as $I(A : B \downarrow\downarrow E) \leq I(A : B \downarrow E)$ for tripartite probability distribution $P(ABE)$ [71]. Hence we open a possibility to study even tighter upper bound on the $K_{DI}^{(iid)}$, upon squashing the $I(A : B \downarrow\downarrow E)$. We focus now on the definition of the aforementioned *squashed nonlocality*.

Definition 3. The squashed nonlocality $\mathcal{N}_{sq}(P)$, of a bipartite nonsignaling device $P := P(AB|XY)$ is

$$\begin{aligned} \mathcal{N}_{sq}(P) &:= \hat{I}(A : B \downarrow E)_{\mathcal{E}(P)(ABE|XYZ)} \\ &= \max_{x,y} \min_z I(A : B \downarrow E)_{(\mathcal{M}_{x,y}^E \otimes \mathcal{M}_z^C) \mathcal{E}(P)(ABE|XYZ)}, \end{aligned}$$

⁶One can also consider the reverse order of optimization, but that opens up a different, uncommon paradigm of key distillation.

where $\mathcal{E}(P) := \mathcal{E}(P)(ABE|XYZ)$ is the complete extension of the device P [56].

We note here, that the above definition is tuned to the definition of $K_{DI}^{(iid)}$. The order of the $\max_{x,y}$ and \min_z stems from the fact that we consider the scenario of key distillation in which Eve knows x, y beforehand. In our case, the inputs (x, y) are fixed before the beginning of the protocol, but in general it could be announced during the protocol's execution. This is important point, as alternative protocols exist in which only one party announces the inputs, and the key is distilled from output of all the inputs [14]. In the latter case, a positive key rate can be obtained even from the quantum *isotropic devices* in the scenario of two binary inputs and two binary outputs, while in the scenario which we consider where both inputs are known to the eavesdropper no positive lower bound on the key rate is known. It is possible that the upper bounds on the protocols such as those from Ref. [14] where x is not announced, are provided in terms of the squashed nonlocality where however $\max_y \min_z \max_x$ appears in front instead of $\max_{xy} \min_z$.

From the definition of a complete extension of a device (see Ref. [56]), we know that in order to construct it, one needs to identify all possible so-called *minimal ensembles* of the device. For example, in the polytope of two binary input and two binary output devices (2,2,2,2), a device lying on the isotropic line between Popescu-Rohrlich and Tsirelson's one⁷ has up to 354 minimal ensembles (achieved for the Tsirelson's device). However, *a priori*, there are 880 946 of ensembles that can be potentially minimal [56]. Hence, obtaining all possible minimal ensembles, and therefore finding out the complete structure of the CE may be an arduous task. However, we observe that to obtain a nontrivial upper bound on the \mathcal{N}_{sq} , not the whole complete extension has to be even known.

We collect below certain properties of the above measure. Some of them are used in what follows, and some of them are of independent interest in the context of Bell nonlocality.

Proposition 1. Besides being nonfaithful, the squashed nonlocality satisfies the following properties:

- (1) Positive. It is a non-negative real function of bipartite nonsignaling devices, and equal to zero for local devices.⁸
- (2) Monotonic with respect to MDLOPC class of operations.
- (3) Convex with respect to the mixture of devices.
- (4) Superadditive over joint nonsignaling devices.
- (5) Additive for product devices.
- (6) Subextensive. $\mathcal{N}_{sq}(P) \leq \log_2(\min\{d_A, d_B\})$.

Proof. For the proof, see Sec. G of Appendix. See also the discussion in Sec. VI.

Note. On the completion of the main results (preliminary version of this paper) contained in Secs. C-F, I and J in Appendix, we have noticed the preprint of the paper by E. Kaur, M. Wilde, and A. Winter [31] also related to upper bounds on

device independent key. The proofs of monotonicity, subadditivity and additivity over tensor product devices (see Secs. G 4 and G 5 of Appendix), were inspired by the analogous result for the squashed intrinsic nonlocality presented there.

Calculating \mathcal{N}_{sq} for an arbitrary bipartite device P is a nontrivial task, but we can use the convexity of this measure to simplify the procedure of finding an upper bound of it. Positivity, monotonicity, and additivity of squashed nonlocality lead to the following Corollary.

Corollary 2. The squashed nonlocality is a measure of nonlocal correlation of the bipartite device P .

We describe now, how to use the convexity of the squashed nonlocality (this technique proposed in this manuscript proved already useful in context of upper bounds on the secure key in QDI scenario [51]). Consider any set of functions $\mathcal{F} = \{F_i(P)\}$, that are convex w.r.t. the mixture of devices, each of which upper bounds the squashed nonlocality $F_i(P) \geq \mathcal{N}_{sq}(P), \forall i$. Then the lower convex hull (LCH) of \mathcal{F} denoted as $F(P) (\equiv \text{LCH}(\mathcal{F}))$ upper bounds $\mathcal{N}_{sq}(P)$, i.e., $\mathcal{N}_{sq}(P) \leq F(P)$, as a consequence of property (3). To exemplify the above convexification process, let $\mathcal{F} = \{\widehat{I}(A : B)_{P(AB|XY)}, \widehat{I}(A : B|E)_{\mathcal{E}(P)(ABE|XYZ)}\}$, then $\mathcal{N}_{sq}(P) \leq F(P) \equiv \text{LCH}(\widehat{I}(A : B)_{P(AB|XY)}, \widehat{I}(A : B|E)_{\mathcal{E}(P)(ABE|XYZ)})$. This fact is used in order to construct Fig. 1: the orange curve is, in fact, a convex hull of several upper bounds that are incomparable with each other.

VI. QUANTITATIVE RESULTS

In Fig. 1, we construct numerically an upper bound on the \mathcal{N}_{sq} , with the help of above specified convexification procedure. We also draw several other squashed quantifiers for the set of (2,2,2,2) devices, lying in the isotropic line, i.e., $P_{iso} = (1 - \varepsilon)PR + \varepsilon\overline{PR}$. Where PR is the famous Popescu-Rohrlich box [30], and \overline{PR} is the anti-PR box.⁹ The nonfaithfulness of our measure, \mathcal{N}_{sq} is visible from the numerical results. The orange curve is the upper bound on \mathcal{N}_{sq} , and we have found that the bound reaches 0 for $\varepsilon = 0.2$ (it remains equal to 0 for $\varepsilon \in [0.2, 0.25]$ due to the convexity of the measure). This is since, in MDLOPC protocol, Eve can perform adaptive general measurements and post-process her output through a classical post-processing channel to reduce the correlations between Alice and Bob. In the range $\varepsilon \in [0.2, 0.25]$, corresponding to each input (x, y) of the honest parties, we have found a measurement and a post-processing channel on Eve, which partitioned the device into an ensemble of product distributions. This proves that there exists nonlocality which can not be turned into security via MDLOPC protocols. Interestingly, these devices are quantum realizable ones. One can conjecture that even the general operation, including the so-called "wirings"¹⁰ can not help in distilling key out of these isotropic devices. Indeed, using wirings that is necessary for the key to be nonzero,

⁷By Tsirelson's device, we mean a one attaining maximal value of violation of the CHSH inequality [60] among quantum (2,2,2,2) devices [72].

⁸By local we mean devices which possess a local hidden variable model [8].

⁹Anti-PR box is a binary input output device, satisfy $\overline{PR}(ab|xy) = \frac{1}{2}\delta_{a \oplus b, xy}, \forall a, b, x, y \in \{0, 1\}$ [73].

¹⁰Operations of feeding input of one device with the output of the other.

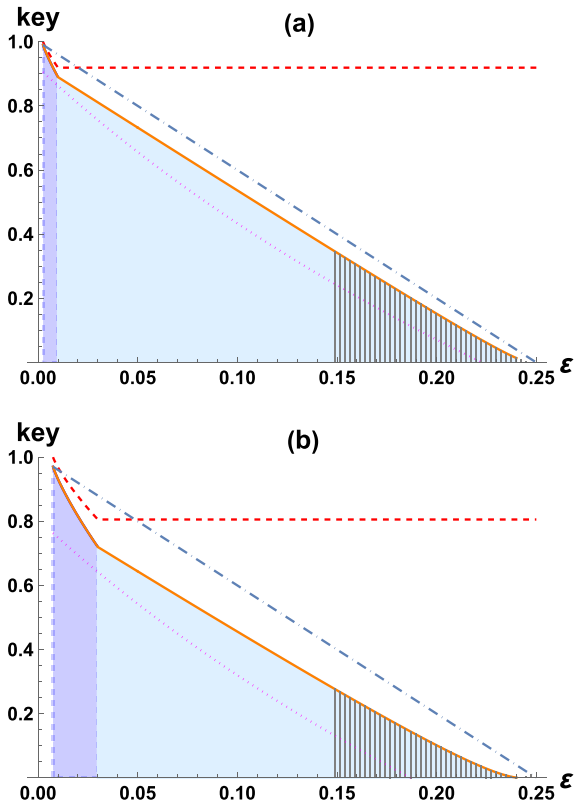


FIG. 4. Plot of several nonsignaling secrecy quantifiers $\widehat{M}(A : B|E)$, as an upper bound on secure key rate $K_{DI}^{(iid)}$, for the bipartite binary input output device P_{HRW} given in Eq. (9) (also in Ref. [17]). The parameters used to draw plot (a) are $\delta = 0.01$, $\epsilon = \frac{1}{16}(3.04 + 12\epsilon)$, and for plot (b) we used $\delta = 0.03$, $\epsilon = \frac{1}{16}(3.12 + 12\epsilon)$. The dashed red line corresponds to the nonsignaling squashed mutual information $\widehat{I}(A : B)_{P_{HRW}}$. The blue straight line represents the nonlocality cost, as well as the nonsignaling squashed conditional mutual information $\widehat{I}(A : B|E)_{\mathcal{E}(P_{HRW})}$ over the complete extension $\mathcal{E}(P_{HRW})$ of the given device P . The solid orange line represents the upper bound on the nonsignaling squashed nonlocality \mathcal{N}_{sq} which is in fact the lower convex hull of the several other upper bounds on \mathcal{N}_{sq} . The magenta dotted line is the key rate $\mathcal{R}(P|_{P_{HRW}})$ of the protocol design by Hänggi, Renner and Wolf [17]. The region with black stripes corresponds to the devices that are quantum realizable ones.

which implies that we enter to some extent the general scenario of key distillation for which there is a wide class of attacks by employing the forward signaling attacks found in Ref. [25,26].

In Figs. 4(a) and 4(b), we plot upper bounds on \mathcal{N}_{sq} for several other sets of (2,2,2,2) devices (nonisotropic), parameterized as in equation (9). In fact, the parametrization that we use is the same as in Ref. [17] as we want to compare our upper bound with the lower bound obtained therein. One can see that there exists some region of nonlocal correlation [Figs. 4(a) and 4(b)], which can be simulated by a quantum device and for which the lower bound obtained by [17,63] is positive, and therefore the secret-key can be generated. As we observe and \mathcal{N}_{sq} is also nontrivial and close to the lower bound in the case considered here. We address the interested reader to Sec. J of Appendix, where more plots are

provided.

$$P_{HRW}(ab|xy)$$

		x		y	
		0	1	0	1
y	b				
	a	0	1	0	1
= 0	0	$\frac{1}{2} - \frac{\delta}{2}$	$\frac{\delta}{2}$	$\frac{3}{8} - \frac{\epsilon}{2}$	$\frac{1}{8} + \frac{\epsilon}{2}$
	1	$\frac{\delta}{2}$	$\frac{1}{2} - \frac{\delta}{2}$	$\frac{1}{8} + \frac{\epsilon}{2}$	$\frac{3}{8} - \frac{\epsilon}{2}$
1	0	$\frac{3}{8} - \frac{\epsilon}{2}$	$\frac{1}{8} + \frac{\epsilon}{2}$	$\frac{1}{8} + \frac{\epsilon}{2}$	$\frac{3}{8} - \frac{\epsilon}{2}$
	1	$\frac{1}{8} + \frac{\epsilon}{2}$	$\frac{3}{8} - \frac{\epsilon}{2}$	$\frac{3}{8} - \frac{\epsilon}{2}$	$\frac{1}{8} + \frac{\epsilon}{2}$

(9)

We note here, that the result presented in Fig. 1 exhibits that in our approach the nonlocality measure based on the intrinsic information can be nonfaithful, i.e., zero for some nonlocal devices. This is inherited after the intrinsic information, which is known to be zero for some tripartite distributions in spite of the fact that the latter are not of the product form $P(A|E)P(B|E)$. We note here, that [74] claimed, that the intrinsic information is nonzero for all devices violating Bell inequality (cf. Ref. [16]). We reformulate the result of [74] as follows:

$$\forall_{P(ABE)} \forall_{\Lambda: E \rightarrow E'} \exists_{(x,y)} P(ABE') \neq P(AE')P(BE') \Leftrightarrow I(A : B|E') > 0. \quad (10)$$

The above implies that if we can adjust the inputs *after* the attack by Eve represented by the map Λ is performed, we will obtain nonzero conditional information. This implies also nonzero intrinsic information as the map can realize the infimum over such maps in the definition of the latter. However this approach does not fit the usual cryptographic scenario: it is that Eve is listening to Alice and Bob and adjusts her measurement to their announcement and not vice versa. Owing to that observation, one should consider the inputs (x, y) to be chosen *before* the map Λ of the attack is performed. This happens, e.g., whenever the input is fixed from advanced as we assume, or when it is announced right after has been made. This change in the paradigm has important consequences. What both Ref. [50] and our result implies goes with no contradiction with the above, as is based on the following fact:

$$\exists_{P(ABE) \neq P(AE)P(BE)} \forall_{(x,y)} \exists_{\Lambda: E \rightarrow E'} I(A : B|E') = 0. \quad (11)$$

Indeed, in the case of the above mentioned quantitative results, we adjust the measurement and post-processing of Eve to the inputs of the honest parties.

Finally we note, that a more common approach to key distribution in device independent scenarios is such that, following A. Ekert [5], one of the honest parties has one more input, which is use to key generation. This so called (3,2,2,2) scenario has been considered in Ref. [16] in context of a nonsignaling adversary, along with a protocol of key distillation and an upper bound on it in terms of the intrinsic information. To see the relation between our results with that of Ref. [16], we show the Eq. (3), that is $\max_{(x,y)} I_{AMP,(x,y)} = \mathcal{N}_{sq}$ (see Sec. G 1 of Appendix). We note here, that by this fact, we show that the bound given in $I_{AMP,(x,y)}$ hold for any MD-LOPC protocol using inputs (x, y) for generating key, closes the problem left open in Ref. [16] concerning possibility of

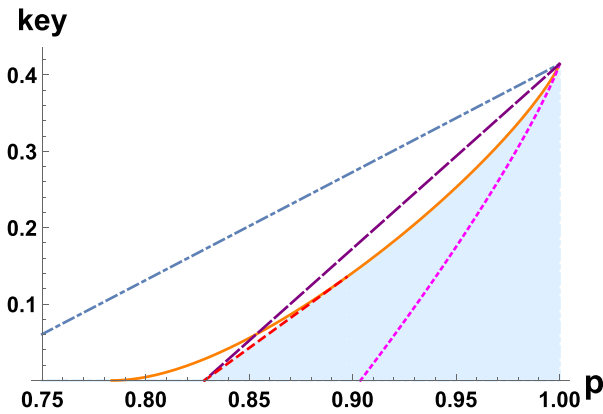


FIG. 5. Plot of nontrivial upper bound on the secret-key rate $K_{DI}^{(iid)}$ given by \mathcal{N}_{sq} , of $\mathcal{P}_{AMP}(ab|xy)$ given in Eq. (J13) (see Appendix), by the blue shaded region under the orange solid line and a red dashed line. The red dashed line is the (segment of) lower convex hull of the orange solid curve and the purple “big-dashed” straight line. The solid orange line is obtained by the lower convex hull of several upper bounds of \mathcal{N}_{sq} , with the help of Eq. (339). Blue dashed-dotted line is the squashed conditional mutual information $\widehat{I}(A : B|E)_{\mathcal{E}(\mathcal{P}_{AMP})}$. The magenta dotted line is the lower bound on the key rate, whereas the purple big-dashed line is the upper bound on intrinsic information of the eavesdropping strategy used in Ref. [16]. We observe that the convexification technique resulting in the convex-hull bound allows to obtain tighter upper bound on \mathcal{N}_{sq} , and therefore the tightest known upper bound on the secret-key rate in the nonsignaling scenario.

key distillation from states that violate CHSH inequality but have zero $I_{AMP,(x,y)}$ bound.

As we will see this fact proves useful, since we have shown that \mathcal{N}_{sq} is convex. This will enable us to use the convexification method to obtain tighter upper bounds. Following [16], as a noise model, we consider the isotropic state $p|\psi_+\rangle\langle\psi_+|_{AB} + \frac{(1-p)}{4}\mathbb{1}_{AB}$ with $|\psi_+\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$ with $p \in [0, 1]$. The bound outperforms existing one [16] in a wide range of a parameter p (see the orange curve in Fig. 5). In general, however, it is incomparable (for the whole range of parameters) with the one given in Ref. [16]. It is possible that a more refined optimization procedure, involving all the extremal points of the nonsignaling polytope in (3,2,2) scenario, would provide a tighter bound. It is, however, computationally involved.

Moreover the convex hull of the bound given in Ref. [16] and ours [which we got by convexification of two upper bounds, given in Eq. (339)], is also an upper bound on the distillable key. This is because \mathcal{N}_{sq} is a lower bound to both upper bounds, and is convex. Hence is less than the convex hull of the latter two bounds. This gives to our knowledge the tightest bound known so far in this scenario.

VII. REPHRASING THE KEY RATE OF THE SKA MODEL

In the SKA model of key distillation, the honest parties and the eavesdropper share a joint probability distribution $P(ABE)$. The task of the honest parties is to perform LOPC operation to distill a secret key, in such a manner that the eavesdropper’s knowledge about the key remains negligibly

small. In the following lines, we propose an alternative definition of the key rate in the aforementioned scenario and prove that it is equivalent to the definition of the secret-key rate introduced in the literature [1,2,34,75]. Rephrasing, the definition of the secret key in the SKA model to the form similar to the one used in quantum cryptography serves not only as a connection between different cryptographic paradigms. Indeed, the theorem below, besides being interesting on its own, is a crucial ingredient used to prove Theorem 1, i.e., our main result.

Theorem 2 (Informal). The secret-key rate $S(A : B|E)$ of SKA cryptographic model [1,2,34,75] is equivalent to the following asymptotic expression:

$$S(A : B|E) = \sup_{\mathcal{P}} \limsup_{N \rightarrow \infty} \frac{\log_2 \dim_A[\mathcal{P}_N(P^{\otimes N}(ABE))]}{N}, \quad (12)$$

with security condition

$$\|\mathcal{P}_N(P^{\otimes N}(ABE)) - P_N^{\text{ideal}}\|_1 \leq \delta_N \xrightarrow{N \rightarrow \infty} 0, \quad (13)$$

where $\mathcal{P} = \cup_{N=1}^{\infty} \{\mathcal{P}_N\}$ is a cryptographic protocol consisting of LOPC operations, acting on N iid copies of the classical probability distribution $P(ABE)$, and P_N^{ideal} is the distribution containing ideal secret key, with adequate dimensions.

Proof. For the proof, see Sec. E of Appendix.

The aim of this rephrasing is to show and utilize a connection between the definition of a secret-key rate in the SKA and NSDI scenarios, as it was done in the case of quantum cryptography [34].

The link we have made in the above theorem, is technical, however important in our method for obtaining the upper bound on the key rate in NSDI scenario. We rephrase the security definition of SKA proposed by U. Maurer [2], with the one based on the trace norm $\|\cdot\|_1$. What is crucial in the choice of the latter criterion is the fact that it is equivalent to the NS norm $\|\cdot\|_{NS}$ for tripartite probability distributions.¹¹ We recall here that the security definition in SKA is based on the control of the correlations (in terms of the mutual information) of the random variable of the honest parties with a random variable representing Eve’s knowledge. These correlations should tend to zero for a large number of copies N . Thus, in other words, in the above theorem, we have modified the security criterion of the SKA to an equivalent form, which is more useful for our purpose. We have done so by demanding that the output distribution of the protocol should be close to an ideal one. The ideal is the distribution representing perfectly correlated uniform random variables, of the honest parties close to being product with the variable of Eve, in trace norm distance $\|\cdot\|_1$. As it will appear later, this technical change turns to be useful when we pass to the case of devices because the NS norm of a device is in fact a trace norm of a distribution coming from this device after measurement.

¹¹Note however that $\|\cdot\|_{NS}$ norm applies also to conditional distributions, i.e., devices. Only for devices with unary input, i.e., distributions, it is equivalent to $\|\cdot\|_1$.

VIII. EQUIVALENCE OF THE SECURITY CONDITIONS

In this section, we show the equivalence between two different known definitions of the security of the secret-key in the NSDI scenario via showing that each of them is equivalent to the one proposed by us. Indeed, we show that the security definition proposed by us that bases on the NS norm is equivalent both to the definition that employs secrecy and correctness as well as the so-called distinguisher [17,24,63,64] and the other one given in Refs. [11,15,23].

A. The definition and the properties of the NS norm

In this section, we provide the explicit description of the NS norm that is an important ingredient of our security criterion. The tensor product should be understood as an algebraic tensor product in \mathbb{R}^N space [76]. To measure the closeness between two devices P and P' , we use the newly defined distance measure, the NS norm which reads

$$\|P - P'\|_{\text{NS}} := \sup_{g \in \mathcal{G}} \frac{1}{2} \|g(P) - g(P')\|_1, \quad (14)$$

where \mathcal{G} is a set of certain operations that map a device to probability distributions and $\|\cdot\|_1$ is a variational distance between two distributions. More precisely, operations from \mathcal{G} are convex combinations of operations that can be composed of the following basic ones (i) fetching an auxiliary device that has single input and single output (a *dice*) (ii) connecting the output of a device/dice to the input of a dice/device respectively, called *wirings* (iii) pre-processing the inputs of device(s) (iv) post-processing inputs and outputs of the devices. We call them *generating* operations,¹² and refer to this norm as to *nonsignaling norm*. The set of generating operations \mathcal{G} is a subset of all linear operations \mathcal{L} mapping device to distribution, that were considered in Ref. [67]. Operational characterization of the set \mathcal{L} is interesting, yet, to our knowledge, unresolved task. However, as we show (see Proposition 2), the set $\mathcal{G} \subseteq \mathcal{L}$ has enough power in discriminating between devices, to be used in security definition in place of \mathcal{L} . Indeed, NS norm via Eq. (14) leads to security definition, which is equivalent to the other two already present in literature (Refs. [11,15,23] and [17,24,63,64]). For more detailed discussion, see Sec. C of Appendix.

After the MDLOPC key distribution protocol, the output of the honest parties reduces to a classical-classical-probability distribution, whereas the part shared by Eve still remains a device, of the form $\Lambda_N(\mathcal{E}(P^{\otimes N}))_{S_A, S_B, Q, E|Z}(s_A, s_B, Q, E|Z)$, where s_A and s_B are the instances of the key shared between Alice and Bob. We will denote random variables S_A, S_B for the secret keys in possession of Alice and Bob, whereas Q stands for all possible classical communications between Alice and Bob; E, Z for Eve's output and input (and the lower case letters are for their values). This distribution, which is, in part a probability distribution, and in part a conditional probability distribution, i.e., device. Hence we will refer to it as to "classical-classical-device" (cc-d) distribution

throughout the paper. The $(P_{\text{ideal}}^{(d_N)})_{S_A, S_B, Q, E|Z}(s_A, s_B, Q, E|Z) = \frac{1}{|S_A|} \delta_{s_A, s_B} \otimes \sum_{s'_A, s'_B} \Lambda_N(\mathcal{E}(P^{\otimes N}))(s'_A, s'_B, Q, E|Z)$, is an *ideal* cc-d distribution which contains uniform and perfectly correlated outcomes shared between the honest parties. Eve is completely uncorrelated in case of this distribution, and it is assumed that Eve's system is the same as she possesses at the end of the real protocol Λ_N .

For the cc-d distribution shared at the end of the MDLOPC protocol, the NS norm given in Eq. (14) takes a more simplified form, stated in the following proposition.

Proposition 2. For the cc-d states P and R shared at the end of the MDLOPC protocol Λ_N , the NS norm can be rephrased with a simplified expression:

$$\begin{aligned} & \|P_{S_A, S_B, Q, E|Z} - R_{S_A, S_B, Q, E|Z}\|_{\text{NS}} \\ &= \frac{1}{2} \sum_{s_A, s_B, q} \max_z \sum_e |P_{S_A, S_B, Q, E|Z}(s_A, s_B, q, e|z) \\ & \quad - R_{S_A, S_B, Q, E|Z}(s_A, s_B, q, e|z)|, \end{aligned} \quad (15)$$

where \max_z stands for the maximization over all possible direct measurements performed by the eavesdropper.

Proof. For the proof, see Sec. C of Appendix.

In the above equality, one can see that the adopted definition of security is equivalent to the one used in Refs. [11,15,23] in the case of the NSDI scenario (the latter is defined as in r.h.s. of the (15)). This justifies our security definition given in Eq. (4), in particular, the choice of the set of operations \mathcal{G} , that define the NS norm $\|\cdot\|_{\text{NS}}$. However, in literature, another definition of security is adopted, given in Refs. [17,24,63,64]. This one is based on assuring high correlations between the parties and low correlations with the eavesdropper. In this approach, Eve can generate *ensembles* of the device of the honest parties i.e., representation of a device as probabilistic mixtures of devices. In later part of this manuscript, we show that the latter definition is also equivalent to the newly proposed one based on the NS norm. By doing so, as a byproduct, we have also proven that our, and the two definitions given in Refs. [11,15,23] and [17,24,63,64] respectively, are equivalent.

B. Equivalence of security criteria

We show that in the NSDI scenario, in analogy to quantum cryptography [77,78], there exist two different, however equivalent definitions of security. One connected to the notion of the so-called distinguisher and the other one based on the proximity in norm [79,80]. In the case of NSDI, Renner, Hänggi, and Wolf [17] present the approach via the notion of distinguisher. Recall here, that to develop the latter approach, we consider the nonsignaling norm, which is a total variational distance for two devices mapped into probability distribution with the so-called *nonsignaling operations*, over which we take a supremum (see Refs. [17,67] in this context). We then focus on tripartite cc-d distributions (classical distribution is isomorphic to a device with unary input) as these are encountered at the end of an NSDI cryptographic protocol. The two classical parts are in the hands of the honest parties, while eavesdropper holds some device. We then show that the NS norm takes for such cc-d distribution a closed-form expression. In particular, we prove that the supremum over

¹²Name for these operations stems from the fact that they are proven in Ref. [56] to generate from the complete extension any possible other nonsignaling extension of a conditional probability distribution.

Eve's operations reduces to a maximization over direct measurements (for the proof, see Sec. C of Appendix).

We present below the theorem, which states that our definition of NS norm security criterion is equivalent to the criteria used by Renner, Hänggi, and Wolf [17]. We do it in analogy to the results of Refs. [77,78] related to quantum device-dependent security, but for nonsignaling devices.

Theorem 3 (Equivalence of the NSDI security criteria). For an MDLOPC protocol Λ , the proximity in the NS norm security criterion is equivalent to the criterion based on secrecy and correctness of the protocol. That is for any $\varepsilon_{\text{sec}} + \varepsilon_{\text{cor}} \equiv \varepsilon \geq \varepsilon_{\text{sec}}, \varepsilon_{\text{cor}} \geq 0$ the following relation holds

$$\begin{aligned} (1 - p_{\text{abort}}) \| P_{S_A, S_B, Q, E|Z}^{\text{real|pass}} - P_{S_A, S_B, Q, E|Z}^{\text{ideal|pass}} \|_{\text{NS}} \\ \leq O(\varepsilon) \iff \{(1 - p_{\text{abort}}) P[S_A \neq S_B | \text{pass}] \leq O(\varepsilon_{\text{cor}}) \\ \wedge (1 - p_{\text{abort}}) \| P_{S_A, Q, E|Z}^{\text{real|pass}} - P_{S_A, Q, E|Z}^{\text{ideal|pass}} \|_{\text{NS}} \leq O(\varepsilon_{\text{sec}})\}, \quad (16) \end{aligned}$$

where p_{abort} is the probability for the protocol to abort and the constant $O(\varepsilon)$ does not depend on any parameter of the protocol.

Proof. For the proof, see Sec. D of Appendix.

Following arguments in Ref. [77], as a consequence of the above Theorem, we can claim that our definition of security is restricted composable [79–81] provided the device is not reused. In that sense, our definition diverges from that of [17] formally in two ways. First, we use the notion of the complete extension. This encapsulates the access of the eavesdropper to all ensembles of the device shared by the honest parties—the fact used in Ref. [17]. Furthermore, in our approach, the memory of Eve is finite and minimal without compromising her eavesdropping power. Second, as we have mentioned, we modify the security criterion, without losing the effect of composability. We use the proximity in NS norm to the ideal classical-classical-device distribution. We show that it is equivalent to the statement that (as it was used in Ref. [17]) the distinguisher can not tell apart the real cc-d distribution from the ideal one.

IX. DISCUSSION AND OPEN PROBLEMS

In this manuscript, we have contributed in three ways to the topics of cryptographic security and Bell nonlocality. We describe them below along with possible directions to follow that naturally appears in consequence.

Firstly, we have initiated a systematic study on the upper bounds on the secret-key rate on the NSDI scenario and defined a computable function, the squashed nonlocality as one of the bounds. We have also demonstrated a direct link between the Secrete Key Agreement scenario and that of NSDI by systematic construction of the bounds in the latter case from the secrecy monotones of the former. Interestingly this method leads among others to a known measure of nonlocality, which is the nonlocality fraction. However, our approach goes much beyond that by offering construction of alternative nonlocality measures, which confirms the generality of our paradigm. Looking for tighter upper bounds stemming from (or even going beyond), the relationship between SKA and NSDI scenarios is a new direction to study.

The numerical estimate of the upper bound suggests that only a limited amount of key can be obtained from quantum

devices with two binary inputs and two binary outputs via direct measurement followed by local operations and public communication. For the family of devices studied here, it is below 40%. Given characterization from Ref. [82] of the boundary of the quantum set, one can find limitations on the key rate obtained via quantum mechanics against a nonsignaling adversary for the set of (2,2,2,2) devices. It appears plausible that employing similar idea to the contextual set of observables may also lead to a novel measure of contextuality which upper bounds their private randomness content [83].

One of the most important problems which arise here is a dual one—whether the isotropic devices in (2,2,2,2) scenario with less than 80% weight of Popescu-Rohrlich box are key undistillable in general. We have shown that one can not distill them by MDLOPC operations, i.e., by direct measurements on device and LOPC operations. However, one might consider that grouping several of such devices together and distilling one of them via the so-called “wirings,” could lead to a positive key if followed by MDLOPC operations. Although one can not exclude this case, it is rather improbable, because an action of wiring, within a group of wired devices, opens a possibility of the forward-signaling attack, as discovered in Ref. [25] and developed in Ref. [26] (the two-way signaling case was excluded already in Ref. [24]). This is the reason why the nonsignaling between individual devices seems necessary precondition of security in NSDI. In any case, extending presented results to a more general class of operations, e.g., including *wirings*, is an important open problem. As a step in this direction, one can consider how the key rate changes if the honest parties have access to randomness private from Eve. Such randomness could be in principle used for performing general measurements. We have also demonstrated applicability of our bound in the (3,2,2,2) scenario, giving a tighter bound to the one provided in Ref. [16]. A more careful study, which takes into account all the extremal points of the nonsignaling polytope in the (3,2,2,2) scenario could be a basis for further tighter bounds.

As the second of the main contributions, we have provided a method of constructing novel measures of nonlocality and proved several important properties for one of them—the squashed nonlocality. Among these properties are the monotonicity, convexity, and additivity. One property which is not considered here, the asymptotic continuity of the squashed nonlocality, will be presented in the forthcoming contribution [84].

Comparing it with the other measure—the relative entropy of nonlocality [55,85,86] may lead to interesting results and possibly the proof that the latter is also an upper bound on the distillable device-independent key. Exploring further the analogy between squashed entanglement and squashed nonlocality may lead to novel analogous results in the realm of quantum devices. We also notice that the squashing procedure can be naturally extended to an arbitrary number of parties. This can be achieved by following Ref. [37], where the multipartite version of the intrinsic information in SKA has been shown to upper bound the conference key in the latter scenario.

As the third contribution, we have realized a idea of incorporating the eavesdropper in the scenario by applying the newly introduced concept of the complete extension [56]. Eve controls the additional interfaces of the extended part. This provides the NSDI protocol a structural definition like the

quantum purification did for QDD and QDI. Although the security condition derived from this approach is equivalent to the former, it shows a direct structural analogy between NSDI and QDD paradigms. In consequence, the complete extension models an adversary with minimal memory required for ultimate eavesdropping power. The amount of memory needed for a given attack in a nonsignaling scenario to best our knowledge has not been studied so far and deserves attention in the future. To formalize security, we considered the NS norm analogous to the trace norm in quantum mechanics. We have proven that this approach is equivalent to the two former ones [11,15,17,23,24,63,64]. We obtained that our definition of security is composable secure if the same device is not reused in composing the protocols (restricted composable). The properties of this NS norm computed for classical-classical devices may become useful also in the context of generalized probabilistic theory [65,66,87]. In this context, it is an important open problem if the class of operations \mathcal{G} , over which supremum is taken in the definition of the NS device norm, is equal to the set of all linear operations \mathcal{L} considered in Ref. [67]. Finding an answer to this problem may lead to the full operational characterization of the set of maps that can be performed on devices.

ACKNOWLEDGMENTS

M.W. thanks Eneet Kaur and Mark Wilde for the discussion during QIP2019. K.H. acknowledges the Fulbright Programm and Mark Wilde for hospitality during the Fulbright scholarship at the School of Electric and Computer Engineering of the Cornell University. The authors acknowledge Ryszard Paweł Kostecki for useful comments. M.W., T.D., and K.H. acknowledge grant Sonata Bis 5 (Grant No. 2015/18/E/ST2/00327) from the National Science Center. M.W., T.D., and K.H. acknowledge partial support by the Foundation for Polish Science through IRAP project co-financed by EU within Smart Growth Operational Programme (Contract No. 2018/MAB/5). The ‘International Centre for Theory of Quantum Technologies’ project (Contract No. MAB/2018/5) is carried out within the International Research Agendas Programme of the Foundation for Polish Science co-financed by the European Union from the funds of the Smart Growth Operational Programme, axis IV: Increasing the research potential (Measure 4.3).

APPENDIX A: DEFINITIONS OF ENTROPIC FUNCTIONS

Notation. In Appendix A, we adapt two different notations for conditional probability distributions (devices). We do this to avoid small fonts in multilevel mathematical expressions that appear in forthcoming parts of this work and hence to make them more readable. For the convenience of the reader, we also provide a table of symbols used throughout the paper (Table I).

In this section, we recall definitions of basic quantities associated with random variables. Suppose A , B , and E are discrete random variables, with outcomes $a \in A$, $b \in B$ and $e \in E$. Let $P(ABE)$ be the joint probability distribution of random variables A , B , and E . Similarly, let $P(A = a, B = b, E = e) \equiv p(abe)$ be the probability for obtaining the outcome $A = a$, $B = b$, and $E = e$.

(i) The Shannon entropy of a random variable (variables) is defined as

$$H(A) = - \sum_a p(a) \log_2 p(a), \quad (\text{A1})$$

$$H(AB) = - \sum_{ab} p(ab) \log_2 p(ab), \quad (\text{A2})$$

$$H(ABE) = - \sum_{abe} p(abe) \log_2 p(abe), \quad (\text{A3})$$

where, $p(ab) = \sum_e p(abe)$ and $p(a) = \sum_b p(ab)$ are the marginal probabilities of the joint probability distribution $P(ABE)$.

(ii) The conditional Shannon entropy of any random variable A with respect to the random variable B , quantifying the lack of knowledge about the outcome of A when one already knows the value of B , is given by

$$H(A|B) = \sum_b p(b) H(A|B = b) = H(AB) - H(B). \quad (\text{A4})$$

(iii) The mutual information $I(A : B)$, measuring the correlations between A and B , is defined as

$$I(A : B) = H(A) + H(B) - H(AB). \quad (\text{A5})$$

(iv) The conditional mutual information $I(A : B|E)$, quantifying the correlation remaining between variables A and B conditioned upon the knowledge about value of third variable E , is given by

$$\begin{aligned} I(A : B|E) &= \sum_e I(A : B|E = e) \quad (\text{A6}) \\ &= H(A|E) + H(B|E) - H(AB|E). \quad (\text{A7}) \end{aligned}$$

(v) The intrinsic mutual information [58,59] $I(A : B \downarrow E)$ is

$$I(A : B \downarrow E) = \inf_{\Theta_{E'|E}} I(A : B|E'), \quad (\text{A8})$$

where $I(A : B|E')$ is the conditional mutual information of the probability distribution $P(ABE') = \sum_e \Theta_{E'|E}(E' = e) P(AB, E = e)$, while the infimum is taken over all possible conditional channels $\Theta_{E'|E}$.

(vi) The reduced intrinsic information [70,88] of random variables A , B and E , denoted by $I(A : B \downarrow\downarrow E)$ is defined as

$$I(A : B \downarrow\downarrow E) = \inf_{\Theta_{U|ABE}} (I(A : B \downarrow EU) + H(U)), \quad (\text{A9})$$

where the infimum is taken over all possible conditional channels $\Theta_{U|ABE}$.

APPENDIX B: THE WORLD OF NONSIGNALING DEVICES AND THE NSDI CRYPTOGRAPHIC SCENARIO

In the NSDI cryptographic scenario, we consider that the honest parties, Alice and Bob, share a cryptographic device of unknown internal structure, identified with a nonsignaling conditional probability distribution $P(AB|XY)$ (we use also $P_{AB|XY}$ notation). We refer to $P(AB|XY)$, as to a nonsignaling device throughout our paper. Here A , B , X , and Y are random variables and $a \in A$, $b \in B$, $x \in X$, and $y \in Y$ are respectively their values. The indices x and y are considered to be choices

TABLE I. List of symbols and abbreviations.

Symbol	Meaning	Symbol	Meaning
$\mathbf{P}(\mathbf{AB} \mathbf{XY})$	Bipartite non-signaling device.	$\mathbf{P}(\mathbf{ABE} \mathbf{XYZ})$	Tripartite non-signaling device.
$\mathbf{P}(\mathbf{ABE})$	Tripartite probability distribution.	$ \psi\rangle_{\mathbf{ABE}}$	A pure tripartite quantum state.
$\mathbf{S}(\mathbf{A} : \mathbf{B} \mathbf{E})$	Secure key rate in SKA model.	$\mathbf{I}(\mathbf{A} : \mathbf{B})$	Mutual information.
$\mathbf{I}(\mathbf{A} : \mathbf{B} \mathbf{E})$	Conditional mutual information.	$\mathbf{I}(\mathbf{A} : \mathbf{B} \downarrow \mathbf{E})$	Intrinsic information.
$\mathbf{K}_D(\rho_{\mathbf{AB}})$	Key rate in QDD scenario.	$\mathbf{I}(\mathbf{A} : \mathbf{B} \downarrow \downarrow \mathbf{E})$	Reduced intrinsic information.
$\mathbf{I}_{\text{sq}}(\rho_{\mathbf{AB}})$	Quantum squashed entanglement.	\mathbf{K}_{DI}	Non-signaling Device independent key rate
$\mathcal{N}_{\text{sq}}(\mathbf{P})$	Non-signaling squashed nonlocality.	$\mathcal{E}(\mathbf{P})$	Complete extension of a device P .
\mathcal{M}^{F}	Full direct measurements.	\mathcal{M}^{G}	General measurements.
Λ	The set of all MDLOPC protocol $\{\Lambda_N\}$	$\tilde{\mathcal{E}}(\mathbf{P})$	Overcomplete extension of the device P .
$\ \mathbf{P} - \mathbf{Q}\ _{\text{NS}}^{\text{res}}$	Restricted NS norm of two devices.	\mathcal{N}_C	Nonlocality cost of a non-signaling device.
$\mathbf{H}(\mathbf{S})$	Entropy of the final key S .	ε	Error in the CHSH game.
\mathbf{Q}	Classical communication variable.	\mathbf{PR}	Popescu Rohrlich box
$\{\mathbf{P}_i, \mathbf{P}^i\}$	An ensemble of a device P .	$\{\mathbf{P}_i, \mathbf{P}_E^i\}$	Pure members ensemble of the device P .
$\mathbf{P}(\mathbf{AB} \mathbf{XY})^{\otimes N}$	Tensor product of N iid copies of the device P .	$\ \mathbf{P} - \mathbf{Q}\ _{\text{NS}}$	Non-signaling device norm of two devices P and Q .
Λ_N	MDLOPC protocol of key distribution acting on N iid copies of a device.	\mathcal{M}	Measurements, maps devices to distributions.
$\mathcal{E}(\mathbf{P}^{\otimes N})$	Complete extension of N iid copies of the device P .	$\mathbf{P}_{\text{ideal}}^{d_N}$	Ideal cc-d distribution of dimension d_N .
\mathcal{O}	All possible linear operations which map a device to a distribution.	\mathbf{S}_A	The set of all possible key string in part of Alice after the MDLOPC operation.
$\mathcal{R}(\Lambda _{\mathbf{P}})$	NSDI key rate for a particular MDLOPC protocol.	\mathbf{S}_B	The set of all possible key string in part of Bob after the MDLOPC operation.
$\mathbf{M}(\mathbf{A} : \mathbf{B} \mathbf{E})$	Secrecy quantifiers of probability distribution $P(\mathbf{ABE})$	$\widehat{\mathbf{M}}(\mathbf{A} : \mathbf{B} \mathbf{E})$	Non-signaling squashed secrecy quantifiers of the device P .
$\overline{\mathbf{PR}}$	Complementary box to Popescu Rohrlich box.	\mathbf{P}_{iso}	Device lying on the isotropic line connecting \mathbf{PR} and $\overline{\mathbf{PR}}$ box.
\mathbf{P}_E	Extremal device in the polytope of all non-signaling devices.	\mathbf{D}	A dice, source of additional randomness.
\mathcal{W}	Variable designate wirings between two devices.	\mathcal{P}_N	LOPC operations on N copies of the distribution.
\mathcal{P}	Class of LOPC operations $\{\mathcal{P}_N\}_{N=1}^{\infty}$, also a protocol for SKA model.	$\mathbf{P}_{\mathbf{B}, \mathbf{A}_1 \mathbf{X}_1}$	A classical-device distribution.
$\mathbf{S}_{\mathbf{ABE}}$	Total state of the system after the MDLOPC protocol.	$\mathbf{P}_{\mathbf{S}_A, \mathbf{S}_B, \mathbf{Q}, \mathbf{E} \mathbf{Z}}^{\text{real}}$	Classical-classical-device distribution after the execution of a real protocol.
$\mathbf{p}^{\text{abort}}$	Probability of aborting the protocol.	$\mathbf{P}_{\mathbf{S}_A, \mathbf{S}_B, \mathbf{Q}, \mathbf{E} \mathbf{Z}}^{\text{real abort}}$	Classical-classical-device distribution after the execution of a real protocol conditioning of aborting.
$\mathbf{P}_{\mathbf{S}_A, \mathbf{S}_B, \mathbf{Q}, \mathbf{E} \mathbf{Z}}^{\text{real pass}}$	Classical-classical-device distribution after the execution of a real protocol conditioning of not aborting.	$\mathbf{P}_{\mathbf{S}_A, \mathbf{S}_B, \mathbf{Q}, \mathbf{E} \mathbf{Z}}^{\text{ideal pass}}$	Classical-classical-device distribution after the execution of an ideal protocol conditioning of not aborting.
$\mathbf{S}_{\mathbf{AE}}$	State of the system after the protocol in part of Alice and Eve.	$\mathcal{D}(\mathbf{P}, \mathbf{Q})$	Distance of two devices P and Q .
$\mathbf{P}[\mathbf{S}_A \neq \mathbf{S}_B]$	Probability of not having the same key strings between Alice and Bob.	\mathbf{C}_i	Message sent from Alice to Bob as part of SKA protocol or vice versa.
\mathbf{C}^t	Collection of all messages $C^t = C_1 C_2 \dots C_t$ sent between Alice and Bob in the t th step.	$\mathbf{I}(\mathbf{S} : \mathbf{C}^t \mathbf{E}^N)$	Mutual information between the final key string and Eve's information.
Λ_N^η	η optimal MDLOPC protocol on N iid copies of the device.	\mathcal{P}_N^η	η optimal LOPC protocol on N iid copies of the distribution.
${}^{x,y}\mathcal{P}_N^\eta$	Measurement dependent η optimal LOPC protocol on N iid copies of the distribution.	Ω_{GMDLOPC}	LOPC operations involve general measurements on the devices.
Λ_{MDLOPC}	LOPC operations involve direct measurements on the devices.	$\mathbf{C}(\mathbf{P})$	Nonlocality fraction of a non-signaling device P .
$\dim_A(\mathcal{P}_N((\mathbf{P}(\mathbf{ABE}))^N))$	Dimension of part A after the LOPC operation on the N copies of the probability distribution.		

of inputs of the honest parties, whereas the respective outcomes are denoted by a and b . The nonsignaling condition for $P_{AB|XY}(ab|xy)$, that roughly speaking forbids faster than light communication between the two parties, is defined as

$$\begin{aligned} P_{A|X}(a|x) &= \sum_b P_{AB|XY}(ab|xy) \\ &= \sum_b P_{AB|XY}(ab|xy') \forall a, x, y, y', \end{aligned} \quad (\text{B1})$$

$$\begin{aligned} P_{B|Y}(b|y) &= \sum_a P_{AB|XY}(ab|xy) \\ &= \sum_a P_{AB|XY}(ab|x'y) \forall b, x, x', y. \end{aligned} \quad (\text{B2})$$

We incorporate the no-signaling eavesdropper (Eve) in the system by giving her the access to the additional interfaces of the *complete extension* (CE) [56], of the shared tripartite nonsignaling device (see next Appendix B 1 for reference on CE). We denote the complete extension of a bipartite device $P(AB|XY)$ as $\mathcal{E}(P)(ABE|XYZ)$, where the additional input $z \in Z$ and the corresponding output $e \in E$, are controlled by Eve. Extending a bipartite device with CE ensures that the nonsignaling constraints also hold between Eve and Alice's and Bob's joint subsystem. Additionally, Eve can also apply local randomness in both her input and output to generate general measurements and to post-processing the output, which gives her the ultimate operational eavesdropping power, as then by construction of CE, she can access all possible ensembles of the extended device [56].

1. The notion of the complete extension

For an arbitrary device $P(A|X)$, one can always find its extension $P(AE|XZ)$ in the space of a larger dimension, such that the nonsignaling constraints are satisfied [see Eqs. (B1) and (B2)]. Some extensions of bipartite nonsignaling boxes have been studied in the past [14–17]. The complete extension defined in Ref. [56], is an extension of the lowest possible dimension, that possesses all basic properties of quantum purification except extremality.

Let us consider a polytope (state space) of nonsignaling devices, with a fixed number of parties and fixed cardinalities of inputs and outputs. An arbitrary device P , in that polytope, can always be expanded as a convex combination of the extremal (pure) devices $\{P_E^i\}$, as $P = \sum_i p_i P_E^i$. The ensemble $\{p_i, P_E^i\}$ will be called a pure members ensemble (PME). The decomposition $\{p_i\}$ is not unique in general [56].

Definition 4 (Minimal ensemble). A pure members ensemble, $\{p_i, P_E^i\}_{i \in \mathcal{I}}$ will be called a *minimal ensemble* of P , if all the members are *pure* and if any proper subset of $\{P_E^i\}_{i \in \mathcal{I}}$ for any new choices of the corresponding probabilities $\{p'_i\}_{i \in \mathcal{I}}$ is not an ensemble of the device P .

We can now invoke the definition of a complete extension. Qualitatively, it is such an extension of a device, which enables to produce all minimal ensembles of it, with the choice of input in the extending part resolving which minimal ensemble will be generated. The complete extension is, by its definition a nonsignaling extension, which makes it a perfect tool for the NSDI cryptography (see Ref. [17] in this context).

Definition 5 (Complete extension [56]). Given a device $P_{\mathcal{A}}(A|X)$, we say that a device $\mathcal{E}(P)_{\mathcal{A}\mathcal{X}}(AE|XZ)$ is its

complete extension to system \mathcal{X} if for any $z \in Z$ and $e \in E$ there holds

$$\mathcal{E}(P)_{\mathcal{A}\mathcal{X}}(A, E = e|X, Z = z) = p(e|z) P_{\mathcal{A}}^{e,z}(A|X), \quad (\text{B3})$$

such that the ensemble $\{p(e|z), P_{\mathcal{A}}^{e,z}(A|X)\}$ is a minimal ensemble of the device $P_{\mathcal{A}}(A|X)$, and corresponding to each minimal ensemble of $P_{\mathcal{A}}(A|X)$, there is exactly one $z \in Z$ which generates it.¹³

Here we slightly abuse the notation, so by $P_{\mathcal{A}}(A|X)$, we mean the device $P(A|X)$ with random variables A and X . The subscript \mathcal{A} denotes that the device is in possession of party \mathcal{A} . Similarly, the subscript \mathcal{X} , for the complete extension $\mathcal{E}(P)_{\mathcal{A}\mathcal{X}}(AE|XZ)$, stands for the extending party \mathcal{X} , who controls the additional interfaces Z and E .

The complete extension satisfies the following properties alike the quantum purification, what makes CE its counterpart [56].

(1) ACCESS. A complete extension of a device P , together with access to arbitrary randomness, gives access to any ensemble of a device P .

(2) GENERATION. The complete extension can be transformed to any other extension.

2. Possible eavesdropping actions

In this section, we define the building blocks of the set of allowed operations that the nonsignaling eavesdropper can perform. In every device-independent key distribution protocol, the honest parties hold a device, the internal structure of which is completely unknown to them. Their task is to share at the end of the protocol a cryptographically secure key, which is perfectly correlated between the honest parties and completely secret with respect to the eavesdropper [78], by use of several copies of the device $P(AB|XY)$. As we are interested in finding the upper bound on the key rate, we consider the attacks by the eavesdropper as an independent and identically distributed (iid) attack as a choice of particular eavesdropping strategy. In this attack, the eavesdropper prepares N iid devices $(P(AB|XY))^{\otimes N} \equiv P^{\otimes N}(AB|XY)$ for Alice and Bob and holds the extending part of the CE $\mathcal{E}(P^{\otimes N})(ABE|XYZ)$, where $\mathbf{A} = A_1 A_2 \cdots A_N$, and similarly for \mathbf{B} , \mathbf{X} , and \mathbf{Y} . At this point we are ready to describe the possible actions of Eve on input and output of the extending system.

(1) Full direct measurement, $\{\mathcal{M}_z^F\}$ defined by choice of input $Z = z$. The inputs correspond to the choices of different minimal ensembles. In a cryptographic sense, some inputs are in favour of Eve, and some are not.

(2) General measurement, $\{\mathcal{M}_z^G\}$, defined by a probabilistic choice of direct measurements $\mathcal{M}_z^G = \sum_{z'} p(z'|z) \mathcal{M}_z^F$. Upon each choice of general measurement on the CE of the shared device, Eve can generate any pure members ensemble of the device. Here $\{p(z'|z)\}$ represents the dice, an external randomness.

(3) Classical post-processing channel $\Theta_{E'|E}$ on the output of the extending subsystem that can also be conditioned upon

¹³The calligraphic \mathcal{X} stands here for the extending system and should not be confused with the input of the system \mathcal{A} .

values of inputs and outputs of the dice. These operations when considered together with a general measurement gives access to all ensembles (possibly mixed) of the part of the device shared by the honest parties.

(4) Eve can also monitor the communication, i.e., collect the classical information exchanged between the honest parties.

The most general strategy of the eavesdropper is to utilize both the general measurement and the post-processing channel. Any other strategy is a specific case of the general one described above. For example, the full direct measurement can be considered as a combination of deterministic dice and an identity post-processing channel.

3. Cryptographic protocol

In this section, we describe the building blocks of the set of operations that the honest parties can perform to generate a cryptographically secure key. In the case of nonsignaling device-independent protocol, the honest parties can perform the following operations on their shared devices.

(1) Full direct measurements on the input, i.e., setting certain values x, y of their inputs X, Y , followed by any composition of operations 2 and 3 below.

(2) Classical post-processing of the distribution.

(3) Public communication.

We call this class of operations as *Measurement on devices followed by local operations and public communications* (MDLOPC) [23]. Here we do not allow the honest parties to perform wirings between their subsystems because the forward signaling between the subsystems has been proved to be an insecure procedure for many important examples of post-processing [25,26]. Limitation from a general measurement to a direct one is because, in the former case, Eve does not have access to correlation with the whole system of Alice and Bob.

In our cryptographic protocol, we prove the security when the Eve's attacking strategy is to prepare N iid copies of a nonsignaling device $P(AB|XY)$ and hands them over to the honest parties. Eve controls the CE of the full system, i.e., $P^{\otimes N}(AB|XY)$. It is important to note that CE of a tensor product of devices is not a tensor product of CE's of these devices. This is the most general eavesdropping strategy (in the iid case) since it gives Eve access to all possible statistical ensembles of the shared device. Incorporating CE in this NSDI scenario encompasses a structural way to access to all ensembles of the extended device, which is the key point in all NSDI security protocol [13–17,21–23].

APPENDIX C: PROPERTIES OF THE NS NORM

The NS norm introduced in Eq. (14) that has its main application in Proposition 2 strongly relies on the notion of the so-called *distinguishing system* [17,24,63]. The *distinguishing system*, also dubbed as the distinguisher, is an external black box type device having the same interfaces as the original device (with one extra output) however, its inputs are interchanged into outputs and vice versa. The structure of the distinguishing system allows it then to be connected to the interfaces of the original device. For each pair of systems

to be distinguished, the distinguisher is devised in such a way that it attains maximal guessing advantage to distinguish between two examined devices. The extra output is used to communicate the guess. For a far more detailed description of the distinguishing system, we refer the reader to Ref. [63].

In this section, we show that in the heuristic approach, the NS norm is a maximal guessing advantage for a distinguisher to distinguish between two devices and plays a role of a distance \mathcal{D} between two conditional probability distributions [63,77]. Devices with unary inputs are isomorphic to probability distributions. For them, the NS norm, is by definition, proportional to the total variational distance.

$$\|P - Q\|_{\text{NS}} = \mathcal{D}(P, Q), \quad (\text{C1})$$

For the sake of cohesion we introduce the NS norm formally.

Definition 6 (Of the NS norm). Let P and P' be any two nonsignaling devices. The following distance measure between P and P' is called the NS norm.

$$\|P - P'\|_{\text{NS}} := \sup_{g \in \mathcal{G}} \frac{1}{2} \|g(P) - g(P')\|_1, \quad (\text{C2})$$

where $\|\cdot\|_1$ is a variational distance between two distributions. Furthermore \mathcal{G} , is a set of generating operations that consists of:

(1) adding an auxiliary device that has single input and single output (a dice),

(2) connecting the output of a device/dice to the input of a dice/device respectively, called wirings,

(3) pre-processing the inputs of device(s),

(4) post-processing inputs and outputs of the devices.

The results of this section, although seem to be highly technical, have a direct implication in distinguishability of the states of devices at the end of the protocol. For an initial tripartite device $P(ABE|XYZ)$, when the honest parties finish the MDLOPC protocol on it, i.e., perform measurements in their respective parts and post-process their data by local operations and public communication, the device is transformed into a *classical-classical-device probability distribution* (c-c-d state). In fact, it is enough to consider classical-device states (c-d states) $P_{B,A_1|X_1}$, and the result still holds for any c-d states, i.e., consisting of many classical subsystems (see Fig. 6). This is because one can always claim that classical variable B is the Cartesian product of many classical variables.

We identify the operations $g \in \mathcal{G}$ that the distinguisher can perform to discriminate between the devices. These can always be decomposed into several basic operations belonging to disjoint sub-classes of different operational meaning, i.e., $g = \mathcal{P} \circ \mathcal{M}^G \circ \mathcal{W}$ considered together with external randomness D . This decomposition guarantees adequate causal order of operations.

(i) The distinguisher can make use of *external randomness*, which in general may depend on the output of the classical part of the system B . We incorporate this randomness by combining systems to be distinguished with an external system, $D_{A_2|X_2,B}$ called a dice.

(ii) A composition of *wirings and prior to input classical communication* (WIPCC), we denote this operation with \mathcal{W} .

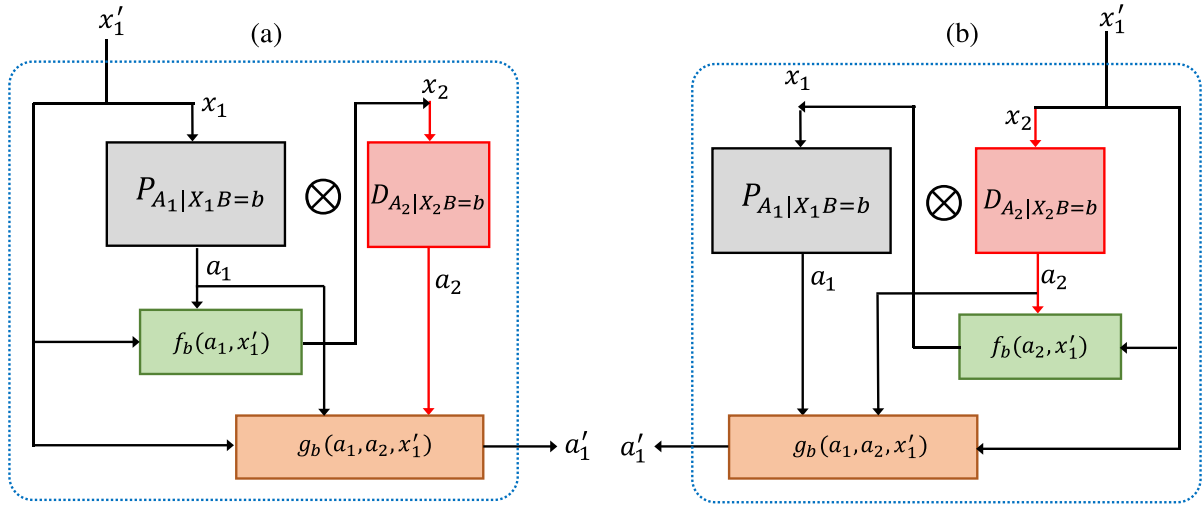


FIG. 6. Schematic diagram of deterministic wiring between the cc-d distribution $P_{A_1|X_1 B}$ and an arbitrary external device (called a dice) $D_{A_2|X_2}$. (a) represents the wiring from the cc-d distribution to the external device, $\mathcal{W}^{\rightarrow}$, and (b) represents the converse one, i.e., wiring from an external device to the cc-d distribution, \mathcal{W}^{\leftarrow} . The diagram is motivated by Ref. [89].

In general, wirings can be adaptive to the outcome of classical variable B , and can be constructed in different manners.

(a) $\mathcal{W}^{\rightarrow}$: deterministic wirings from c-d system to dice.

(b) \mathcal{W}^{\leftarrow} : deterministic wirings from a dice into the input of the c-d system.

(c) A mixture of the above.

(iii) Direct or general measurements

(a) Full direct measurement \mathcal{M}_x^F : A full direct measurement acting on a device $P(A|X) \equiv P_{A|X}$, is equivalent to choosing an input $x \in X$, resulting with a conditional probability distribution,

$$\mathcal{M}_x^F(P(A|X)) = P(A|X = x). \quad (C3)$$

Different x correspond to different measurements (inputs).

(b) General measurement \mathcal{M}_x^G : A general measurement is a mixture of direct measurements, $\mathcal{M}_x^G = \sum_x p(x|x') \mathcal{M}_x^F$, and its action is described as

$$\begin{aligned} \mathcal{M}_x^G(P(A|X)) &= \sum_x p(x|x') \mathcal{M}_x^F(P(A|X)) \\ &= \sum_x p(x|x') P(A|X = x), \end{aligned} \quad (C4)$$

with a conditional probability distribution $p(x|x')$ satisfying $\sum_x p(x|x') = 1 \forall x'$. Here different x' indicate different choices of a general measurement.

(iv) Classical data post-processing we denote with \mathcal{P} .

In the proof of the following proposition, we consider supremum over external systems $D_{A_2|X_2 B}$. Hence without loss of generality, we can consider only wirings employing deterministic functions. The notation for wirings is adapted from Ref. [89], as depicted in Fig. 6 above. The domains and codomains of functions f_b and g_b , which determine wirings, are always adapted to the sizes of inputs and outputs. We consider deterministic wiring, so the sets of $\{f_b\}$ and $\{g_b\}$ are always finite. For the sake of simplicity, in the proof, we omit a unary input in the places where it does not lead to any ambiguity.

Proposition 2. For the c-d states (alike those shared at the end of the MDLOPC-protocol Λ_N), i.e., the many parties nonsignaling device for which only a single party has not unary input, the NS norm takes the form

$$\begin{aligned} & \left\| P_{B,A_1|X_1}^1 - P_{B,A_1|X_1}^2 \right\|_{\text{NS}} \\ &= \frac{1}{2} \sum_b \sup_{\mathcal{M}_{x_1}^F} \sum_a \left| \mathcal{M}_{x_1}^F(P_{B,A_1|X_1}^1)(b, a) \right. \\ & \quad \left. - \mathcal{M}_{x_1}^F(P_{B,A_1|X_1}^2)(b, a) \right|, \end{aligned} \quad (C5)$$

where $b \in B$ is a multivariable corresponding to outputs of c part of the c-d distribution.

From now on, for the sake of the ease of notation we make the following identification: $\mathcal{M}^F \equiv \mathcal{M}_x^F$ and $\mathcal{M}^G \equiv \mathcal{M}_x^G$, where x should be understood from the context. Note that wherever fiducial measurements are considered the sup operator can be used here interchangeably with max operator, as they act in the set with a finite number of elements.

Proof. To attain the supremum over all operations given in Eq. (14), we have to consider all possible actions of the distinguisher. For the proof, it is sufficient to consider the single most general operation instead of a mixture. This is because a norm defined with supremum of some distance is a convex function and attains maximum at the boundaries of the set over which the supremum is evaluated.

$$\begin{aligned} & \sup_{g \in \mathcal{G}} \|g(P) - g(Q)\|_1 \\ &= \sup_{\{\lambda_i\}} \sup_{\{\tilde{g}_i\} \subseteq \mathcal{G}} \left\| \sum_i \lambda_i \tilde{g}_i(P) - \sum_i \lambda_i \tilde{g}_i(Q) \right\|_1 \\ &\leq \sup_{\{\lambda_i\}} \sup_{\{\tilde{g}_i\} \subseteq \mathcal{G}} \sum_i \lambda_i \|\tilde{g}_i(P) - \tilde{g}_i(Q)\|_1 \end{aligned} \quad (C6)$$

$$\leq \sup_{\{\lambda_i\}} \sum_i \lambda_i \sup_{\tilde{g} \in \mathcal{G}} \|\tilde{g}(P) - \tilde{g}(Q)\|_1 = \sup_{\tilde{g} \in \mathcal{G}} \|\tilde{g}(P) - \tilde{g}(Q)\|_1, \quad (C7)$$

where $\tilde{g} \in \tilde{\mathcal{G}}$ are pure operations, i.e., they are not a mixture of others.

Following the arguments of the previous paragraphs the NS norm can be phrased as

$$\begin{aligned} & \left\| P_{B,A_1|X_1}^1 - P_{B,A_1|X_1}^2 \right\|_{\text{NS}} \\ &= \sup_{g \in \mathcal{G}} \frac{1}{2} \left\| g(P_{B,A_1|X_1}^1) - g(P_{B,A_1|X_1}^2) \right\|_1 \end{aligned} \quad (\text{C8})$$

$$\begin{aligned} &= \sup_D \sup_{\mathcal{W}} \sup_{\mathcal{M}^G} \sup_P \frac{1}{2} \left\| (\mathcal{P} \circ \mathcal{M}^G \circ \mathcal{W})(P_{B,A_1|X_1}^1 \otimes D_{A_2|X_2,B}) \right. \\ &\quad \left. - (\mathcal{P} \circ \mathcal{M}^G \circ \mathcal{W})(P_{B,A_1|X_1}^2 \otimes D_{A_2|X_2,B}) \right\|_1, \end{aligned} \quad (\text{C9})$$

where the suprema are taken over operations being adaptive with respect to the output B . When acting on the systems with a fixed value of classical output B , with a little abuse of notation, this can be rephrased using the same symbols for nonadaptive operations.

$$\begin{aligned} & \left\| P_{B,A_1|X_1}^1 - P_{B,A_1|X_1}^2 \right\|_{\text{NS}} \\ &= \frac{1}{2} \sum_b \sup_D \sup_{\mathcal{W}} \sup_{\mathcal{M}^G} \sup_P \left\| (\mathcal{P} \circ \mathcal{M}^G \circ \mathcal{W}) \right. \\ &\quad \times (P_{B=b,A_1|X_1}^1 \otimes D_{A_2|X_2,B=b}) \\ &\quad \left. - (\mathcal{P} \circ \mathcal{M}^G \circ \mathcal{W})(P_{B=b,A_1|X_1}^2 \otimes D_{A_2|X_2,B=b}) \right\|_1. \end{aligned} \quad (\text{C10})$$

The first step to simplify the expression above is to notice that $\|\cdot\|_1$ is contractive under classical post-processing on probability distributions. Since the trivial post-processing is always accessible, we obtain

$$\begin{aligned} & \left\| P_{B,A_1|X_1}^1 - P_{B,A_1|X_1}^2 \right\|_{\text{NS}} \\ &= \frac{1}{2} \sum_b \sup_D \sup_{\mathcal{W}} \sup_{\mathcal{M}^G} \left\| (\mathcal{M}^G \circ \mathcal{W})(P_{B=b,A_1|X_1}^1 \otimes D_{A_2|X_2,B=b}) \right. \\ &\quad \left. - (\mathcal{M}^G \circ \mathcal{W})(P_{B=b,A_1|X_1}^2 \otimes D_{A_2|X_2,B=b}) \right\|_1. \end{aligned} \quad (\text{C11})$$

As it was stated informally above, the general wiring, \mathcal{W} , can be constructed adaptively upon the knowledge of the values of the output B , as a probabilistic combination of two types of wirings $\text{conv}\{\mathcal{W}^\rightarrow, \mathcal{W}^\leftarrow\}$ (see Fig. 6). In the following lines, we show that the strategy of mixing is not optimal. However, in general, the cardinalities of inputs and outputs in different (types) of wiring can be different. In order to overcome this obstacle, we consider a common supremum over a convex set of wirings composed with measurements. From an operational point of view, this procedure means that the knowledge about the preparation was discarded after the optimal measurement for each type of wiring had already been chosen.

$$\begin{aligned} & \left\| P_{B,A_1|X_1}^1 - P_{B,A_1|X_1}^2 \right\|_{\text{NS}} \\ &= \frac{1}{2} \sum_b \sup_D \sup_{\{p_b^\leftarrow, p_b^\rightarrow\}} \sup_{\mathcal{M}^G \circ \mathcal{W}^\leftarrow} \sup_{\mathcal{M}^G \circ \mathcal{W}^\rightarrow} \left\| (p_b^\leftarrow (\mathcal{M}^G \circ \mathcal{W}^\leftarrow)(P_{B=b,A_1|X_1}^1 \otimes D_{A_2|X_2,B=b}) \right. \\ &\quad \left. + p_b^\rightarrow (\mathcal{M}^G \circ \mathcal{W}^\rightarrow)(P_{B=b,A_1|X_1}^1 \otimes D_{A_2|X_2,B=b}) \right. \\ &\quad \left. - (p_b^\leftarrow (\mathcal{M}^G \circ \mathcal{W}^\leftarrow)(P_{B=b,A_1|X_1}^2 \otimes D_{A_2|X_2,B=b}) \right. \\ &\quad \left. + p_b^\rightarrow (\mathcal{M}^G \circ \mathcal{W}^\rightarrow)(P_{B=b,A_1|X_1}^2 \otimes D_{A_2|X_2,B=b}) \right\|_1 \end{aligned} \quad (\text{C12})$$

$$\begin{aligned} & \leq \frac{1}{2} \sum_b \sup_D \sup_{\{p_b^\leftarrow, p_b^\rightarrow\}} \left(\sup_{\mathcal{M}^G \circ \mathcal{W}^\leftarrow} p_b^\leftarrow \left\| (\mathcal{M}^G \circ \mathcal{W}^\leftarrow)(P_{B=b,A_1|X_1}^1 \otimes D_{A_2|X_2,B=b}) \right. \right. \\ &\quad \left. \left. - (\mathcal{M}^G \circ \mathcal{W}^\leftarrow)(P_{B=b,A_1|X_1}^2 \otimes D_{A_2|X_2,B=b}) \right\|_1 + \sup_{\mathcal{M}^G \circ \mathcal{W}^\rightarrow} p_b^\rightarrow \left\| (\mathcal{M}^G \circ \mathcal{W}^\rightarrow)(P_{B=b,A_1|X_1}^1 \otimes D_{A_2|X_2,B=b}) \right. \right. \\ &\quad \left. \left. - (\mathcal{M}^G \circ \mathcal{W}^\rightarrow)(P_{B=b,A_1|X_1}^2 \otimes D_{A_2|X_2,B=b}) \right\|_1 \right) \end{aligned} \quad (\text{C13})$$

$$\begin{aligned} & \leq \frac{1}{2} \sum_b \sup_D \max \left\{ \sup_{\mathcal{M}^G \circ \mathcal{W}^\leftarrow} \left\| (\mathcal{M}^G \circ \mathcal{W}^\leftarrow)(P_{B=b,A_1|X_1}^1 \otimes D_{A_2|X_2,B=b}) \right. \right. \\ &\quad \left. \left. - (\mathcal{M}^G \circ \mathcal{W}^\leftarrow)(P_{B=b,A_1|X_1}^2 \otimes D_{A_2|X_2,B=b}) \right\|_1, \sup_{\mathcal{M}^G \circ \mathcal{W}^\rightarrow} \left\| (\mathcal{M}^G \circ \mathcal{W}^\rightarrow)(P_{B=b,A_1|X_1}^1 \otimes D_{A_2|X_2,B=b}) \right. \right. \\ &\quad \left. \left. - (\mathcal{M}^G \circ \mathcal{W}^\rightarrow)(P_{B=b,A_1|X_1}^2 \otimes D_{A_2|X_2,B=b}) \right\|_1 \right\}. \end{aligned} \quad (\text{C14})$$

In the two following paragraphs, we investigate probability distributions, obtained after the wirings \mathcal{W}^\rightarrow and \mathcal{W}^\leftarrow .

\mathcal{W}^\rightarrow) The first thing to do now is to identify a probability distribution we obtain after wiring. The state of the system after distinguisher obtains a classical output $B = b$, which is prior to input in the considered scenario, is given by $P_{A_1|X_1, B=b} \otimes D_{A_2|X_2, B=b}$, see Fig. 6(a). The distinguisher can apply wirings from P to D , controlled by f_b, g_b , which can depend on outcome b . The probability distribution after the wiring \mathcal{W}^\rightarrow (for a fixed value of outcome B) is given by

$$\mathcal{W}^\rightarrow(P_{A_1|X_1, B} \otimes D_{A_2|X_2, B})_{A_1|X_1, B}(a_1|x_1', b) = \sum_{a_1, a_2: g_b(a_1, a_2, x_1')=a_1'} P_{A_1|X_1, B}(a_1|x_1', b) D_{A_2|X_2, B}(a_2|f_b(a_1, x_1'), b) \quad (\text{C15})$$

Hence the probability distribution for the device after a wiring is given by

$$\overline{P}_{f_b, g_b, B, A_1|X_1'}(b, a_1|x_1') := P_{B|X_1}(b|x_1') \sum_{a_1, a_2: g_b(a_1, a_2, x_1')=a_1'} P_{A_1|X_1, B}(a_1|x_1', b) D_{A_2|X_2, B}(a_2|f_b(a_1, x_1'), b) \quad (\text{C16})$$

$$= \sum_{a_1, a_2: g_b(a_1, a_2, x_1')=a_1'} P_{B, A_1|X_1}(b, a_1|x_1') D_{A_2|X_2, B}(a_2|f_b(a_1, x_1'), b). \quad (\text{C17})$$

\mathcal{W}^{\leftarrow}). The first thing to do is again to identify a probability distribution after wiring. However, we are now in a comfortable situation, as it is enough to interchange inputs of $P_{A_1|X_1,B}$ and $D_{A_2|X_2,B}$ systems, see Fig. 6(b).

$$\overline{\overline{P}}_{f_b, g_b, B, A_1|X_1'}(b, a_1|x_1') := \sum_{a_1, a_2: g_b(a_1, a_2, x_1')=a_1'} P_{B, A_1|X_1}(b, a_1|f_b(a_2, x_1')) D_{A_2|X_2, B}(a_2|x_1', b). \quad (\text{C18})$$

At this point, we are ready to calculate both terms in Eq. (C14) separately.

(a) In the first term $\forall_{b \in B} \forall_{D}$, we have

$$\sup_{\mathcal{M}^G \circ \mathcal{W}^{\rightarrow}} \left\| (\mathcal{M}^G \circ \mathcal{W}^{\rightarrow})(P_{B=b, A_1|X_1}^1 \otimes D_{A_2|X_2, B=b}) - (\mathcal{M}^G \circ \mathcal{W}^{\rightarrow})(P_{B=b, A_1|X_1}^2 \otimes D_{A_2|X_2, B=b}) \right\|_1 \quad (\text{C19})$$

$$= \sup_{f_b, g_b} \sup_{\mathcal{M}^G} \sum_{a_1'} \left| \mathcal{M}^G(\overline{\overline{P}}_{f_b, g_b, B, A_1|X_1'}^1)(b, a_1') - \mathcal{M}^G(\overline{\overline{P}}_{f_b, g_b, B, A_1|X_1'}^2)(b, a_1') \right| \quad (\text{C20})$$

$$= \sup_{f_b, g_b} \sup_{\{\omega_i\}} \sum_{a_1'} \left| \sum_i \omega_i \mathcal{M}_i^F(\overline{\overline{P}}_{f_b, g_b, B, A_1|X_1'}^1)(b, a_1') - \sum_i \omega_i \mathcal{M}_i^F(\overline{\overline{P}}_{f_b, g_b, B, A_1|X_1'}^2)(b, a_1') \right| \quad (\text{C21})$$

$$\leq \sup_{f_b, g_b} \sup_{\{\omega_i\}} \sum_{a_1'} \sum_i \omega_i \left| \mathcal{M}_i^F(\overline{\overline{P}}_{f_b, g_b, B, A_1|X_1'}^1)(b, a_1') - \mathcal{M}_i^F(\overline{\overline{P}}_{f_b, g_b, B, A_1|X_1'}^2)(b, a_1') \right| \quad (\text{C22})$$

$$\leq \sup_{f_b, g_b} \max_{x_1'} \sum_{a_1'} \left| \overline{\overline{P}}_{f_b, g_b, B, A_1|X_1'}^1(b, a_1'|x_1') - \overline{\overline{P}}_{f_b, g_b, B, A_1|X_1'}^2(b, a_1'|x_1') \right| \quad (\text{C23})$$

$$= \sup_{f_b, g_b} \max_{x_1'} \sum_{a_1'} \left| \sum_{a_1, a_2: g_b(a_1, a_2, x_1')=a_1'} P_{B, A_1|X_1}^1(b, a_1|x_1') D_{A_2|X_2, B}(a_2|f_b(a_1, x_1'), b) - \sum_{a_1, a_2: g_b(a_1, a_2, x_1')=a_1'} P_{B, A_1|X_1}^2(b, a_1|x_1') D_{A_2|X_2, B}(a_2|f_b(a_1, x_1'), b) \right| \quad (\text{C24})$$

$$= \sup_{f_b, g_b} \max_{x_1'} \sum_{a_1'} \left| \sum_{a_1, a_2: g_b(a_1, a_2, x_1')=a_1'} D_{A_2|X_2, B}(a_2|f_b(a_1, x_1'), b) (P_{B, A_1|X_1}^1(b, a_1|x_1') - P_{B, A_1|X_1}^2(b, a_1|x_1')) \right| \quad (\text{C25})$$

$$\leq \sup_{f_b, g_b} \max_{x_1'} \sum_{a_1'} \sum_{a_1, a_2: g_b(a_1, a_2, x_1')=a_1'} D_{A_2|X_2, B}(a_2|f_b(a_1, x_1'), b) |P_{B, A_1|X_1}^1(b, a_1|x_1') - P_{B, A_1|X_1}^2(b, a_1|x_1')| \quad (\text{C26})$$

$$= \sup_{f_b, g_b} \max_{x_1'} \sum_{a_1, a_2} D_{A_2|X_2, B}(a_2|f_b(a_1, x_1'), b) |P_{B, A_1|X_1}^1(b, a_1|x_1') - P_{B, A_1|X_1}^2(b, a_1|x_1')| \quad (\text{C27})$$

$$= \max_{x_1'} \sum_{a_1} |P_{B, A_1|X_1}^1(b, a_1|x_1') - P_{B, A_1|X_1}^2(b, a_1|x_1')| \quad (\text{C28})$$

$$= \sup_{\mathcal{M}^F} \sum_{a_1} |\mathcal{M}^F(P_{B, A_1|X_1}^1)(b, a_1) - \mathcal{M}^F(P_{B, A_1|X_1}^2)(b, a_1)|. \quad (\text{C29})$$

The important point is to notice that $\sum_{a_1'} \sum_{a_1, a_2: g_b(a_1, a_2, x_1')=a_1'} h(a_1, a_2) = \sum_{a_1, a_2} h(a_1, a_2)$.

(b) Now in the second term $\forall_{b \in B} \forall_{D}$, we have

$$\sup_{\mathcal{M}^G \circ \mathcal{W}^{\leftarrow}} \left\| (\mathcal{M}^G \circ \mathcal{W}^{\leftarrow})(P_{B=b, A_1|X_1}^1 \otimes D_{A_2|X_2, B=b}) - (\mathcal{M}^G \circ \mathcal{W}^{\leftarrow})(P_{B=b, A_1|X_1}^2 \otimes D_{A_2|X_2, B=b}) \right\|_1 \quad (\text{C30})$$

$$= \sup_{f_b, g_b} \sup_{\mathcal{M}^G} \sum_{a_1'} \left| \mathcal{M}^G(\overline{\overline{P}}_{f_b, g_b, B, A_1|X_1'}^1)(b, a_1') - \mathcal{M}^G(\overline{\overline{P}}_{f_b, g_b, B, A_1|X_1'}^2)(b, a_1') \right| \quad (\text{C31})$$

$$= \sup_{f_b, g_b} \sup_{\{\omega_i\}} \sum_{a_1'} \left| \sum_i \omega_i \mathcal{M}_i^F(\overline{\overline{P}}_{f_b, g_b, B, A_1|X_1'}^1)(b, a_1') - \sum_i \omega_i \mathcal{M}_i^F(\overline{\overline{P}}_{f_b, g_b, B, A_1|X_1'}^2)(b, a_1') \right| \quad (\text{C32})$$

$$\leq \frac{1}{2} \sup_{f_b, g_b} \sup_{\{\omega_i\}} \sum_{a_1'} \sum_i \omega_i \left| \mathcal{M}_i^F(\overline{\overline{P}}_{f_b, g_b, B, A_1|X_1'}^1)(b, a_1') - \mathcal{M}_i^F(\overline{\overline{P}}_{f_b, g_b, B, A_1|X_1'}^2)(b, a_1') \right| \quad (\text{C33})$$

$$\leq \frac{1}{2} \sup_{f_b, g_b} \max_{x'_1} \sum_{a'_1} \left| \overline{P^1_{f_b, g_b, B, A_1 | X'_1}}(b, a'_1 | x'_1) - \overline{P^2_{f_b, g_b, B, A_1 | X'_1}}(b, a'_1 | x'_1) \right| \tag{C34}$$

$$= \sup_{f_b, g_b} \max_{x'_1} \sum_{a'_1} \left| \sum_{a_1, a_2: g_b(a_1, a_2, x'_1)=a'_1} P^1_{B, A_1 | X_1}(b, a_1 | f_b(a_2, x'_1)) D_{A_2 | X_2, B}(a_2 | x'_1, b) \right. \tag{C35}$$

$$\left. - \sum_{a_1, a_2: g_b(a_1, a_2, x'_1)=a'_1} P^2_{B, A_1 | X_1}(b, a_1 | f_b(a_2, x'_1)) D_{A_2 | X_2, B}(a_2 | x'_1, b) \right| \tag{C36}$$

$$= \sup_{f_b, g_b} \max_{x'_1} \sum_{a'_1} \left| \sum_{a_1, a_2: g_b(a_1, a_2, x'_1)=a'_1} D_{A_2 | X_2, B}(a_2 | x'_1, b) (P^1_{B, A_1 | X_1}(b, a_1 | f_b(a_2, x'_1)) \right. \tag{C37}$$

$$\left. - P^2_{B, A_1 | X_1}(b, a_1 | f_b(a_2, x'_1)) \right| \tag{C38}$$

$$\leq \sup_{f_b, g_b} \max_{x'_1} \sum_{a'_1} \sum_{a_1, a_2: g_b(a_1, a_2, x'_1)=a'_1} D_{A_2 | X_2, B}(a_2 | x'_1, b) |P^1_{B, A_1 | X_1}(b, a_1 | f_b(a_2, x'_1)) \tag{C39}$$

$$- P^2_{B, A_1 | X_1}(b, a_1 | f_b(a_2, x'_1))| \tag{C40}$$

$$= \sup_{f_b} \max_{x'_1} \sum_{a_1, a_2} D_{A_2 | X_2, B}(a_2 | x'_1, b) |P^1_{B, A_1 | X_1}(b, a_1 | f_b(a_2, x'_1)) - P^2_{B, A_1 | X_1}(b, a_1 | f_b(a_2, x'_1))| \tag{C41}$$

$$= \sup_{f_b} \max_{x'_1} \sum_{a_2} D_{A_2 | X_2, B}(a_2 | x'_1, b) \sum_{a_1} |P^1_{B, A_1 | X_1}(b, a_1 | f_b(a_2, x'_1)) - P^2_{B, A_1 | X_1}(b, a_1 | f_b(a_2, x'_1))| \tag{C42}$$

$$\leq \sup_{f_b} \max_{x'_1} \sum_{a_2} D_{A_2 | X_2, B}(a_2 | x'_1, b) \max_{a'_2} \sum_{a_1} |P^1_{B, A_1 | X_1}(b, a_1 | f_b(a'_2, x'_1)) - P^2_{B, A_1 | X_1}(b, a_1 | f_b(a'_2, x'_1))| \tag{C43}$$

$$= \sup_{f_b} \max_{x'_1} \max_{a'_2} \sum_{a_1} |P^1_{B, A_1 | X_1}(b, a_1 | f_b(a'_2, x'_1)) - P^2_{B, A_1 | X_1}(b, a_1 | f_b(a'_2, x'_1))| \tag{C44}$$

$$= \max_{x_1} \sum_{a_1} |P^1_{B, A_1 | X_1}(b, a_1 | x_1) - P^2_{B, A_1 | X_1}(b, a_1 | x_1)| \tag{C45}$$

$$= \sup_{\mathcal{M}^F} \sum_{a_1} |\mathcal{M}^F(P^1_{B, A_1 | X_1})(b, a_1) - \mathcal{M}^F(P^2_{B, A_1 | X_1})(b, a_1)|. \tag{C46}$$

From (a), (b), and Eq. (C14), we conclude that

$$\|P^1_{B, A_1 | X_1} - P^2_{B, A_1 | X_1}\|_{NS} \tag{C47}$$

$$\leq \frac{1}{2} \sum_b \sup_D \max \left\{ \sup_{\mathcal{M}^F} \sum_{a_1} |\mathcal{M}^F(P^1_{B, A_1 | X_1})(b, a_1) - \mathcal{M}^F(P^2_{B, A_1 | X_1})(b, a_1)| \right. \tag{C48}$$

$$\left. , \sup_{\mathcal{M}^F} \sum_{a_1} |\mathcal{M}^F(P^1_{B, A_1 | X_1})(b, a_1) - \mathcal{M}^F(P^2_{B, A_1 | X_1})(b, a_1)| \right\} \tag{C49}$$

As the right-hand side (r.h.s.) of the expression above realizes a particular strategy of the distinguisher within considered NS norm, the above inequality can be always saturated, what yields:

$$\|P^1_{B, A_1 | X_1} - P^2_{B, A_1 | X_1}\|_{NS}$$

$$= \frac{1}{2} \sum_b \sup_{\mathcal{M}^F} \sum_{a_1} |\mathcal{M}^F(P^1_{B, A_1 | X_1})(b, a_1) - \mathcal{M}^F(P^2_{B, A_1 | X_1})(b, a_1)|. \tag{C50}$$

■

Corollary 1. For the cc-d states shared at the end of the MDLOPC protocol Λ , the NS norm can be rephrased with a simplified expression:

$$\begin{aligned} & \|P_{S_A, S_B, Q, E|Z} - Q_{S_A, S_B, Q, E|Z}\|_{\text{NS}} \\ &= \frac{1}{2} \sum_{S_A, S_B, Q} \max_z \sum_e |P_{S_A, S_B, Q, E|Z}(S_A, S_B, q, e|z) \\ &\quad - Q_{S_A, S_B, Q, E|Z}(S_A, S_B, q, e|z)|, \end{aligned} \quad (\text{C51})$$

where \max_z stands for the maximization over all possible direct measurements performed by the eavesdropper.

Proof. The proof follows directly from substituting $B \equiv (S_A, S_B, Q)$, $A_1 \equiv E$ and $X_1 \equiv Z$ in the result of Proposition (9). In this way, we obtain cc-d states that are shared at the end of the MDLOPC protocol Λ , and hence we arrive at the claim:

$$\begin{aligned} & \|P_{S_A, S_B, Q, E|Z} - Q_{S_A, S_B, Q, E|Z}\|_{\text{NS}} \\ &= \frac{1}{2} \sum_{S_A, S_B, Q} \max_z \sum_e |P_{S_A, S_B, Q, E|Z}(S_A, S_B, q, e|z) \\ &\quad - Q_{S_A, S_B, Q, E|Z}(S_A, S_B, q, e|z)|, \end{aligned} \quad (\text{C52})$$

where the \max_z is the maximization over direct measurements in the part of Eve. ■

Remark 1. The norm on the space of no-signaling conditional probability distributions based on trace distance introduced by M. Christandl and B. Toner [67] is based on a supremum over all possible linear operations. According to our best knowledge, these operations have not been characterized yet in the literature. In this section, we do not target to describe this class of operations. Instead, via the set \mathcal{G} , we constructed a particular action of the distinguishing system on c-d states, which is sufficient for cryptographic purpose as it yields equivalent security criterion to [23].

APPENDIX D: EQUIVALENCE BETWEEN SECURITY CRITERIA FOR NSDI PROTOCOLS

The iid NSDI key rate in Definition 1 is implicitly dependent on proximity in the NS norm security criterion in Eq. (4). In the quantum case, it was shown that the proximity in the norm (of a state to the ideal one) is equivalent to *the correctness and secrecy* of a protocol [77,78]. These two notions are employed in a protocol independent definition of security [80]. In this section, we show that security criterion based on NS norm is equivalent to the one based on secrecy and correctness of MDLOPC protocol.

In what follows, we employ the notions of *real*, *ideal*, and *intermediate systems*. A real system is a device shared by the parties at the end of a protocol. An ideal device is the one which possesses the same distribution on Eve's side as a real device, however, possesses perfect (uniform) correlations between Alice and Bob, that are completely uncorrelated with Eve. An intermediate device is another kind of device in which Alice and Bob always share fully correlated keys. However, the distribution of the keys is not uniform (Eve's part stays unchanged). The usual part of any protocol employing nonlocal correlations is an acceptance phase in which honest parties

decide (upon some test) whether to abort or to proceed with the protocol.

Composability concept in security is an area of research concerned with composing cryptographic primitives into more complex ones while keeping high security level. In the universal composability approach, a cryptographic primitive is said to be *universally composable* if any functionality using this primitive is as secure as an ideal one [80,81]. The composable security is considered as the strongest notion of security [80,81]. However, in the device independent scenario, so far, it was not rigorously proven that this scheme is ultimately secure. Furthermore, the results of Ref. [68] strongly suggest that it is not the case, so the problem arises when one wants to reuse the device. In particular, if the device used for composition has some memory, then it can leak the key of the previous use. This implies that, in general, the protocol is composable secure as long as the same device is not reused in the protocol. We refer to this notion of security to be restricted composable.

Theorem 3 is essential to compare the secret key of our scenario to these of other cryptographic schemes or even certain protocols, in particular to the results of Hänggi, Renner, and Wolf [17], with the upper bounds that will be presented in this paper. We start with a few definitions.

Definition 7 (State of the device at the end of protocol).

The state of the device after the MDLOPC protocol is a conditional probability distribution (c-d state) denoted by $P_{S_A, S_B, Q, E|Z}^{\text{real}}$:

$$P_{S_A, S_B, Q, E|Z}^{\text{real}} = P_{\text{abort}} P_{S_A, S_B, Q, E|Z}^{\text{real|abort}} + (1 - P_{\text{abort}}) P_{S_A, S_B, Q, E|Z}^{\text{real|pass}}. \quad (\text{D1})$$

The random variables S_A , S_B , and E are respectively outputs of Alice, Bob, and Eve conditioned upon input Z of Eve. S_A and S_B are the key strings hold by Alice and Bob after the protocol, respectively. Q is the random denoting public communication. During the protocol, Q is shared by the three parties, although Alice and Bob use it only to distill the final key and discard it after the protocol is finished. For this reason, we treat Q to be the random variable of Eve that she can use for the choice of her input. Despite the fact that in the notation adopted by as variables of outputs are conditioned upon variables of inputs, Eve's choice of input Z can still depend on the value of Q . The superscripts abort and pass indicate whether protocol passed the acceptance phase.

Definition 8 (Ideal output state). The ideal output state of the device is the one that possesses perfect correlations between honest parties that are completely uncorrelated with the eavesdropper. Local outcomes of the eavesdropper and communication simulate the real system.

$$\begin{aligned} & P_{S_A, S_B, Q, E|Z}^{\text{ideal|pass}}(S_A, S_B, q, e|z) \\ &= \frac{\delta_{S_A, S_B}}{|S_A|} \sum_{S'_A, S'_B} P_{S_A, S_B, Q, E|Z}^{\text{real|pass}}(S'_A, S'_B, q, e|z). \end{aligned} \quad (\text{D2})$$

Since the honest parties are uncorrelated with Eve, the ideal system can be decomposed according to tensor rule formula for independent systems in the following way:

$$P_{S_A, S_B, Q, E|Z}^{\text{ideal|pass}} = P_{S_A, S_B}^{\text{ideal|pass}} \otimes P_{Q, E|Z}^{\text{ideal|pass}}. \quad (\text{D3})$$

Definition 9 (State of the intermediate system). An intermediate system is the one that bears fully correlated key strings between the honest parties, but the distribution they possess is not uniform; hence correlations are not perfect in a cryptographic sense. The eavesdropper is not completely uncorrelated with the honest parties.

$$\begin{aligned} P_{S_A, S_B, Q, E|Z}^{\text{int|pass}}(s_A, s_B, q, e|z) \\ = \delta_{s_A, s_B} \sum_{s_B} P_{S_A, S_B, Q, E|Z}^{\text{real|pass}}(s_A, s_B, q, e|z). \end{aligned} \quad (\text{D4})$$

Since the states of the intermediate and the ideal systems are constructed with respect to the state of the real system, the p_{abort} is the same in all cases (later, we consider the protocol after the acceptance phase, for which $p_{\text{abort}} = 0$). The same is true for all states conditioned on aborting, i.e., they are trivially the same.

For the sake of cohesion, we provide definitions of secrecy, correctness, and security of a cryptographic protocol in case of nonsignaling devices.

Definition 10 (ε -secrecy of a protocol). An MDLOPC key distribution protocol is ε -secret if it outputs a device for which conditional probability distribution shared between Alice (Bob) and Eve at the end of the protocol (and the protocol does not abort) satisfies

$$(1 - p_{\text{abort}}) \|P_{S_A, Q, E|Z}^{\text{real|pass}} - P_{S_A, Q, E|Z}^{\text{ideal|pass}}\|_{\text{NS}} \leq \varepsilon, \quad (\text{D5})$$

Observation 1. The following equality holds.

$$\|P_{S_A, S_B, Q, E|Z}^{\text{real}} - P_{S_A, S_B, Q, E|Z}^{\text{ideal}}\|_{\text{NS}} = (1 - p_{\text{abort}}) \|P_{S_A, S_B, Q, E|Z}^{\text{real|pass}} - P_{S_A, S_B, Q, E|Z}^{\text{ideal|pass}}\|_{\text{NS}}. \quad (\text{D9})$$

Proof.

$$\begin{aligned} & \|P_{S_A, S_B, Q, E|Z}^{\text{real}} - P_{S_A, S_B, Q, E|Z}^{\text{ideal}}\|_{\text{NS}} \\ &= \|p_{\text{abort}} P_{S_A, S_B, Q, E|Z}^{\text{real|abort}} + (1 - p_{\text{abort}}) P_{S_A, S_B, Q, E|Z}^{\text{real|pass}} - p_{\text{abort}} P_{S_A, S_B, Q, E|Z}^{\text{ideal|abort}} - (1 - p_{\text{abort}}) P_{S_A, S_B, Q, E|Z}^{\text{ideal|pass}}\|_{\text{NS}} \end{aligned} \quad (\text{D10})$$

$$= \|p_{\text{abort}} (P_{S_A, S_B, Q, E|Z}^{\text{real|abort}} - P_{S_A, S_B, Q, E|Z}^{\text{ideal|abort}}) + (1 - p_{\text{abort}}) (P_{S_A, S_B, Q, E|Z}^{\text{real|pass}} - P_{S_A, S_B, Q, E|Z}^{\text{ideal|pass}})\|_{\text{NS}} \quad (\text{D11})$$

$$\stackrel{(I)}{=} (1 - p_{\text{abort}}) \|P_{S_A, S_B, Q, E|Z}^{\text{real|pass}} - P_{S_A, S_B, Q, E|Z}^{\text{ideal|pass}}\|_{\text{NS}}, \quad (\text{D12})$$

(I) - we use the fact that $P_{S_A, S_B, Q, E|Z}^{\text{real|abort}}$ and $P_{S_A, S_B, Q, E|Z}^{\text{ideal|abort}}$ are the same when the protocol is aborted [78]. ■

Lemma 1. The NS norm evaluated for real and intermediate states quantifies the probability of Alice and Bob to share different key strings at the end of the protocol.

$$\|P_{S_A, S_B, Q, E|Z}^{\text{real}} - P_{S_A, S_B, Q, E|Z}^{\text{int}}\| = (1 - p_{\text{abort}}) \text{P}[S_A \neq S_B | \text{pass}]. \quad (\text{D13})$$

Proof. From Observation 1, we have

$$\|P_{S_A, S_B, Q, E|Z}^{\text{real}} - P_{S_A, S_B, Q, E|Z}^{\text{int}}\|_{\text{NS}} \stackrel{(I)}{=} (1 - p_{\text{abort}}) \|P_{S_A, S_B, Q, E|Z}^{\text{real|pass}} - P_{S_A, S_B, Q, E|Z}^{\text{int|pass}}\|_{\text{NS}}. \quad (\text{D14})$$

Now, using Proposition 2:

$$\begin{aligned} & \|P_{S_A, S_B, Q, E|Z}^{\text{real|pass}} - P_{S_A, S_B, Q, E|Z}^{\text{int|pass}}\|_{\text{NS}} \\ &= \frac{1}{2} \sum_{s_A, s_B, q} \max_z \sum_e |P_{S_A, S_B, Q, E|Z}^{\text{real|pass}}(s_A, s_B, q, e|z) - P_{S_A, S_B, Q, E|Z}^{\text{int|pass}}(s_A, s_B, q, e|z)| \end{aligned} \quad (\text{D15})$$

$$= \frac{1}{2} \sum_{s_A, s_B, q} \max_z \sum_e \left| P_{S_A, S_B, Q, E|Z}^{\text{real|pass}}(s_A, s_B, q, e|z) - \delta_{s_A, s_B} \sum_{s_B} P_{S_A, S_B, Q, E|Z}^{\text{real|pass}}(s_A, s_B, q, e|z) \right| \quad (\text{D16})$$

$$= \frac{1}{2} \sum_{s_A, s_B, q} \max_z \sum_e \left| P_{S_A, S_B, Q, E|Z}^{\text{real|pass}}(s_A, s_B, q, e|z) - \delta_{s_A, s_B} \sum_{s_B} P_{S_A, S_B, Q, E|Z}^{\text{real|pass}}(s_A, s_B, q, e|z) \right| \delta_{s_A, s_B}$$

where

$$P_{S_A, Q, E|Z}^{\text{real(ideal)|pass}}(s_A, q, e|z) := \sum_{s_B} P_{S_A, S_B, Q, E|Z}^{\text{real(ideal)|pass}}(s_A, s_B, q, e|z). \quad (\text{D6})$$

Definition 11 (ε correctness). An MDLOPC key distribution protocol is ε -correct if the probability (and the protocol does not abort) for Alice and Bob not to share the same output keys satisfies

$$(1 - p_{\text{abort}}) \text{P}[S_A \neq S_B | \text{pass}] \leq \varepsilon. \quad (\text{D7})$$

Definition 12 (ε -security of a protocol). Let $P_{S_A, S_B, Q, E|Z}^{\text{real|pass}}$ be the state of the system shared between Alice, Bob, and Eve after the protocol (and the protocol does not abort). Then the protocol is ε -secure if

$$(1 - p_{\text{abort}}) \|P_{S_A, S_B, Q, E|Z}^{\text{real|pass}} - P_{S_A, S_B, Q, E|Z}^{\text{ideal|pass}}\|_{\text{NS}} \leq \varepsilon, \quad (\text{D8})$$

where p_{abort} is the probability of aborting (which is the same for the real and ideal protocols).

To prove the equivalence between security criterion based on NS norm and the one based on security and correctness, we provide technical Lemmas, showing that proximity in NS norm implies secrecy and correctness, and vice versa.

$$+ \frac{1}{2} \sum_{s_A, q} \sum_{s_B \neq s_A} \max_z \sum_e \left| P_{S_A, S_B, Q, E|Z}^{\text{real|pass}}(s_A, s_B, q, e|z) - \delta_{s_A, s_B} \sum_{s_B} P_{S_A, S_B, Q, E|Z}^{\text{real|pass}}(s_A, s_B, q, e|z) \right| \quad (\text{D17})$$

$$= \frac{1}{2} \sum_{s_A, q} \max_z \sum_e \left| P_{S_A, S_B, Q, E|Z}^{\text{real|pass}}(s_A, s_A, q, e|z) - \sum_{s_B} P_{S_A, S_B, Q, E|Z}^{\text{real|pass}}(s_A, s_B, q, e|z) \right| \\ + \frac{1}{2} \sum_{s_A, q} \sum_{s_B \neq s_A} \max_z \sum_e P_{S_A, S_B, Q, E|Z}^{\text{real|pass}}(s_A, s_B, q, e|z) \quad (\text{D18})$$

$$\stackrel{(I)}{=} \frac{1}{2} \sum_{s_A, q} \max_z \sum_e \left(\sum_{s_B} P_{S_A, S_B, Q, E|Z}^{\text{real|pass}}(s_A, s_B, q, e|z) - P_{S_A, S_B, Q, E|Z}^{\text{real|pass}}(s_A, s_A, q, e|z) \right) \\ + \frac{1}{2} \sum_{s_A, q} \sum_{s_B \neq s_A} \sum_e P_{S_A, S_B, Q, E|Z}^{\text{real|pass}}(s_A, s_B, q, e|z) \quad (\text{D19})$$

$$= \frac{1}{2} \sum_{s_A, q} \max_z \sum_e \sum_{s_B \neq s_A} P_{S_A, S_B, Q, E|Z}^{\text{real|pass}}(s_A, s_B, q, e|z) + \frac{1}{2} \sum_{s_A, q} \sum_{s_B \neq s_A} \sum_e P_{S_A, S_B, Q, E|Z}^{\text{real|pass}}(s_A, s_B, q, e|z) \quad (\text{D20})$$

$$\stackrel{(II)}{=} \sum_{s_A, q} \sum_{s_B \neq s_A} \sum_e P_{S_A, S_B, Q, E|Z}^{\text{real|pass}}(s_A, s_A, q, e|z) \quad (\text{D21})$$

$$= \text{P}[S_A \neq S_B | \text{pass}], \quad (\text{D22})$$

where (I) and (II) are due to nonsignaling condition on Eves's input z . Finally we obtain

$$\| P_{S_A, S_B, Q, E|Z}^{\text{real}} - P_{S_A, S_B, Q, E|Z}^{\text{int}} \|_{\text{NS}} = (1 - \text{p}_{\text{abort}}) \text{P}[S_A \neq S_B | \text{pass}]. \quad (\text{D23})$$

Lemma 2 (Secrecy and correctness imply security). If a protocol is ε_{sec} -secret and ε_{cor} -correct then the protocol is ε -secure, where $\varepsilon = \varepsilon_{\text{sec}} + \varepsilon_{\text{cor}}$.

$$\left\{ (1 - \text{p}_{\text{abort}}) \| P_{S_A, Q, E|Z}^{\text{real|pass}} - P_{S_A, Q, E|Z}^{\text{ideal|pass}} \|_{\text{NS}} \leq \varepsilon_{\text{sec}} \text{ and } (1 - \text{p}_{\text{abort}}) \text{P}[S_A \neq S_B | \text{pass}] \leq \varepsilon_{\text{cor}} \right\} \\ \Rightarrow (1 - \text{p}_{\text{abort}}) \| P_{S_A, S_B, Q, E|Z}^{\text{real|pass}} - P_{S_A, S_B, Q, E|Z}^{\text{ideal|pass}} \|_{\text{NS}} \leq \varepsilon_{\text{sec}} + \varepsilon_{\text{cor}} = \varepsilon. \quad (\text{D24})$$

Proof. To prove the security of the protocol, we can decompose the left-hand side (l.h.s.) of Eq. (D8) in the following way:

$$\| P_{S_A, S_B, Q, E|Z}^{\text{real|pass}} - P_{S_A, S_B, Q, E|Z}^{\text{ideal|pass}} \|_{\text{NS}} \leq \| P_{S_A, S_B, Q, E|Z}^{\text{real|pass}} - P_{S_A, S_B, Q, E|Z}^{\text{int|pass}} \|_{\text{NS}} + \| P_{S_A, S_B, Q, E|Z}^{\text{int|pass}} - P_{S_A, S_B, Q, E|Z}^{\text{ideal|pass}} \|_{\text{NS}}, \quad (\text{D25})$$

where we used the triangle inequality for the NS norm. From proposition 2, we have

$$\| P_{S_A, S_B, Q, E|Z}^{\text{int|pass}} - P_{S_A, S_B, Q, E|Z}^{\text{ideal|pass}} \|_{\text{NS}} \quad (\text{D26})$$

$$= \frac{1}{2} \sum_{s_A, s_B, q} \max_z \sum_e \left| \delta_{s_A, s_B} \sum_{s_B} P_{S_A, S_B, Q, E|Z}^{\text{real|pass}}(s_A, s_B, q, e|z) - \frac{\delta_{s_A, s_B}}{|S_A|} \sum_{s'_A, s'_B} P_{S_A, S_B, Q, E|Z}^{\text{real|pass}}(s'_A, s'_B, q, e|z) \right| \quad (\text{D27})$$

$$= \frac{1}{2} \sum_{s_A, s_B, q} \max_z \sum_e \delta_{s_A, s_B} \left| \sum_{s_B} P_{S_A, S_B, Q, E|Z}^{\text{real|pass}}(s_A, s_B, q, e|z) - \frac{1}{|S_A|} \sum_{s'_A, s'_B} P_{S_A, S_B, Q, E|Z}^{\text{real|pass}}(s'_A, s'_B, q, e|z) \right| \quad (\text{D28})$$

$$= \frac{1}{2} \sum_{s_A} \max_z \sum_e \left| \sum_{s_B} P_{S_A, S_B, Q, E|Z}^{\text{real|pass}}(s_A, s_B, q, e|z) - \sum_{s_B} \frac{\delta_{s_A, s_B}}{|S_A|} \sum_{s'_A, s'_B} P_{S_A, S_B, Q, E|Z}^{\text{real|pass}}(s'_A, s'_B, q, e|z) \right| \quad (\text{D29})$$

$$= \| P_{S_A, Q, E|Z}^{\text{real|pass}} - P_{S_A, Q, E|Z}^{\text{ideal|pass}} \|_{\text{NS}}. \quad (\text{D30})$$

Using now Lemma 1 and Eq. (D25), we have

$$\| P_{S_A, S_B, Q, E|Z}^{\text{real|pass}} - P_{S_A, S_B, Q, E|Z}^{\text{ideal|pass}} \|_{\text{NS}} \leq \text{P}[S_A \neq S_B | \text{pass}] + \| P_{S_A, Q, E|Z}^{\text{real|pass}} - P_{S_A, Q, E|Z}^{\text{ideal|pass}} \|_{\text{NS}}. \quad (\text{D31})$$

Hence,

$$(1 - \text{p}_{\text{abort}}) \| P_{S_A, S_B, Q, E|Z}^{\text{real|pass}} - P_{S_A, S_B, Q, E|Z}^{\text{ideal|pass}} \|_{\text{NS}} \leq (1 - \text{p}_{\text{abort}}) \text{P}[S_A \neq S_B | \text{pass}] + (1 - \text{p}_{\text{abort}}) \| P_{S_A, Q, E|Z}^{\text{real|pass}} - P_{S_A, Q, E|Z}^{\text{ideal|pass}} \|_{\text{NS}}. \quad (\text{D32})$$

Using the above inequality if a protocol is ε_{sec} secret and ε_{cor} correct it is also at least $(\varepsilon_{\text{sec}} + \varepsilon_{\text{cor}})$ secure.

$$\left\{ (1 - \text{p}_{\text{abort}}) \| P_{S_A, Q, E|Z}^{\text{real|pass}} - P_{S_A, Q, E|Z}^{\text{ideal|pass}} \|_{\text{NS}} \leq \varepsilon_{\text{sec}} \text{ and } (1 - \text{p}_{\text{abort}}) \text{P}[S_A \neq S_B | \text{pass}] \leq \varepsilon_{\text{cor}} \right\} \quad (\text{D33})$$

$$\Rightarrow (1 - \text{p}_{\text{abort}}) \| P_{S_A, S_B, Q, E|Z}^{\text{real|pass}} - P_{S_A, S_B, Q, E|Z}^{\text{ideal|pass}} \|_{\text{NS}} \leq \varepsilon_{\text{sec}} + \varepsilon_{\text{cor}} = \varepsilon. \quad (\text{D34})$$

We proved that if the protocol is ε_{sec} secret and ε_{cor} correct then its output is $\varepsilon_{\text{sec}} + \varepsilon_{\text{cor}}$ close to ideal device in NS norm, and by definition is $\varepsilon_{\text{sec}} + \varepsilon_{\text{cor}}$ secure. To prove equivalence of security criteria, we now show the proof in the opposite direction, i.e., we show that if an output device of the protocol is ε close in NS norm to the ideal one, then the protocol is at least ε secret and ε correct.

Lemma 3 (Security implies secrecy and correctness). If a protocol is ε -secure, then it is at least ε -secret and ε -correct.

$$(1 - p_{\text{abort}}) \|P_{S_A, S_B, Q, E|Z}^{\text{real|pass}} - P_{S_A, S_B, Q, E|Z}^{\text{ideal|pass}}\|_{\text{NS}} \leq \varepsilon$$

$$\Rightarrow \{(1 - p_{\text{abort}})P[S_A \neq S_B|\text{pass}] \leq \varepsilon \text{ and } (1 - p_{\text{abort}}) \|P_{S_A, Q, E|Z}^{\text{real|pass}} - P_{S_A, Q, E|Z}^{\text{ideal|pass}}\|_{\text{NS}} \leq \varepsilon\} \quad (\text{D35})$$

Proof of Lemma 3. Let us prove the following first.

$$\|P_{S_A, S_B, Q, E|Z}^{\text{real|pass}} - P_{S_A, S_B, Q, E|Z}^{\text{ideal|pass}}\|_{\text{NS}} \geq P[S_A \neq S_B|\text{pass}]. \quad (\text{D36})$$

To proceed with this task we employ Definition 8 of the ideal system and Proposition 2.

$$\|P_{S_A, S_B, Q, E|Z}^{\text{real|pass}} - P_{S_A, S_B, Q, E|Z}^{\text{ideal|pass}}\|_{\text{NS}} \quad (\text{D37})$$

$$= \frac{1}{2} \sum_{s_A, s_B, q} \max_z \sum_e \left| P_{S_A, S_B, Q, E|Z}^{\text{real|pass}}(s_A, s_B, q, e|z) - \frac{\delta_{s_A, s_B}}{|S_A|} \sum_{s'_A, s'_B} P_{S_A, S_B, Q, E|Z}^{\text{real|pass}}(s'_A, s'_B, q, e|z) \right| \quad (\text{D38})$$

$$= \frac{1}{2} \sum_{s_A, q} \max_z \sum_e \left| P_{S_A, S_B, Q, E|Z}^{\text{real|pass}}(s_A, s_A, q, e|z) - \frac{1}{|S_A|} \sum_{s'_A, s'_B} P_{S_A, S_B, Q, E|Z}^{\text{real|pass}}(s'_A, s'_B, q, e|z) \right|$$

$$+ \frac{1}{2} \sum_{s_A} \sum_{s_B \neq s_A} \max_z \sum_e P_{S_A, S_B, Q, E|Z}^{\text{real|pass}}(s_A, s_B, q, e|z) \quad (\text{D39})$$

$$\stackrel{(I)}{\geq} \frac{1}{2} \sum_{s_A, q} \max_z \left| \sum_e P_{S_A, S_B, Q, E|Z}^{\text{real|pass}}(s_A, s_A, q, e|z) - \frac{1}{|S_A|} \sum_{s'_A, s'_B, e} P_{S_A, S_B, Q, E|Z}^{\text{real|pass}}(s'_A, s'_B, q, e|z) \right|$$

$$+ \frac{1}{2} \sum_{s_A} \sum_{s_B \neq s_A} \max_z \sum_e P_{S_A, S_B, Q, E|Z}^{\text{real|pass}}(s_A, s_B, q, e|z) \quad (\text{D40})$$

$$\stackrel{(II)}{\geq} \frac{1}{2} \sum_{s_A, q} \left| \sum_e P_{S_A, S_B, Q, E|Z}^{\text{real|pass}}(s_A, s_A, q, e|z) - \frac{1}{|S_A|} \sum_{s'_A, s'_B, e} P_{S_A, S_B, Q, E|Z}^{\text{real|pass}}(s'_A, s'_B, q, e|z) \right|$$

$$+ \frac{1}{2} \sum_{s_A} \sum_{s_B \neq s_A} \sum_e P_{S_A, S_B, Q, E|Z}^{\text{real|pass}}(s_A, s_B, q, e|z) \quad (\text{D41})$$

$$\stackrel{(III)}{\geq} \frac{1}{2} \left| \sum_{s_A, q} \left(\sum_e P_{S_A, S_B, Q, E|Z}^{\text{real|pass}}(s_A, s_A, q, e|z) - \frac{1}{|S_A|} \sum_{s'_A, s'_B, e} P_{S_A, S_B, Q, E|Z}^{\text{real|pass}}(s'_A, s'_B, q, e|z) \right) \right| + \frac{1}{2} P[S_A \neq S_B|\text{pass}] \quad (\text{D42})$$

$$= \frac{1}{2} \left| \sum_{s_A, q, e} P_{S_A, S_B, Q, E|Z}^{\text{real|pass}}(s_A, s_A, q, e|z) - \sum_{s_A} \frac{1}{|S_A|} \sum_{s'_A, s'_B, q, e} P_{S_A, S_B, Q, E|Z}^{\text{real|pass}}(s'_A, s'_B, q, e|z) \right| + \frac{1}{2} P[S_A \neq S_B|\text{pass}] \quad (\text{D43})$$

$$= \frac{1}{2} \sum_{s'_A} \sum_{s'_B \neq s'_A} \sum_{e, q} P_{S_A, S_B, Q, E|Z}^{\text{real|pass}}(s'_A, s'_B, q, e|z) + \frac{1}{2} P[S_A \neq S_B|\text{pass}] \quad (\text{D44})$$

$$= P[S_A \neq S_B|\text{pass}], \quad (\text{D45})$$

where we used the triangle inequality used in (I) and (III), and the nonsignaling condition in the Eve's subsystems used in (II). Hence

$$(1 - p_{\text{abort}}) \|P_{S_A, S_B, Q, E|Z}^{\text{real|pass}} - P_{S_A, S_B, Q, E|Z}^{\text{ideal|pass}}\|_{\text{NS}} \geq (1 - p_{\text{abort}}) P[S_A \neq S_B|\text{pass}]. \quad (\text{D46})$$

The above inequality verifies that ε security implies ε correctness.

In the next step, we prove

$$(1 - p_{\text{abort}}) \|P_{S_A, S_B, Q, E|Z}^{\text{real|pass}} - P_{S_A, S_B, Q, E|Z}^{\text{ideal|pass}}\|_{\text{NS}} \geq (1 - p_{\text{abort}}) \|P_{S_A, Q, E|Z}^{\text{real|pass}} - P_{S_A, Q, E|Z}^{\text{ideal|pass}}\|_{\text{NS}}. \quad (\text{D47})$$

Let us use Proposition 2 again.

$$\|P_{S_A, S_B, Q, E|Z}^{\text{real|pass}} - P_{S_A, S_B, Q, E|Z}^{\text{ideal|pass}}\|_{\text{NS}} \quad (\text{D48})$$

$$= \frac{1}{2} \sum_{s_A, s_B, q} \max_z \sum_e \left| P_{S_A, S_B, Q, E|Z}^{\text{real|pass}}(s_A, s_B, q, e|z) - \frac{\delta_{s_A, s_B}}{|S_A|} \sum_{s'_A, s'_B} P_{S_A, S_B, Q, E|Z}^{\text{real|pass}}(s'_A, s'_B, q, e|z) \right| \quad (\text{D49})$$

$$= \frac{1}{2} \sum_{s_A, q} \max_z \sum_e \left| P_{S_A, S_B, Q, E|Z}^{\text{real|pass}}(s_A, s_A, q, e|z) - \frac{1}{|S_A|} \sum_{s'_A, s'_B} P_{S_A, S_B, Q, E|Z}^{\text{real|pass}}(s'_A, s'_B, q, e|z) \right| + \frac{1}{2} \sum_{s_A} \sum_{s_B \neq s_A} \max_z \sum_e P_{S_A, S_B, Q, E|Z}^{\text{real|pass}}(s_A, s_B, q, e|z) \quad (\text{D50})$$

$$\stackrel{(I)}{=} \frac{1}{2} \sum_{s_A, q} \max_z \sum_e \left| P_{S_A, S_B, Q, E|Z}^{\text{real|pass}}(s_A, s_A, q, e|z) - \frac{1}{|S_A|} \sum_{s'_A, s'_B} P_{S_A, S_B, Q, E|Z}^{\text{real|pass}}(s'_A, s'_B, q, e|z) \right| + \frac{1}{2} \text{P}[S_A \neq S_B|\text{pass}] \quad (\text{D51})$$

$$= \frac{1}{2} \sum_{s_A, q} \max_z \sum_e \left| \left(P_{S_A, S_B, Q, E|Z}^{\text{real|pass}}(s_A, s_A, q, e|z) - \sum_{s_B} P_{S_A, S_B, Q, E|Z}^{\text{real|pass}}(s_A, s_B, q, e|z) \right) + \left(\sum_{s_B} P_{S_A, S_B, Q, E|Z}^{\text{real|pass}}(s_A, s_B, q, e|z) - \frac{1}{|S_A|} \sum_{s'_A, s'_B} P_{S_A, S_B, Q, E|Z}^{\text{real|pass}}(s'_A, s'_B, q, e|z) \right) \right| + \frac{1}{2} \text{P}[S_A \neq S_B|\text{pass}] \quad (\text{D52})$$

$$\stackrel{(II)}{\geq} \frac{1}{2} \sum_{s_A, q} \max_z \sum_e \left| \left| P_{S_A, S_B, Q, E|Z}^{\text{real|pass}}(s_A, s_A, q, e|z) - \sum_{s_B} P_{S_A, S_B, Q, E|Z}^{\text{real|pass}}(s_A, s_B, q, e|z) \right| - \left| \sum_{s_B} P_{S_A, S_B, Q, E|Z}^{\text{real|pass}}(s_A, s_B, q, e|z) - \frac{1}{|S_A|} \sum_{s'_A, s'_B} P_{S_A, S_B, Q, E|Z}^{\text{real|pass}}(s'_A, s'_B, q, e|z) \right| \right| + \frac{1}{2} \text{P}[S_A \neq S_B|\text{pass}] \quad (\text{D53})$$

$$\stackrel{(III)}{\geq} \frac{1}{2} \left| \sum_{s_A, q} \max_z \sum_e \left| P_{S_A, S_B, Q, E|Z}^{\text{real|pass}}(s_A, s_A, q, e|z) - \sum_{s_B} P_{S_A, S_B, Q, E|Z}^{\text{real|pass}}(s_A, s_B, q, e|z) \right| - \sum_{s_A, q} \max_z \sum_e \left| \sum_{s_B} P_{S_A, S_B, Q, E|Z}^{\text{real|pass}}(s_A, s_B, q, e|z) - \frac{1}{|S_A|} \sum_{s'_A, s'_B} P_{S_A, S_B, Q, E|Z}^{\text{real|pass}}(s'_A, s'_B, q, e|z) \right| \right| + \frac{1}{2} \text{P}[S_A \neq S_B|\text{pass}] \quad (\text{D54})$$

$$\stackrel{(IV)}{=} \left| \frac{1}{2} \text{P}[S_A \neq S_B|\text{pass}] - \|P_{S_A, Q, E|Z}^{\text{real|pass}} - P_{S_A, Q, E|Z}^{\text{ideal|pass}}\|_{\text{NS}} \right| + \frac{1}{2} \text{P}[S_A \neq S_B|\text{pass}], \quad (\text{D55})$$

where in (I) the second component is treated like in the previous step, reverse triangle inequality has been used in (II), triangle inequality in (III) and in (IV) we use the results given in Eqs. (D22) and (D30). We have

$$(1 - \text{p}_{\text{abort}}) \|P_{S_A, S_B, Q, E|Z}^{\text{real|pass}} - P_{S_A, S_B, Q, E|Z}^{\text{ideal|pass}}\|_{\text{NS}} \quad (\text{D56})$$

$$\geq \left| \frac{1}{2} (1 - \text{p}_{\text{abort}}) \text{P}[S_A \neq S_B|\text{pass}] - (1 - \text{p}_{\text{abort}}) \|P_{S_A, Q, E|Z}^{\text{real|pass}} - P_{S_A, Q, E|Z}^{\text{ideal|pass}}\|_{\text{NS}} \right| + \frac{1}{2} (1 - \text{p}_{\text{abort}}) \text{P}[S_A \neq S_B|\text{pass}]. \quad (\text{D57})$$

One should now go through two separate cases:

Case I. ($\frac{1}{2} \text{P}[S_A \neq S_B|\text{pass}] \geq \|P_{S_A, Q, E|Z}^{\text{real|pass}} - P_{S_A, Q, E|Z}^{\text{ideal|pass}}\|_{\text{NS}}$):

$$(1 - \text{p}_{\text{abort}}) \|P_{S_A, S_B, Q, E|Z}^{\text{real|pass}} - P_{S_A, S_B, Q, E|Z}^{\text{ideal|pass}}\|_{\text{NS}} \geq \frac{1}{2} (1 - \text{p}_{\text{abort}}) \text{P}[S_A \neq S_B|\text{pass}] - (1 - \text{p}_{\text{abort}}) \|P_{S_A, Q, E|Z}^{\text{real|pass}} - P_{S_A, Q, E|Z}^{\text{ideal|pass}}\|_{\text{NS}} + \frac{1}{2} (1 - \text{p}_{\text{abort}}) \text{P}[S_A \neq S_B|\text{pass}] \quad (\text{D58})$$

$$= (1 - \text{p}_{\text{abort}}) \text{P}[S_A \neq S_B|\text{pass}] - (1 - \text{p}_{\text{abort}}) \|P_{S_A, Q, E|Z}^{\text{real|pass}} - P_{S_A, Q, E|Z}^{\text{ideal|pass}}\|_{\text{NS}} \quad (\text{D59})$$

$$\geq 2(1 - \text{p}_{\text{abort}}) \|P_{S_A, Q, E|Z}^{\text{real|pass}} - P_{S_A, Q, E|Z}^{\text{ideal|pass}}\|_{\text{NS}} - (1 - \text{p}_{\text{abort}}) \|P_{S_A, Q, E|Z}^{\text{real|pass}} - P_{S_A, Q, E|Z}^{\text{ideal|pass}}\|_{\text{NS}} \quad (\text{D60})$$

$$\geq (1 - \text{p}_{\text{abort}}) \|P_{S_A, Q, E|Z}^{\text{real|pass}} - P_{S_A, Q, E|Z}^{\text{ideal|pass}}\|_{\text{NS}}. \quad (\text{D61})$$

Case 2. ($\frac{1}{2}\mathbb{P}[S_A \neq S_B|\text{pass}] < \|P_{S_A, Q, E|Z}^{\text{real|pass}} - P_{S_A, Q, E|Z}^{\text{ideal|pass}}\|_{\text{NS}}$):

$$(1 - p_{\text{abort}})\|P_{S_A, S_B, Q, E|Z}^{\text{real|pass}} - P_{S_A, S_B, Q, E|Z}^{\text{ideal|pass}}\|_{\text{NS}} \geq (1 - p_{\text{abort}})\|P_{S_A, Q, E|Z}^{\text{real|pass}} - P_{S_A, Q, E|Z}^{\text{ideal|pass}}\|_{\text{NS}} \quad (\text{D62})$$

$$-\frac{1}{2}(1 - p_{\text{abort}})\mathbb{P}[S_A \neq S_B|\text{pass}] + \frac{1}{2}(1 - p_{\text{abort}})\mathbb{P}[S_A \neq S_B|\text{pass}]$$

$$= (1 - p_{\text{abort}})\|P_{S_A, Q, E|Z}^{\text{real|pass}} - P_{S_A, Q, E|Z}^{\text{ideal|pass}}\|_{\text{NS}}. \quad (\text{D63})$$

Finally,

$$(1 - p_{\text{abort}})\|P_{S_A, S_B, Q, E|Z}^{\text{real|pass}} - P_{S_A, S_B, Q, E|Z}^{\text{ideal|pass}}\|_{\text{NS}} \geq (1 - p_{\text{abort}})\|P_{S_A, Q, E|Z}^{\text{real|pass}} - P_{S_A, Q, E|Z}^{\text{ideal|pass}}\|_{\text{NS}}. \quad (\text{D64})$$

If protocol is ε -secure, we see from (D36) and (D47) that

$$(1 - p_{\text{abort}})\|P_{S_A, S_B, Q, E|Z}^{\text{real|pass}} - P_{S_A, S_B, Q, E|Z}^{\text{ideal|pass}}\|_{\text{NS}} \leq \varepsilon \\ \Rightarrow \{(1 - p_{\text{abort}})\mathbb{P}[S_A \neq S_B|\text{pass}] \leq \varepsilon \text{ and } (1 - p_{\text{abort}})\|P_{S_A, Q, E|Z}^{\text{real|pass}} - P_{S_A, Q, E|Z}^{\text{ideal|pass}}\|_{\text{NS}} \leq \varepsilon\}. \quad (\text{D65})$$

Once we proved the above Lemmas, we can state the Theorem regarding the equivalence between the secrecy and correctness and proximity in NS norm criteria of security for a protocol we have considered.

Theorem 3 (Equivalence of security criteria). For an MDLOPC protocol Λ , the proximity in the NS norm security criterion is equivalent to the criterion based on security and correctness. That is for any $\varepsilon_{\text{sec}} + \varepsilon_{\text{cor}} \equiv \varepsilon \geq \varepsilon_{\text{sec}}, \varepsilon_{\text{cor}} \geq 0$ the following equivalence relation holds:

$$(1 - p_{\text{abort}})\|P_{S_A, S_B, Q, E|Z}^{\text{real|pass}} - P_{S_A, S_B, Q, E|Z}^{\text{ideal|pass}}\|_{\text{NS}} \leq O(\varepsilon) \iff \{(1 - p_{\text{abort}})\mathbb{P}[S_A \neq S_B|\text{pass}] \leq O(\varepsilon_{\text{cor}}) \\ \wedge (1 - p_{\text{abort}})\|P_{S_A, Q, E|Z}^{\text{real|pass}} - P_{S_A, Q, E|Z}^{\text{ideal|pass}}\|_{\text{NS}} \leq O(\varepsilon_{\text{sec}})\}, \quad (\text{D66})$$

where p_{abort} is the probability for the protocol to abort and the constant in $O(\varepsilon)$ does not depend on any parameter of the protocol.

Proof. From Lemma 2, we have

$$\{(1 - p_{\text{abort}})\|P_{S_A, Q, E|Z}^{\text{real|pass}} - P_{S_A, Q, E|Z}^{\text{ideal|pass}}\|_{\text{NS}} \leq \varepsilon_{\text{sec}} \text{ and } (1 - p_{\text{abort}})\mathbb{P}[S_A \neq S_B|\text{pass}] \leq \varepsilon_{\text{cor}}\} \\ \Rightarrow (1 - p_{\text{abort}})\|P_{S_A, S_B, Q, E|Z}^{\text{real|pass}} - P_{S_A, S_B, Q, E|Z}^{\text{ideal|pass}}\|_{\text{NS}} \leq \varepsilon_{\text{sec}} + \varepsilon_{\text{cor}} = \varepsilon, \quad (\text{D67})$$

and from Lemma 3,

$$(1 - p_{\text{abort}})\|P_{S_A, S_B, Q, E|Z}^{\text{real|pass}} - P_{S_A, S_B, Q, E|Z}^{\text{ideal|pass}}\|_{\text{NS}} \leq \varepsilon \\ \Rightarrow \{(1 - p_{\text{abort}})\mathbb{P}[S_A \neq S_B|\text{pass}] \leq \varepsilon \text{ and } (1 - p_{\text{abort}})\|P_{S_A, Q, E|Z}^{\text{real|pass}} - P_{S_A, Q, E|Z}^{\text{ideal|pass}}\|_{\text{NS}} \leq \varepsilon\} \quad (\text{D68})$$

By combining the above implications under $\varepsilon_{\text{sec}} + \varepsilon_{\text{cor}} \equiv \varepsilon \geq \varepsilon_{\text{sec}}, \varepsilon_{\text{cor}} \geq 0$ constraints, we obtain

$$(1 - p_{\text{abort}})\|P_{S_A, S_B, Q, E|Z}^{\text{real|pass}} - P_{S_A, S_B, Q, E|Z}^{\text{ideal|pass}}\|_{\text{NS}} \leq O(\varepsilon) \iff \{(1 - p_{\text{abort}})\mathbb{P}[S_A \neq S_B|\text{pass}] \leq O(\varepsilon_{\text{cor}}) \quad (\text{D69})$$

$$\wedge (1 - p_{\text{abort}})\|P_{S_A, Q, E|Z}^{\text{real|pass}} - P_{S_A, Q, E|Z}^{\text{ideal|pass}}\|_{\text{NS}} \leq O(\varepsilon_{\text{sec}})\}, \quad (\text{D70})$$

hence the corresponding notions are cryptographically equivalent. ■

Remark 2. In the rest of this paper, we assume that the protocol is after the acceptance phase. However, for the full generality in this section, we took a step back and also considered the possibility of aborting. From now, we set the probability of aborting to zero.

APPENDIX E: REPHRASING THE KEY RATE IN THE SECRET-KEY AGREEMENT SCENARIO

The secret-key agreement (SKA) scenario is a cryptographic scheme in which the honest parties and the eavesdropper share many copies of a classical joint probability distribution $P(ABE)$ [1,2]. The honest parties task is to agree on the secret key, by employing local operations and public

communication (LOPC), in such a manner that the eavesdropper's knowledge about the key is negligible. In the following lines, we propose an alternative definition of the secret-key rate $S(A : B|E)$ in the aforementioned scenario and prove that the definition we propose is equivalent to those present in the literature [2,70,75]. This technical result intends to show and utilize a connection between the definition of secret-key rate in SKA and NSDI scenarios, as it was done in the case of quantum cryptography [34].

Before we begin with the proof of Theorem 2, let us recall two definitions of secret-key rate in SKA scenario [2,75].

Definition 13 (The weak secret-key rate [2,75]). The (weak) secret-key rate of A and B with respect to E, denoted $\bar{S}(A : B|Z)$, is the maximal $R \geq 0$ such that for every $\varepsilon > 0$

and for all $N \geq N_0(\epsilon)$ there exists a protocol, using public communication over an insecure but an authenticated channel, such that Alice and Bob, who receive $A^N = [A_1, \dots, A_N]$ and $B^N = [B_1, \dots, B_N]$, can compute keys S_A and S_B , respectively, with the following properties. First, $S_A = S_B$ hold with probability at least $1 - \epsilon$, and second,

$$\frac{1}{N}I(S_A : CE^N) \leq \epsilon \quad \text{and} \quad \frac{1}{N}H(S_A) \geq R - \epsilon \quad (\text{E1})$$

hold. Here, C denotes the collection of messages sent over the insecure channel by Alice and Bob.

Definition 14 (The strong secret-key rate [75]). The strong secret-key rate of A and B with respect to E , denoted by $\bar{S}(A : B|Z)$, is defined in the same way as $\bar{S}(A : B|Z)$ with the modifications that Alice and Bob compute strings S_A and S_B which are with probability at least $1 - \epsilon$ both equal to a string S with the properties

$$I(S : CE^N) \leq \epsilon \quad \text{and} \quad H(S) = \log_2 |S| \geq N \cdot (R - \epsilon). \quad (\text{E2})$$

The above definitions of the secret-key rate were proven to be equivalent [75], i.e., $\bar{S}(A : B|Z) = \bar{S}(A : B|Z)$, for every distribution $P(ABE)$ shared between the parties before the protocol. We propose an alternative definition of the secret-key rate based on proximity in the trace distance (total variational distance).

Definition 15 (The secret-key rate). Let $P(ABE)$ be the joint distribution of three discrete random variables A , B , and E . The secret-key rate $S(A : B|E)$ is given by

$$S(A : B|E)_{P(ABE)} := \sup_{\mathcal{P}} \limsup_{N \rightarrow \infty} \frac{\log_2 \dim_{S_A} [\mathcal{P}_N(P^{\otimes N}(ABE))]}{N}, \quad (\text{E3})$$

where $\mathcal{P} = \cup_{N=1}^{\infty} \{\mathcal{P}_N\}$ is a LOPC protocol that satisfies

$$\|P_N^{\text{real}} - P_N^{\text{ideal}}\|_1 \leq \delta_N \xrightarrow{N \rightarrow \infty} 0, \quad (\text{E4})$$

for

$$P_N^{\text{real}} \equiv P_N^{\text{real}}(S_A S_B C E^N) := \mathcal{P}_N(P^{\otimes N}(ABE)), \quad (\text{E5})$$

$$\begin{aligned} P_N^{\text{ideal}} &\equiv P_N^{\text{ideal}}(S_A S_B C E^N) \\ &:= \left(\frac{\delta_{S_A, S_B}}{|S_A|} \right) \otimes \sum_{S_A, S_B} P_N^{\text{real}}(S_A = s_A, S_B = s_B, C E^N). \end{aligned} \quad (\text{E6})$$

Theorem 2. The secret-key rate $S(A : B|E)$ introduced in Definition 15 is equal to secret-key rates $\bar{S}(A : B|E)$ and $\bar{S}(A : B|Z)$ provided in Definitions 13 and 14, respectively.

Before we show the proof of Theorem 2, we present the basic tools that will be used. For two joint probability distributions $P \equiv P(XY)$ and $Q \equiv Q(XY)$, that are close by according to the trace distance, their Shannon entropies, and the mutual information functions satisfy the asymptotic continuity relations [90,91], which is

$$|H(X)_P - H(X)_Q| \leq \epsilon \log_2 (\dim_X(P) - 1) + h_2(\epsilon), \quad (\text{E7})$$

$$|I(X : Y)_P - I(X : Y)_Q| \leq 2\epsilon \log_2 d + 2g(\epsilon), \quad (\text{E8})$$

where $\epsilon = \frac{1}{2} \|P(XY) - Q(XY)\|_1 \in [0, 1]$, $h_2(\epsilon) := -\epsilon \log_2 \epsilon - (1 - \epsilon) \log_2 (1 - \epsilon)$ is the binary Shannon entropy, $g(\epsilon) := -\epsilon \log_2 \epsilon + (1 + \epsilon) \log_2 (1 + \epsilon)$, and $d = \min\{\dim_X(P), \dim_Y(P)\}$. Functions h_2 and g are equal at $\epsilon = 0$ and for $\epsilon > 0$ $h_2(\epsilon) < g(\epsilon)$. It is also useful to observe that $\|P(X) - Q(X)\|_1 \leq \|P(XY) - Q(XY)\|_1$ for $P(X)$ and $Q(X)$ being marginal probability distributions of $P(XY)$ and $Q(XY)$ respectively.

Another relation that we need is the so-called Pinsker's inequality. It states that if P and Q are two probability distributions, then

$$\frac{1}{2} \|P - Q\|_1 \leq \sqrt{\frac{1}{2} D_{\text{KL}}(P||Q)}, \quad (\text{E9})$$

where $D_{\text{KL}}(P||Q)$ is the Kullback-Leibler divergence. One of the properties of this function is its relation to mutual information, i.e., for a joint probability distribution $P(XY)$ and $P(X)$, $P(Y)$ being its marginal distributions we have: $D_{\text{KL}}(P(XY)||P(X)P(Y)) = I(X : Y)_{P(XY)}$.

The last mathematical property we describe before the proof is the Fano's inequality stating that

$$\begin{aligned} H(X|Y) &\leq h(\epsilon) + P(\epsilon) \log_2 (|X| - 1), \\ P(\epsilon) &= \text{Prob}[X \neq \tilde{X}], \end{aligned} \quad (\text{E10})$$

where $h(x)$ is the binary entropy and $\tilde{X} = f(Y)$ is an approximate version of X .

In the proof, we also use the notions of real and ideal systems. The real system P_N^{real} is a tripartite probability distribution shared by the honest parties after N th round of an LOPC protocol \mathcal{P} . The ideal system P_N^{ideal} is the one in which the honest parties are perfectly correlated (with uniform distribution), and Eve's marginal distribution is the same as for the real system, however completely uncorrelated with the honest parties.

$$P_N^{\text{real}} \equiv P_N^{\text{real}}(S_A S_B C E^N) := \mathcal{P}_N(P^{\otimes N}(ABE)), \quad (\text{E11})$$

$$\begin{aligned} P_N^{\text{ideal}} &\equiv P_N^{\text{ideal}}(S_A S_B C E^N) \\ &:= \left(\frac{\delta_{S_A, S_B}}{|S_A|} \right) \otimes \sum_{S_A, S_B} P_N^{\text{real}}(S_A = s_A, S_B = s_B, C E^N), \end{aligned} \quad (\text{E12})$$

where $P(ABE)$ is tripartite probability distribution shared by all parties at the beginning of SKA protocol, i.e., input state of the protocol, $|S| = \dim_S(P_N^{\text{real}})$ and dimensions of P_N^{real} and P_N^{ideal} are equal. By $\left(\frac{\delta_{S_A, S_B}}{|S_A|} \right)$ we denote a distribution of perfectly and uniformly correlated random variables S_A and S_B .

Proof of Theorem 2. We begin the proof by showing that the weak secret-key rate $\bar{S}(A : B|Z)$, constitutes an upper bound on $S(A : B|Z)$. We do this by showing that every protocol that satisfies the condition in Eq. (E4) also satisfies conditions in Definition 13.

We denote protocol that satisfy security condition in Eq. (E4) with \mathcal{P} . From asymptotic continuity of the mutual information and the fact that $I(S_A : CE^N)_{P_N^{\text{ideal}}} = 0$ by the

construction of P_N^{ideal} , we read

$$\begin{aligned} & \forall_{\mathcal{P}} \forall_N I(S_A : CE^N)_{P_N^{\text{real}}} \\ &= I(S_A : CE^N)_{P_N^{\text{real}}} - I(S_A : CE^N)_{P_N^{\text{ideal}}} \\ &\leq |I(S_A : CE^N)_{P_N^{\text{real}}} - I(S_A : CE^N)_{P_N^{\text{ideal}}}| \\ &\leq 2\delta_N \log_2 d_{S_A} + 2g(\delta_N), \end{aligned} \quad (\text{E13})$$

where $d_{S_A} := \dim_{S_A}(P_N^{\text{real}}) \geq \min\{\dim_{S_A}(P_N^{\text{real}}), \dim_{CE^N}(P_N^{\text{real}})\}$ and $\delta_N \geq \frac{1}{2} \|P_N^{\text{real}} - P_N^{\text{ideal}}\|_1$. Because in any reasonable LOPC protocol dimension of the output is smaller than the dimension of the input, and we observe that¹⁴

$$\begin{aligned} & \forall_{\mathcal{P}} \forall_N d_S \\ &= \dim_{S_A}(P_N^{\text{real}}) = \dim_{S_A}[\mathcal{P}_N((P(ABE))^{\otimes N})] \\ &\leq \dim_A[(P(ABE))^{\otimes N}] \\ &= [\dim_A(P(ABE))]^N. \end{aligned} \quad (\text{E14})$$

Hence,

$$\begin{aligned} & \forall_{\mathcal{P}} \forall_N \frac{1}{N} I(S_A : CE^N)_{P_N^{\text{real}}} \\ &\leq \frac{2\delta_N \log_2 [\dim_A(P(ABE))]^N + 2g(\delta_N)}{N} \\ &= 2\delta_N \log_2 [\dim_A(P(ABE))] + \frac{2g(\delta_N)}{N} \end{aligned} \quad (\text{E15})$$

Hence if a protocol satisfies the trace norm security condition $\|P_N(P^{\otimes N}(ABE)) - P_N^{\text{ideal}}\|_1 \leq \delta_N \xrightarrow{N \rightarrow \infty} 0$ then

$$\forall_{\mathcal{P}} \forall_{\varepsilon > 0} \exists N_1(\varepsilon) \forall_{N \geq N_1(\varepsilon)} \frac{1}{N} I(S_A : CE^N)_{P_N^{\text{real}}} < \varepsilon, \quad (\text{E16})$$

as r.h.s. of Eq. (E15) approaches 0 when N goes to infinity.

Another condition in Definition 13 we call correctness of a protocol, requiring that $S_A = S_B$ with probability at least $1 - \varepsilon$ (equivalently $\text{Prob}[S_A \neq S_B] \leq \varepsilon$) is satisfied¹⁵ by virtue of Theorem 3, with $|Z| = 1$ and $p_{\text{abort}} = 0$.

This is because the NS norm computed for classical probability distributions is equal to the trace distance. Therefore, from the condition in Eq. (E4) and Theorem 3, we have

$$\forall_{\mathcal{P}} \forall_{\varepsilon} \exists N_2(\varepsilon) \forall_{N \geq N_2(\varepsilon)} \text{Prob}[S_A \neq S_B] \leq \delta_N. \quad (\text{E17})$$

Let us show now the upper bound. We first observe that for all protocols the following is true:

$$\begin{aligned} & \forall_{\mathcal{P}} \forall_N H(S_A)_{P_N^{\text{real}}} \leq H(S_A)_{P_N^{\text{ideal}}} \\ &= \log_2 \dim_{S_A}(P_N^{\text{ideal}}) = \log_2 \dim_{S_A}(P_N^{\text{real}}), \end{aligned} \quad (\text{E18})$$

where the inequality is due to the definition of ideal system in which S_A is uniformly distributed and of the same dimension

¹⁴This follows from: $S(A : B|E)_P \leq I(A : B \downarrow E)_P \leq \log_2 \dim_A(P)$.

¹⁵Devices with unary input are isomorphic with unconditional probability distributions.

as in real system. From asymptotic continuity of the Shannon entropy, we have

$$\begin{aligned} & \forall_{\mathcal{P}} \forall_N \frac{1}{N} H(S_A)_{P_N^{\text{real}}} \\ &\geq \frac{1}{N} H(S_A)_{P_N^{\text{ideal}}} - \frac{1}{N} (\delta_N \log_2 (\dim_{S_A}(P_N^{\text{real}}) - 1) + h_2(\delta_N)) \end{aligned} \quad (\text{E19})$$

$$\geq \frac{1}{N} H(S_A)_{P_N^{\text{ideal}}} - \left(2\delta_N \log_2 (\dim_A(P(ABE))) + \frac{2g(\delta_N)}{N} \right), \quad (\text{E20})$$

where the second inequality is a consequence of the similar arguments as in Eq. (E14) and the fact that $\forall_{x>0} h_2(x) < g(x)$.

Let us define $L(N) := \frac{\log_2 \dim_{S_A}(P_N^{\text{real}})}{N}$. In particular there exists $0 < \eta(N) \xrightarrow{N \rightarrow \infty} 0$ such that $L(N) = \limsup_{N \rightarrow \infty} \frac{\log_2 \dim_{S_A}(P_N^{\text{real}})}{N} - \eta(N)$. Hence, we have the following inequality.

$$\begin{aligned} & \forall_{\mathcal{P}} \forall_{\varepsilon > 0} \exists N_3(\varepsilon) \forall_{N > N_3(\varepsilon)} \frac{1}{N} H(S_A)_{P_N^{\text{real}}} \\ &\geq \limsup_{N \rightarrow \infty} \frac{\log_2 \dim_{S_A}(P_N^{\text{real}})}{N} - \varepsilon. \end{aligned} \quad (\text{E21})$$

Let us define now $N_0(\varepsilon) := \max\{N_1(\varepsilon), N_2(\varepsilon), N_3(\varepsilon)\}$. All conditions in Definition 13, are now satisfied as for all $\varepsilon > 0$ and for all $N \geq N_0(\varepsilon)$, with $R = \limsup_{N \rightarrow \infty} \frac{\log_2 \dim_{S_A}(P_N^{\text{real}})}{N}$. The weak secret-key rate is by Definition 13 maximal R , for which second inequality in Eq. (E1) is satisfied, hence to achieve $\bar{S}(A : B|E)$ one has to take a supremum over rates of all protocols.

$$\bar{S}(A : B|E)_{P(ABE)} = \sup_{\bar{\mathcal{P}}} R, \quad (\text{E22})$$

where $\bar{\mathcal{P}}$ are the protocols that satisfy conditions in Definition 13. As we have shown that condition (E4) in Definition 13 implies conditions in Definition 13, it is clear that $\{\mathcal{P}\} \subseteq \{\bar{\mathcal{P}}\}$, and hence:

$$\begin{aligned} & \bar{S}(A : B|E)_{P(ABE)} \\ &= \sup_{\bar{\mathcal{P}}} R \geq \sup_{\mathcal{P}} R \\ &= \sup_{\mathcal{P}} \limsup_{N \rightarrow \infty} \frac{\log_2 \dim_{S_A}(P_N^{\text{real}})}{N}. \end{aligned} \quad (\text{E23})$$

Let us now show that the secret-key rate $S(A : B|E)$ is lower bounded with the strong secret-key rate $\bar{S}(A : B|E)$. In this part, we refer again to results in Appendix D. It is enough to show that conditions in Definition 14 imply secrecy and correctness of a protocol, as by virtue of Theorem 3 and the same arguments regarding the connection between the NS norm and the trace distance, these conditions imply proximity in the trace distance.

We start with the condition of secrecy (see Definition 10). Let $P_N^{\text{real}}(S_A S_B S C E^N)$ be an extension of $P_N^{\text{real}}(S_A S_B C E^N)$,

such it satisfies conditions in Eq. (E2).

$$\begin{aligned}
\forall_{\varepsilon>0} \varepsilon_N^{\text{sec}} &:= \frac{1}{2} \left\| P_N^{\text{real}}(S_A C E^N) - P_N^{\text{ideal}}(S_A C E^N) \right\|_1 = \frac{1}{2} \left\| P_N^{\text{real}}(S_A C E^N) - \left(\frac{1}{|S_A|} \right) \otimes P_N^{\text{real}}(C E^N) \right\|_1 \\
&= \frac{1}{2} \left\| P_N^{\text{real}}(S_A C E^N) - P_N^{\text{real}}(S_A) \otimes P_N^{\text{real}}(C E^N) + P_N^{\text{real}}(S_A) \otimes P_N^{\text{real}}(C E^N) - \left(\frac{1}{|S_A|} \right) \otimes P_N^{\text{real}}(C E^N) \right\|_1 \\
&\leq \frac{1}{2} \left\| P_N^{\text{real}}(S_A C E^N) - P_N^{\text{real}}(S_A) \otimes P_N^{\text{real}}(C E^N) \right\|_1 + \frac{1}{2} \left\| P_N^{\text{real}}(S_A) - \left(\frac{1}{|S_A|} \right) \right\|_1 \\
&\leq \frac{1}{2} \left\| \overline{P}_N^{\text{real}}(S_A S C E^N) - \overline{P}_N^{\text{real}}(S_A S) \otimes P_N^{\text{real}}(C E^N) \right\|_1 + \frac{1}{2} \left\| P_N^{\text{real}}(S_A) - \left(\frac{1}{|S_A|} \right) \right\|_1, \tag{E24}
\end{aligned}$$

where $\left(\frac{1}{|S_A|}\right)$ denotes uniform distribution, and we identify $|S_A|$ with $|S|$. The first term in the equation above can be upper bounded via Pinsker's inequality, and the first inequality in (E2):

$$\begin{aligned}
\forall_{\varepsilon>0} \exists_{N_0(\varepsilon)} \forall_{N>N_0(\varepsilon)} \frac{1}{2} \left\| \overline{P}_N^{\text{real}}(S_A S C E^N) - \overline{P}_N^{\text{real}}(S_A S) \otimes P_N^{\text{real}}(C E^N) \right\|_1 &\leq \sqrt{\frac{1}{2} D_{\text{KL}}(\overline{P}_N^{\text{real}}(S_A S C E^N) \| \overline{P}_N^{\text{real}}(S_A S) \otimes P_N^{\text{real}}(C E^N))} \\
&\stackrel{(I)}{=} \sqrt{\frac{1}{2} I(S_A S : C E^N)_{\overline{P}_N^{\text{real}}}} = \frac{1}{\sqrt{2}} \sqrt{I(S : C E^N)_{\overline{P}_N^{\text{real}}} + I(S_A : C E^N | S)_{\overline{P}_N^{\text{real}}}} \leq \frac{1}{\sqrt{2}} \sqrt{\varepsilon + I(S_A : C E^N | S)_{\overline{P}_N^{\text{real}}}}, \tag{E25}
\end{aligned}$$

where (I) follows from the properties of the Kullback–Leibler divergence. Let us upper bound $I(S_A : C E^N | S)_{\overline{P}_N^{\text{real}}}$ in the next step.

$$\begin{aligned}
I(S_A : C E^N | S)_{\overline{P}_N^{\text{real}}} &= H(S_A | S)_{\overline{P}_N^{\text{real}}} - H(S_A | S, C E^N)_{\overline{P}_N^{\text{real}}} \stackrel{(II)}{\leq} H(S_A | S)_{\overline{P}_N^{\text{real}}} \stackrel{(III)}{\leq} H(e)_{P(e)} + P(e) \log_2(|S| - 1) \stackrel{(IV)}{\leq} h(\varepsilon) \\
&\quad + \varepsilon \log_2(|S| - 1), \tag{E26}
\end{aligned}$$

where $h(x)$ is the binary entropy and in (II) we used non-negativity of the conditional entropy, (III) follows from Fano's inequality for $P(e) = \text{Prob}[S \neq S_A]$, and the last step (IV) is a consequence of $\text{Prob}[S_A = S_B = S] \geq 1 - \varepsilon$ and an assumption that $\varepsilon \leq \frac{1}{2}$. This assumption is well justified in cryptography. From inequalities (E25) and (E26), we have

$$\forall_{\frac{1}{2} \geq \varepsilon > 0} \exists_{N_0(\varepsilon)} \forall_{N>N_0(\varepsilon)} \frac{1}{2} \left\| \overline{P}_N^{\text{real}}(S_A S C E^N) - \overline{P}_N^{\text{real}}(S_A S) \otimes P_N^{\text{real}}(C E^N) \right\|_1 \leq \frac{1}{\sqrt{2}} \sqrt{\varepsilon + h(\varepsilon) + \varepsilon \log_2(|S| - 1)}. \tag{E27}$$

In order to upper bound the second term, we make the following observations:

$$\forall_{\varepsilon>0} \exists_{N_0(\varepsilon)} \forall_{N>N_0(\varepsilon)}$$

$$(a) \text{Prob}[S_A = S_B = S] > 1 - \varepsilon \Rightarrow \text{Prob}[S_A = S] > 1 - \varepsilon \Leftrightarrow \text{Prob}[S_A \neq S] < \varepsilon \Leftrightarrow \sum_{s_A} \sum_{s \neq s_A} P_N^{\text{real}}(s_A s) < \varepsilon, \tag{E28}$$

$$(b) H(S) = \log_2 |S| \Rightarrow P_N^{\text{real}}(s) = \frac{1}{|S|}, \tag{E29}$$

$$(c) \forall_s \frac{1}{|S_A|} = \sum_{s_A} P_N^{\text{real}}(s_A s) = \sum_{s_A \neq s} P_N^{\text{real}}(s_A s) + P_N^{\text{real}}(s s) \geq P_N^{\text{real}}(s s). \tag{E30}$$

Therefore we have

$$\begin{aligned}
\forall_{\varepsilon>0} \exists_{N_0(\varepsilon)} \forall_{N>N_0(\varepsilon)} \frac{1}{2} \left\| P_N^{\text{real}}(S_A) - \left(\frac{1}{|S_A|} \right) \right\|_1 &= \frac{1}{2} \sum_{s_A} \left| P_N^{\text{real}}(s_A) - \frac{1}{|S_A|} \right| = \frac{1}{2} \sum_{s_A} \left| \sum_s P_N^{\text{real}}(s_A s) - \frac{1}{|S_A|} \right| \\
&= \frac{1}{2} \sum_{s_A} \left| \sum_{s \neq s_A} P_N^{\text{real}}(s_A s) + P_N^{\text{real}}(s_A s_A) - \frac{1}{|S_A|} \right| \stackrel{(I)}{\leq} \frac{1}{2} \sum_{s_A} \sum_{s \neq s_A} P_N^{\text{real}}(s_A s) + \frac{1}{2} \sum_{s_A} \left| \frac{1}{|S_A|} - P_N^{\text{real}}(s_A s_A) \right| \\
&= \frac{1}{2} \sum_{s_A} \sum_{s \neq s_A} P_N^{\text{real}}(s_A s) + \frac{1}{2} \sum_s \left| \frac{1}{|S_A|} - P_N^{\text{real}}(s s) \right| \stackrel{(II)}{=} \frac{1}{2} \sum_{s_A} \sum_{s \neq s_A} P_N^{\text{real}}(s_A s) + \frac{1}{2} \sum_{s_A} \left(\frac{1}{|S_A|} - P_N^{\text{real}}(s_A s_A) \right) \\
&= \frac{1}{2} \sum_{s_A} \sum_{s \neq s_A} P_N^{\text{real}}(s_A s) + \frac{1}{2} \left(1 - \sum_{s_A} P_N^{\text{real}}(s_A s_A) \right) = \sum_{s_A} \sum_{s \neq s_A} P_N^{\text{real}}(s_A s) \leq \varepsilon, \tag{E31}
\end{aligned}$$

where (I) follows from triangle inequality, (II) is due to observation (c), and in the last step, we used (a). From eqs. (E24), (E25), and (E31), we conclude that $\varepsilon_N^{\text{sec}} \leq \frac{1}{\sqrt{2}}\sqrt{\varepsilon + h(\varepsilon) + \varepsilon \log_2(|S| - 1)} + \varepsilon$.

The correctness of a protocol is explicitly stated in Definition 14, i.e., $\text{Prob}[S_A = S_B = S] > 1 - \varepsilon$ (see Definition 11 for reference). Hence we have $\varepsilon_N^{\text{cor}} := \varepsilon$. From Theorem 3, we obtain>

$$\begin{aligned} & \forall_{\frac{1}{2} \geq \varepsilon > 0} \exists_{N_0(\varepsilon)} \forall_{N > N_0(\varepsilon)} \\ & \frac{1}{2} \left\| P_N^{\text{real}}(S_A S_B C E^N) - P_N^{\text{ideal}}(S_A S_B C E^N) \right\|_1 \leq \varepsilon_N^{\text{cor}} + \varepsilon_N^{\text{sec}} \\ & \leq \frac{1}{\sqrt{2}} \sqrt{\varepsilon + h(\varepsilon) + \varepsilon \log_2(|S| - 1)} + 2\varepsilon, \end{aligned} \quad (\text{E32})$$

or equivalently

$$\left\| P_N^{\text{real}} - P_N^{\text{ideal}} \right\|_1 \leq \delta_N \xrightarrow{N \rightarrow \infty} 0. \quad (\text{E33})$$

From the second inequality (E2) and Eq. (E18), we have that

$$\forall_{\varepsilon > 0} \exists_{N_0(\varepsilon)} \forall_{N > N_0(\varepsilon)} L(N) \geq R - \varepsilon, \quad (\text{E34})$$

for $L(N) = \frac{\log_2 \dim_{S_A}(P_N^{\text{real}})}{N}$ and hence by performing a limit $N \rightarrow \infty$, and condition of R being maximal number so that the above is satisfied we have $R = \limsup_{N \rightarrow \infty} \frac{\log_2 \dim_{S_A}(P_N^{\text{real}})}{N}$. The strong secret-key rate is defined as

$$\bar{S}(A : B || E)_{P(ABE)} = \sup_{\bar{P}} R, \quad (\text{E35})$$

where \bar{P} are protocols that satisfy conditions in Definition 14. Because conditions in Definition 14 imply Condition (E4), we have $\{\bar{P}\} \subseteq \{P\}$, and therefore,

$$\begin{aligned} & \bar{S}(A : B || E)_{P(ABE)} \\ & = \sup_{\bar{P}} R = \sup_{\bar{P}} \limsup_{N \rightarrow \infty} \frac{\log_2 \dim_{S_A}(P_N^{\text{real}})}{N} \\ & \leq \sup_{P} \limsup_{N \rightarrow \infty} \frac{\log_2 \dim_{S_A}(P_N^{\text{real}})}{N}. \end{aligned} \quad (\text{E36})$$

By combining equations (E23) and (E36), we have

$$\begin{aligned} & \bar{S}(A : B || E)_{P(ABE)} \\ & \leq \sup_{P} \limsup_{N \rightarrow \infty} \frac{\log_2 \dim_{S_A}(P_N^{\text{real}})}{N} \\ & \leq \bar{S}(A : B || E)_{P(ABE)}. \end{aligned} \quad (\text{E37})$$

However, in Ref. [75], it was shown that $\forall_{P(ABE)} \bar{S}(A : B || E)_{P(ABE)} = \bar{S}(A : B || E)_{P(ABE)}$, hence we conclude that

$$\begin{aligned} \bar{S}(A : B || E)_{P(ABE)} & = S(A : B || E)_{P(ABE)} \\ & = \bar{S}(A : B || E)_{P(ABE)}, \end{aligned} \quad (\text{E38})$$

with $S(A : B || E)_{P(ABE)} = \sup_{P} \limsup_{N \rightarrow \infty} \frac{\log_2 \dim_{S_A}[\mathcal{P}_N(P^{\otimes N}(ABE))]}{N}$, and therefore all three definitions are equivalent.

APPENDIX F: UPPER BOUND ON DEVICE INDEPENDENT KEY

In this section, we prove our main result. Namely, we show that the secrecy quantifiers, that provide upper bounds on the key rate in the SKA model [2,75], can serve us to construct upper bounds in device-independent key agreement scenario via operation of squashing. The secret-key agreement scenario (SKA) is a well established area of cryptography, where upper bounds on the key rate are well known and given by entropic functions. The connection between upper bounds in SKA and NSDI cryptographic paradigms that we show in this section may simplify further studies on the latter.

Theorem 1. The secret-key rate, in the nonsignaling device-independent *iid* scenario achieved with MDLOPC operations, $K_{DI}^{(\text{iid})}$, from a device P is upper bounded by any nonsignaling squashed secrecy quantifier evaluated for the complete extension of P :

$$\forall_P \widehat{M}(A : B || E)_{\mathcal{E}(P)} \geq K_{DI}^{(\text{iid})}(P), \quad (\text{F1})$$

where $P \equiv P(AB|XY)$ is a single copy of a bipartite nonsignaling device shared by the honest parties, and $\mathcal{E}(P) \equiv \mathcal{E}(P)(ABE|XYZ)$ is its complete extension to the eavesdropper's system.

Proof of Theorem 1. We start the proof by modifying the equality in Eq. (E3), in Definition 15 in the following way:

$$\max_{x,y} \min_z S(A : B || E)_{(\mathcal{M}_{x,y}^F \otimes \mathcal{M}_z^G) \mathcal{E}(P)(ABE|XYZ)} = \max_{x,y} \min_z \sup_{\mathcal{P}^{x,y,z}} \limsup_{N \rightarrow \infty} \frac{\log_2 \dim_{S_A} [\mathcal{P}_N^{x,y,z} (((\mathcal{M}_{x,y}^F \otimes \mathcal{M}_z^G) \mathcal{E}(P)(ABE|XYZ))^{\otimes N})]}{N}, \quad (\text{F2})$$

where $\mathcal{P}^{x,y,z}$ is a LOPC protocol secure with respect to probability distribution that arises after x, y, z choice of inputs (see Sec. E of Appendix for reference), and $\mathcal{M}_{x,y}^F, \mathcal{M}_z^G$ are fiducial and general measurements of Alice, Bob, and Eve, respectively, described before in Sec. B 2 of Appendix.

Let us notice that for each choice of x and y there exists $z = \bar{z}_{x,y}$ such that

$$\max_{x,y} \sup_{\mathcal{P}^{x,y,\bar{z}_{x,y}}} \limsup_{N \rightarrow \infty} \frac{\log_2 \dim_{S_A} [\mathcal{P}_N^{x,y,\bar{z}_{x,y}} (((\mathcal{M}_{x,y}^F \otimes \mathcal{M}_{\bar{z}_{x,y}}^G) \mathcal{E}(P)(ABE|XYZ))^{\otimes N})]}{N} \quad (\text{F3})$$

$$:= \max_{x,y} \min_z \sup_{\mathcal{P}^{x,y,z}} \limsup_{N \rightarrow \infty} \frac{\log_2 \dim_{S_A} [\mathcal{P}_N^{x,y,z} (((\mathcal{M}_{x,y}^F \otimes \mathcal{M}_z^G) \mathcal{E}(P)(ABE|XYZ))^{\otimes N})]}{N}. \quad (\text{F4})$$

Now, when the optimization domains are explicitly stated, we can make use of max-min inequality to obtain

$$\max_{x,y} \min_z S(A : B||E)_{(\mathcal{M}_{x,y}^F \otimes \mathcal{M}_z^G) \mathcal{E}(P)(ABE|XYZ)} \quad (\text{F5})$$

$$\geq \max_{x,y} \sup_{\mathcal{P}^{x,y,\bar{z}_{x,y}}} \min_z \limsup_{N \rightarrow \infty} \frac{\log_2 \dim_{S_A} [\mathcal{P}_N^{x,y,\bar{z}_{x,y}} (((\mathcal{M}_{x,y}^F \otimes \mathcal{M}_z^G) \mathcal{E}(P)(ABE|XYZ))^{\otimes N})]}{N}. \quad (\text{F6})$$

We notice that the minimization of Eve's choice of input (\min_z) is void in the r.h.s. of the Eq. (F5) above. This is because the value of r.h.s. depends only on the value of $\dim_{S_A}(\cdot)$ that is determined by choice of x, y , and hence by the protocol. Therefore we can write the following sequence of equalities where we swap from classical probability distributions to cc-d states:

$$\forall_{x,y} \forall_{\mathcal{P}^{x,y,\bar{z}_{x,y}}} \min_z \limsup_{N \rightarrow \infty} \frac{\log_2 \dim_{S_A} [\mathcal{P}_N^{x,y,\bar{z}_{x,y}} (((\mathcal{M}_{x,y}^F \otimes \mathcal{M}_z^G) \mathcal{E}(P)(ABE|XYZ))^{\otimes N})]}{N} \quad (\text{F7})$$

$$= \limsup_{N \rightarrow \infty} \frac{\log_2 \dim_{S_A} [\mathcal{P}_N^{x,y,\bar{z}_{x,y}} (((\mathcal{M}_{x,y}^F \otimes \mathbb{1}) \mathcal{E}(P)(ABE|XYZ))^{\otimes N})]}{N} \quad (\text{F8})$$

$$= \limsup_{N \rightarrow \infty} \frac{\log_2 \dim_{S_A} [\mathcal{P}_N^{x,y,\bar{z}_{x,y}} ((\mathcal{M}_{x,y}^F \otimes \mathbb{1})^{\otimes N} \mathcal{E}^{\otimes N}(P)(ABE|XYZ))]}{N} \quad (\text{F9})$$

$$= \limsup_{N \rightarrow \infty} \frac{\log_2 \dim_{S_A} [\mathcal{P}_N^{x,y,\bar{z}_{x,y}} ((\mathcal{M}_{x,y}^F \otimes \mathbb{1})^{\otimes N} \mathcal{E}(P^{\otimes N})(ABE|XYZ))]}{N} \quad (\text{F10})$$

$$= \limsup_{N \rightarrow \infty} \frac{\log_2 \dim_{S_A} [(\mathcal{P}_N^{x,y,\bar{z}_{x,y}} \circ (\mathcal{M}_{x,y}^F)^{\otimes N})(\mathcal{E}(P^{\otimes N})(ABE|XYZ))]}{N}. \quad (\text{F11})$$

In the third equality above, we again used the fact that the dimension of Alice's subsystem (when the protocol is already fixed) is independent of Eve's action and her marginal distribution. This is the reason why we can substitute $\mathcal{E}(P^{\otimes N})$ in the place of $\mathcal{E}^{\otimes N}(P)$. Moreover, in the last equality we use a notation that explicitly shows the composition between a measurement and a LOPC protocol. With a little abuse of notation $\mathbb{1}$ in Eve's part is abandoned.

We notice now that each composition of measurement x, y and protocol $\mathcal{P}^{x,y,\bar{z}_{x,y}}$ is a candidate for MDLOPC protocol $\Lambda := \{\Lambda_N\} = \{\mathcal{P}_N^{x,y,\bar{z}_{x,y}} \circ (\mathcal{M}_{x,y}^F)^{\otimes N}\}$. However we require that

the distribution after the protocol is secure in NS-norm, i.e.,

$$\|\Lambda_N(\mathcal{E}(P^{\otimes N})) - P_{\text{ideal}}^{(d_N)}\|_{\text{NS}} \leq \varepsilon_N \xrightarrow{N \rightarrow \infty} 0, \quad (\text{F12})$$

what implies security not only with respect to Eve choosing $\bar{z}_{x,y}$, but against eavesdropper that has access to all inputs of $\mathcal{E}(P^{\otimes N})$, hence possibly more powerful attacks. This is also a reason why we stay general even if there is any other good choice of $\bar{z}_{x,y}$ in Eq. (F3). Having this in mind, we can write the inequalities below:

$$\max_{x,y} \min_z S(A : B||E)_{(\mathcal{M}_{x,y}^F \otimes \mathcal{M}_z^G) \mathcal{E}(P)(ABE|XYZ)} \quad (\text{F13})$$

$$\geq \max_{x,y} \sup_{\mathcal{P}^{x,y,\bar{z}_{x,y}}} \limsup_{N \rightarrow \infty} \frac{\log_2 \dim_{S_A} [(\mathcal{P}_N^{x,y,\bar{z}_{x,y}} \circ (\mathcal{M}_{x,y}^F)^{\otimes N})(\mathcal{E}(P^{\otimes N})(ABE|XYZ))]}{N} \quad (\text{F14})$$

$$\geq \sup_{\Lambda} \limsup_{N \rightarrow \infty} \frac{\log_2 \dim_{S_A} [\Lambda_N(\mathcal{E}(P^{\otimes N})(ABE|XYZ))]}{N} = K_{DI}^{(\text{iid})}(P), \quad (\text{F15})$$

where the second inequality is due to the fact that now optimization is over a smaller set (not larger), i.e., only these combinations of measurements and LOPC operations that satisfy security condition in Eq. (F12). Moreover, in the equality we identified MDLOPC (iid) secret-key rate from Definition 1.

For the second part of the proof, we need to recall some properties of a family of secrecy quantifiers $\{M(A : B||E)\}$ of SKA model [88]. Each function that upper bounds secret-key rate in the SKA paradigm can be squashed according to the following procedure. For any function among them:

$$\forall_{Q(ABE)} M(A : B||E)_{Q(ABE)} \geq S(A : B||E)_{Q(ABE)}. \quad (\text{F16})$$

By extending the above inequality to any tripartite nonsignaling device $P(ABE|XYZ)$ and general measurement for input Z , one can write

$$\forall_{P(ABE|XYZ)} \forall_{x,y,z} M(A : B||E)_{(\mathcal{M}_{x,y}^F \otimes \mathcal{M}_z^G) P(ABE|XYZ)} \geq S(A : B||E)_{(\mathcal{M}_{x,y}^F \otimes \mathcal{M}_z^G) P(ABE|XYZ)}. \quad (\text{F17})$$

Without loss of generality, we fix the device $P(ABE|XYZ)$ for now. Let us denote $\bar{z}_{x,y}$ as such an adaptive choice of z that

$$\forall_{x,y} M(A : B||E)_{(\mathcal{M}_{x,y}^F \otimes \mathcal{M}_{\bar{z}_{x,y}}^G) P(ABE|XYZ)} := \min_z M(A : B||E)_{(\mathcal{M}_{x,y}^F \otimes \mathcal{M}_z^G) P(ABE|XYZ)}. \quad (\text{F18})$$

The immediate consequence is

$$\begin{aligned} \forall_{x,y} \min_z M(A : B|E)_{(\mathcal{M}_{x,y}^F \otimes \mathcal{M}_z^G)P(ABE|XYZ)} \\ = M(A : B|E)_{(\mathcal{M}_{x,y}^F \otimes \mathcal{M}_{x,y}^G)P(ABE|XYZ)} \end{aligned} \quad (\text{F19})$$

$$\begin{aligned} \geq S(A : B|E)_{(\mathcal{M}_{x,y}^F \otimes \mathcal{M}_{x,y}^G)P(ABE|XYZ)} \\ \geq \min_z S(A : B|E)_{(\mathcal{M}_{x,y}^F \otimes \mathcal{M}_z^G)P(ABE|XYZ)}. \end{aligned} \quad (\text{F20})$$

Employing a similar technique again, let us choose \tilde{x}, \tilde{y} such that

$$\begin{aligned} \min_z S(A : B|E)_{(\mathcal{M}_{\tilde{x},\tilde{y}}^F \otimes \mathcal{M}_z^G)P(ABE|XYZ)} \\ := \max_{x,y} \min_z S(A : B|E)_{(\mathcal{M}_{x,y}^F \otimes \mathcal{M}_z^G)P(ABE|XYZ)}. \end{aligned} \quad (\text{F21})$$

This yields

$$\begin{aligned} \max_{x,y} \min_z S(A : B|E)_{(\mathcal{M}_{x,y}^F \otimes \mathcal{M}_z^G)P(ABE|XYZ)} \\ = \min_z S(A : B|E)_{(\mathcal{M}_{\tilde{x},\tilde{y}}^F \otimes \mathcal{M}_z^G)P(ABE|XYZ)} \end{aligned} \quad (\text{F22})$$

$$\begin{aligned} \leq \min_z M(A : B|E)_{(\mathcal{M}_{\tilde{x},\tilde{y}}^F \otimes \mathcal{M}_z^G)P(ABE|XYZ)} \\ \leq \max_{x,y} \min_z M(A : B|E)_{(\mathcal{M}_{x,y}^F \otimes \mathcal{M}_z^G)P(ABE|XYZ)}. \end{aligned} \quad (\text{F23})$$

On the r.h.s. we recognize $\widehat{M}(A : B|E)_{P(ABE|XYZ)}$ from Definition 2. Using the result in Eqs. (F13)–(F15) from the first part of the proof, and substituting the complete extension of $P(AB|XY)$ as a tripartite device, we obtain

$$\forall_{P(AB|XY)} \widehat{M}(A : B|E)_{\mathcal{E}(P)(ABE|XYZ)} \geq K_{\text{DI}}^{(\text{id})}(P(AB|XY)). \quad (\text{F24})$$

APPENDIX G: PROOF OF THE PROPERTIES OF NONSIGNALING SQUASHED NONLOCALITY

In this section, we give the proofs of the properties of the nonsignaling squashed nonlocality. Before we start with the proof, let us recall the definition of intrinsic information $I(A : B \downarrow E)$, given in Sec. X. We will rewrite the definition in two new ways. One of them is in full analogy to the forms of the squashed entanglement [61,62]. Indeed, one can write the latter measure in terms of the minimization over all possible extensions: $E_{\text{sq}}(\rho_{AB}) := \inf_{\sigma_{ABE} : \text{Tr}_E \sigma_{ABE} = \rho_{AB}} I(A : B|E)_{\sigma_{ABE}}$. The second form of the squashed nonlocality involves ensembles induced by measurements on the extending system and resembles the definition of the so-called *classical squashed entanglement* [62].

The intrinsic information involves an optimization over all possible conditional probability distributions $\Theta_{E|E}$. Moreover, in the squashing procedure, an optimization over the measurements on the CE of a bipartite device $P(AB|XY)$, has been involved. The nonsignaling squashed intrinsic information is

$$\begin{aligned} \widehat{I}(A : B \downarrow E)_{\mathcal{E}(P)(ABE|XYZ)} \\ = \max_{x,y} \min_z I(A : B \downarrow E)_{(\mathcal{M}_{x,y}^F \otimes \mathcal{M}_z^G)\mathcal{E}(P)(ABE|XYZ)} \end{aligned} \quad (\text{G1})$$

$$= \max_{x,y} \min_z \inf_{\Theta_{E|E}^z} I(A : B|E')_{(\Theta_{E|E}^z)(\mathcal{M}_{x,y}^F \otimes \mathcal{M}_z^G)\mathcal{E}(P)(ABE|XYZ)}, \quad (\text{G2})$$

where $\mathcal{M}_{x,y}^F$ is the direct measurement on the inputs X and Y , and \mathcal{M}_z^G is a general measurement on Z . According to Theorem 4 of Ref. [56], $(\Theta_{E|E}^z)(\mathbb{1} \otimes \mathcal{M}_z^G)\mathcal{E}(P)(ABE|XYZ) = \sum_e \Theta_{E|E}^z \sum_z p(z|z') \mathcal{E}(P)(ABE = e|XYZ = z) = \tilde{P}(ABE'|XYZ' = z')$, is an arbitrary ensemble (possibly mixed) of the device $P(AB|XY)$, where $\mathbb{1}$ is the identity operator on the system of the honest parties. Hence, for a fixed input randomizer (dice $p(z|z')$) and a fixed channel, one can generate an arbitrary extension $\tilde{P}(ABE'|XY)$ with unary input. All possible choices of input randomizer and post-processing channel lead to all possible extensions, hence $\min_z \inf_{\Theta_{E|E}^z} = \inf_{\tilde{P}(ABE'|XY)}$. And hence, it follows that Definition 3 of the squashed nonlocality is equivalent to

$$\mathcal{N}_{\text{sq}}(P(AB|XY)) = \max_{x,y} \inf_{\tilde{P}(ABE|XY)} I(A : B|E)_{\mathcal{M}_{x,y}^F \tilde{P}(ABE|XY)}. \quad (\text{G3})$$

This arbitrary extension of a form $\tilde{P}(ABE|XY)$, gives rise to an arbitrary but fixed ensemble of the bipartite device $P(AB|XY) = \sum_e \tilde{P}(ABE = e|XY) = \sum_e p(e)P^e(AB|XY)$, where $P^e(AB|XY)$ is an arbitrary device corresponding to each output $E = e$, and belongs to the same polytope (state space) as $P(AB|XY)$. Moreover, all possible choices of $\tilde{P}(ABE|XY)$ give rise to all possible ensembles of $P(AB|XY)$. The set of all ensembles of a given device $P(AB|XY)$, reads

$$\mathcal{S}^{\text{all}} := \left\{ \{p_i, P^i(AB|XY)\} : \sum_i p_i P^i(AB|XY) = P(AB|XY) \right\}. \quad (\text{G4})$$

Hence, $\inf_{\tilde{P}(ABE|XY)} = \inf_{\{p_i, P^i(AB|XY)\} \in \mathcal{S}^{\text{all}}}$, and by virtue of Eq. (G3) we can rewrite definition 3 of the squashed nonlocality in the following way:

$$\begin{aligned} \mathcal{N}_{\text{sq}}(P(AB|XY)) = \max_{x,y} \inf_{\{p_i, P^i(AB|XY)\} \in \mathcal{S}^{\text{all}}} \\ \times \sum_i p_i I(A : B)_{\mathcal{M}_{x,y}^F P^i(AB|XY)}. \end{aligned} \quad (\text{G5})$$

From Eq. (G5), it is clear that the squashed nonlocality reduces to the convex roof extension of the mutual information function. This is analogous to the definition of entanglement for mixed quantum states [92], the only difference is that here we are not restricting the device to be decomposable in terms of only pure (extremal) devices (see in this context [37]).

1. Relation to the bound of Ref. [16]

To describe the relation between our results and the results in Ref. [16], we prove that $\max_{x,y} I_{\text{AMP},(x,y)} = \mathcal{N}_{\text{sq}}$. This allows us to compare the bounds on equal footing, and by showing that \mathcal{N}_{sq} is convex, to use the convexification method to achieve tighter bound than given in Ref. [16].

We first show that the \geq inequality. Indeed, let us fix (x, y) arbitrarily. Let $\{p(E = e)^*, P(ABE = e|XY)^*\}$ be an optimal ensemble achieving $I_{\text{AMP},(x,y)}$. By definition of the

complete extension [56], there exists a measurement¹⁶ z on its Eve's system E that generates this ensemble: $\{P(E = e|Z = z), P(ABE = e|XY, Z = z)\}$ so that $P(E = e|Z = z) = P(E = e)^*$ and $P(ABE = e|XY, Z = z) = P(ABE = e|XY)^*$. Since (x, y) was arbitrary and the z could be suboptimal for the definition of \mathcal{N}_{sq} we get the inequality $\max_{x,y} \mathbb{I}_{\text{AMP},(x,y)} \geq \mathcal{N}_{\text{sq}}$. To see that $\max_{x,y} \mathbb{I}_{\text{AMP},(x,y)} \leq \mathcal{N}_{\text{sq}}$, let x, y be fixed arbitrarily and $z(x, y)$ such that the value of $\inf_z \mathbb{I}(A : B \downarrow E)_{P(ABE|X=x, Y=y, Z=z)}$ is minimal. Then $\{P(E = e|Z = z(xy)), P(ABE = e|X = x, Y = y, Z = z(x, y))\}$ is a particular ensemble of $P(AB|XY)$, which may be suboptimal, i.e., not attaining infimum in definition of $\mathbb{I}_{\text{AMP},(x,y)}$, we get that $\mathbb{I}_{\text{AMP},(x,y)} \leq \inf_z \mathbb{I}(A : B \downarrow E)_{P(ABE|X=x, Y=y, Z=z)}$. Since (x, y) was arbitrary, we can take max over (x, y) on both sides, and on the r.h.s. we obtain \mathcal{N}_{sq} while the bound $\max_{(x,y)} \mathbb{I}_{\text{AMP},(x,y)}$ is on the l.h.s., which proves the claimed equality.

2. Positivity of the measure

Proposition 3. The squashed nonlocality is a positive semidefinite function of bipartite nonsignaling devices $P(AB|XY)$,

$$\mathcal{N}_{\text{sq}}(P(AB|XY)) \geq 0, \quad (\text{G6})$$

and the equality holds if the device P admits a local hidden variable model [8].

Proof. The intrinsic conditional mutual information satisfy $\mathbb{I}(A : B \downarrow E) \geq 0$ for all distributions $P(ABE)$, hence the positive semi-definiteness directly follows from its definition:

$$\begin{aligned} \mathcal{N}_{\text{sq}}(P(AB|XY)) &= \max_{x,y} \min_z \mathbb{I}(A : B \downarrow E)_{(\mathcal{M}_{x,y}^E \otimes \mathcal{M}_z^G) \mathcal{E}(P)(ABE|XYZ)} \\ &\geq \max_{x,y} \min_z 0 = 0. \end{aligned} \quad (\text{G7})$$

Now we have to show that it is zero for all local devices. Let us assume $P_L(AB|XY)$ is a local device, i.e., there exists a hidden variable model λ , such that $P_L(AB|XY) = \sum_{\lambda} P(A|X, \lambda) \otimes P(B|Y, \lambda) \rho(\lambda)$. This leads to an ensemble $\{\rho(\lambda), P(A|X, \lambda) \otimes P(B|Y, \lambda)\}$ whose members are tensor products of local devices, hence from Eq. (G5), we can directly write

$$\begin{aligned} \mathcal{N}_{\text{sq}}(P_L(AB|XY)) &= \max_{x,y} \sum_i \rho(\lambda_i) \mathbb{I}(A : B)_{\mathcal{M}_{x,y}^E(P(A|X, \lambda_i) \otimes P(B|Y, \lambda_i))} = 0. \end{aligned} \quad (\text{G8})$$

3. Convexity

Proposition 4. $\mathcal{N}_{\text{sq}}(P)$ is a convex function, i.e., if $P(AB|XY)$ and $Q(AB|XY)$ are two bipartite nonsignaling devices in the same polytope, then

$$\begin{aligned} \mathcal{N}_{\text{sq}}(\lambda P(AB|XY) + (1 - \lambda)Q(AB|XY)) &\leq \lambda \mathcal{N}_{\text{sq}}(P(AB|XY)) + (1 - \lambda) \mathcal{N}_{\text{sq}}(Q(AB|XY)) \end{aligned} \quad (\text{G9})$$

¹⁶Here, we mean the generalized measurement that gives the eavesdropper the access to any ensemble of the device (see Appendix B 2 for details).

$\forall \lambda \in [0, 1]$.

Proof. Consider the convex combination of the devices

$$\tilde{P}(AB|XY) = \lambda P(AB|XY) + (1 - \lambda)Q(AB|XY). \quad (\text{G10})$$

In particular there exists an extension $\tilde{P}_{\text{ext}}(ABE \wedge |XY)$ of $\tilde{P}(AB|XY)$, such that

$$\tilde{P}_{\text{ext}}(ABE \wedge = 0|XY) = p(\Lambda = 0) \tilde{P}(ABE|XY), \quad (\text{G11})$$

$$\tilde{P}_{\text{ext}}(ABE \wedge = 1|XY) = p(\Lambda = 1) \tilde{Q}(ABE|XY), \quad (\text{G12})$$

with $p(\Lambda = 0) = \lambda$ and $p(\Lambda = 1) = 1 - \lambda$. We consider that the devices $\tilde{P}(ABE|XY)$ and $\tilde{Q}(ABE|XY)$ are arbitrary extensions of the devices $P(AB|XY)$ and $Q(AB|XY)$ respectively, as discussed above.

Hence, from Eq. (G3), we have

$$\begin{aligned} \forall x, y \quad \inf_{\tilde{P}(ABE|XY)} \mathbb{I}(A : B|E)_{\mathcal{M}_{x,y}^E \tilde{P}(ABE|XY)} &\leq \mathbb{I}(A : B|E \wedge)_{\mathcal{M}_{x,y}^E \tilde{P}_{\text{ext}}(ABE \wedge |XY)} \\ &= \lambda \mathbb{I}(A : B|E)_{\mathcal{M}_{x,y}^E \tilde{P}(ABE|XY)} \\ &\quad + (1 - \lambda) \mathbb{I}(A : B|E)_{\mathcal{M}_{x,y}^E \tilde{Q}(ABE|XY)}, \end{aligned} \quad (\text{G13})$$

$$+ (1 - \lambda) \mathbb{I}(A : B|E)_{\mathcal{M}_{x,y}^E \tilde{Q}(ABE|XY)}, \quad (\text{G14})$$

where $\tilde{P}(ABE|XY)$ are such that $\sum_e \tilde{P}(ABE = e|XY) = \tilde{P}(AB|XY)$. The above relation holds for an arbitrary extensions of P and Q , the $\tilde{P}(ABE|XY)$ and $\tilde{Q}(ABE|XY)$ respectively. Hence, it is also true for the optimal extensions

$$\begin{aligned} \forall x, y \quad \inf_{\tilde{P}(ABE|XY)} \mathbb{I}(A : B|E)_{\mathcal{M}_{x,y}^E \tilde{P}(ABE|XY)} &\leq \lambda \inf_{\hat{P}(ABE|XY)} \mathbb{I}(A : B|E)_{\mathcal{M}_{x,y}^E \hat{P}(ABE|XY)} \\ &\quad + (1 - \lambda) \inf_{\hat{Q}(ABE|XY)} \mathbb{I}(A : B|E)_{\mathcal{M}_{x,y}^E \hat{Q}(ABE|XY)}, \end{aligned} \quad (\text{G15})$$

where $\hat{P}(ABE|XY)$ are such that $\sum_e \hat{P}(ABE = e|XY) = P(AB|XY)$ and $\hat{Q}(ABE|XY)$ are such that $\sum_e \hat{Q}(ABE = e|XY) = Q(AB|XY)$. Consider direct measurements \bar{x} and \bar{y} that maximize l.h.s. of inequality (G15). Then from Eq. (G3), we have

$$\begin{aligned} \mathcal{N}_{\text{sq}}(\tilde{P}(AB|XY)) &= \max_{x,y} \inf_{\tilde{P}(ABE|XY)} \mathbb{I}(A : B|E)_{\mathcal{M}_{x,y}^E \tilde{P}(ABE|XY)} \\ &= \inf_{\tilde{P}(ABE|XY)} \mathbb{I}(A : B|E)_{\mathcal{M}_{\bar{x}, \bar{y}}^E \tilde{P}(ABE|XY)} \end{aligned} \quad (\text{G16})$$

$$\begin{aligned} &\stackrel{(I)}{\leq} \lambda \inf_{\hat{P}(ABE|XY)} \mathbb{I}(A : B|E)_{\mathcal{M}_{\bar{x}, \bar{y}}^E \hat{P}(ABE|XY)} \\ &\quad + (1 - \lambda) \inf_{\hat{Q}(ABE|XY)} \mathbb{I}(A : B|E)_{\mathcal{M}_{\bar{x}, \bar{y}}^E \hat{Q}(ABE|XY)}. \end{aligned} \quad (\text{G17})$$

$$\begin{aligned} &\stackrel{(II)}{\leq} \lambda \max_{x,y} \inf_{\hat{P}(ABE|XY)} \mathbb{I}(A : B|E)_{\mathcal{M}_{x,y}^E \hat{P}(ABE|XY)} \\ &\quad + (1 - \lambda) \max_{x,y} \inf_{\hat{Q}(ABE|XY)} \mathbb{I}(A : B|E)_{\mathcal{M}_{x,y}^E \hat{Q}(ABE|XY)}. \end{aligned} \quad (\text{G18})$$

$$= \lambda \mathcal{N}_{\text{sq}}(P(AB|XY)) + (1 - \lambda) \mathcal{N}_{\text{sq}}(Q(AB|XY)), \quad (\text{G19})$$

where in (I), we use the inequality (G15), with $x = \bar{x}$ and $y = \bar{y}$. In (II), we use the fact that direct measure-

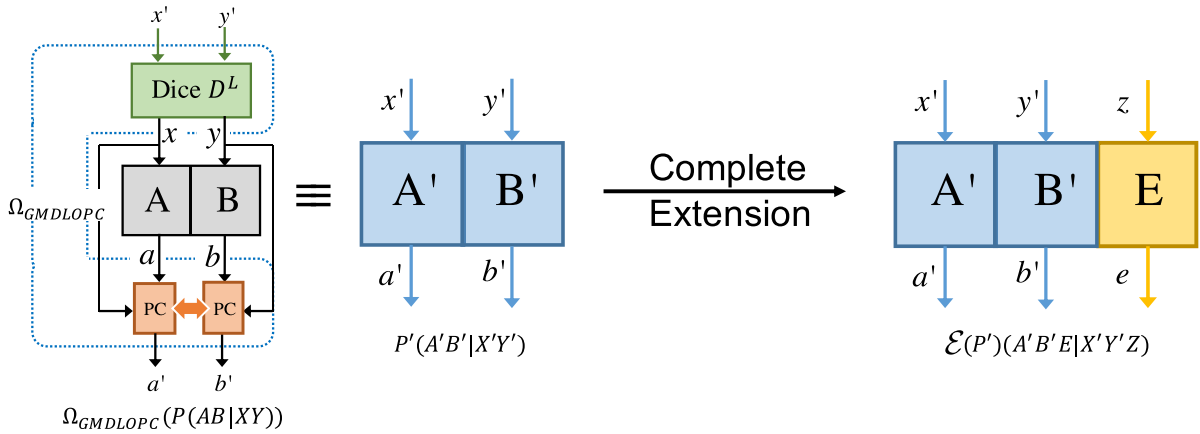


FIG. 7. Schematic diagram of the Ω_{GMDLOPC} operation, where the inputs of the devices shared by the honest parties are chosen by a local randomizer $D_{XY|X'Y'}^L(xy|x'y')$ as given in Eq. (G28). Similarly, the outputs are also connected through a post-processing channel $PC_{A'B'|ABXY}^L(a'b'|abxy)$ which also depends on the inputs x, y and has a local hidden variable model given in Eq. (G29).

ments \bar{x} and \bar{y} , may not maximize terms at r.h.s. of the inequality (G15). ■

4. Inheritance of monotonicity: Monotonicity under MDLOPC class of operation

In this section, we will show that any secrecy monotone (functional, nonincreasing under LOPC operations), after squashing procedure yields a functional which is monotonic under MDLOPC operations.

Proposition 5. (Inheritance of monotonicity) Any secrecy quantifier $M(A : B|E)$, which is nonincreasing under LOPC operations, after the squashing procedure is nonincreasing under MDLOPC operations.

Proof. Let us consider arbitrary MDLOPC operation Λ_{MDLOPC} . By definition it is a composition of the form $\Lambda_{\text{MDLOPC}} = \Lambda_{\text{LOPC}} \circ \mathcal{M}_{x_0, y_0}^F$. Let us also choose arbitrary device $P(ABE|XYZ)$ and let us fix arbitrarily $z = z_0$. As a consequence we can write a sequence of (in)equalities which we comment below, where for the sake of clarity of the proof we will use a short notation: $M(A : B|E) \equiv M$ and $\widehat{M}(A : B|E) \equiv \widehat{M}$.

$$\begin{aligned} & \widehat{M}(\Lambda_{\text{MDLOPC}}(P(ABE|XYZ))) \\ &= \max_{x,y} \min_z M(\Lambda_{\text{MDLOPC}}(P(ABE|X=x, Y=y, Z=z))) \end{aligned} \quad (\text{G20})$$

$$= \min_z M(\Lambda_{\text{LOPC}}(P(ABE|X=x_0, Y=y_0, Z=z))) \quad (\text{G21})$$

$$\leq M(\Lambda_{\text{LOPC}}(P(ABE|X=x_0, Y=y_0, Z=z_0))) \quad (\text{G22})$$

$$\leq M(P(ABE|X=x_0, Y=y_0, Z=z_0)) \quad (\text{G23})$$

$$\leq \max_{x,y} M(P(ABE|X=x, Y=y, Z=z_0)) \quad (\text{G24})$$

$$= \max_{x,y} \min_z M(P(ABE|X=x, Y=y, Z=z)) \quad (\text{G25})$$

$$= \widehat{M}(P(ABE|XYZ)). \quad (\text{G26})$$

In the first equality, we use the definition of \widehat{M} . In Eq. (G21), we use the fact that the device P after measurement \mathcal{M}_{x_0, y_0}^F has unary inputs in part of the honest parties (it becomes a distribution in that part), hence there is no parameter x, y to maximise over. The inequality (G22) follows from the property of minimum (over z). The inequality (G23) is due to the monotonicity of M under Λ_{LOPC} . The inequality (G24) is because x_0, y_0 may be suboptimal in $\max_{x,y}$ over $M(P(ABE|X=x, Y=y, Z=z_0))$. The equality (G25) comes from the fact that the choice of z_0 was arbitrary, so it is true for the z_0 that attains \min_z in (G25). The last equality comes from definition of $\widehat{M} \equiv \widehat{M}(A : B|E)$, which ends the proof. ■

From Proposition 5, it directly follows that the squashed nonlocality, is monotonic under Λ_{MDLOPC} , as it is defined based on the secrecy quantifier, intrinsic mutual information $I(A : B \downarrow E)$. And it is monotonic under LOPC operation [34,70].

Without using the above proposition, we can also independently prove that \mathcal{N}_{sq} is monotonic under MDLOPC operation, or in principle, under a larger class of operations, the GMDLOPC. We have mentioned in the main text, that the MDLOPC class of operations involve (i) direct measurement, changing devices into distributions followed by (ii) Local operations and Public communication. If we relax the measurement procedure and include all possible general measurements, then we will have the GMDLOPC class of operations, as shown in the schematic diagram in Fig. 7. Clearly $\text{MDLOPC} \subset \text{GMDLOPC}$, and one particular operation of GMDLOPC class will be denoted as Ω_{GMDLOPC} . Hence the monotonicity:

*Proposition 6.*¹⁷ The nonsignaling squashed nonlocality of any nonsignaling bipartite device P satisfies

$$\forall_{\Omega_{\text{GMDLOPC}}} \mathcal{N}_{\text{sq}}(\Omega_{\text{GMDLOPC}}(P)) \leq \mathcal{N}_{\text{sq}}(P), \quad (\text{G27})$$

Proof. To prove the monotonicity under GMDLOPC, we will use the equivalent definition of \mathcal{N}_{sq} given in Eq. (G5).

¹⁷The result of this section is partially based on Ref. [31].

Under the GMDLOPC operation $\Omega_{\text{GMDLOPC}} \in \text{GMDLOPC}$, the honest parties can choose general measurements in the input of the shared device $P(AB|XY)$. The general measurements can be chosen by using a public shared local randomness generator $D_{XY|X'Y'}^L(xy|x'y')$ (as depicted in Fig. 7), with $x' \in X', y' \in Y'$ the input and $x \in X, y \in Y$ are the output. As the output of D^L will be feeded to the input of $P(AB|XY)$, hence, we will assume without any loss of generality that both cardinality are same. Moreover, as D^L is

local randomness generator, hence

$$D_{XY|X'Y'}^L(xy|x'y') = \sum_{\lambda_1} \mu(\lambda_1) D^1(x|x'\lambda_1) D^2(y|y'\lambda_1), \quad (\text{G28})$$

where $\lambda_1 \in \Lambda_1$ is the local hidden variable and $\sum_{\lambda_1} \mu(\lambda_1) = 1$. Similarly, the outputs are also passed through a local post-processing channel $PC_{A'B'|ABXY}^L(a'b'|abxy)$, which also depends on the inputs of the initial device, as shown in Fig. 7. Additionally, the locality condition give rise to

$$PC_{A'B'|ABXY}^L(a'b'|abxy) = \sum_{\lambda_2} v(\lambda_2) PC^1(a'|ax\lambda_2) PC^2(b'|by\lambda_2), \quad (\text{G29})$$

with $\lambda_2 \in \Lambda_2$ and $\sum_{\lambda_2} v(\lambda_2) = 1$. Hence, under Ω_{GMDLOPC} , the given device $P(AB|XY)$ transforms into

$$P'_{A'B'|X'Y'}(a'b'|x'y') = \Omega_{\text{GMDLOPC}}(P(AB|XY)) \quad (\text{G30})$$

$$= \sum_{xy} D_{XY|X'Y'}^L(xy|x'y') \sum_{ab} P_{AB|XY}(ab|xy) PC_{A'B'|ABXY}^L(a'b'|abxy) \quad (\text{G31})$$

$$= \sum_{xy} \sum_{\lambda_1} \mu(\lambda_1) D^1(x|x'\lambda_1) D^2(y|y'\lambda_1) \sum_{ab} P(ab|xy) \sum_{\lambda_2} v(\lambda_2) PC^1(a'|ax\lambda_2) PC^2(b'|by\lambda_2). \quad (\text{G32})$$

Now the \mathcal{N}_{sq} of P' as in Eq. (G5) is

$$\mathcal{N}_{\text{sq}}(P') = \widehat{I}(A' : B' \downarrow E)_{\mathcal{E}(P')(A'B'E|X'Y'Z)} = \max_{x'y'} \inf_{\{p_i, P^i\} \in S^{\text{all}}(P')} \sum_i p_i I(A' : B')_{P^i}, \quad (\text{G33})$$

where $\mathcal{E}(P')(A'B'E|X'Y'Z)$, is the CE of $P'(A'B'|X'Y')$, and $S^{\text{all}}(P')$ denotes all possible ensembles of P' .

Consider the following tripartite device, resulting upon performing the Ω_{GMDLOPC} on the CE of $P(AB|XY)$,

$$\Omega_{\text{GMDLOPC}} \otimes \mathbb{1}_E(\mathcal{E}(P)(ABE|XYZ)) = \sum_{xy} D_{XY|X'Y'}^L(xy|x'y') \sum_{ab} \mathcal{E}(P)(abe|xyz) PC_{A'B'|ABXY}^L(a'b'|abxy) \quad (\text{G34})$$

$$= \sum_{xy} \sum_{\lambda_1} \mu(\lambda_1) D^1(x|x'\lambda_1) D^2(y|y'\lambda_1) \sum_{ab} \mathcal{E}(P)(abe|xyz) \\ \times \sum_{\lambda_2} v(\lambda_2) PC^1(a'|ax\lambda_2) PC^2(b'|by\lambda_2). \quad (\text{G35})$$

Here $\mathbb{1}_E$ means the identity operator in Eve's subsystem.

Consider the ensemble $\{p(e|z)\mu(\lambda_1)v(\lambda_2), P^{ez\lambda_1\lambda_2}(a'b'|x'y')\}$, where

$$P^{ez\lambda_1\lambda_2}(a'b'|x'y') = \sum_{xy} D^1(x|x'\lambda_1) D^2(y|y'\lambda_1) \sum_{ab} P^{ez}(ab|xy) PC^1(a'|ax\lambda_2) PC^2(b'|by\lambda_2). \quad (\text{G36})$$

Now we will show that the above ensemble will be an ensemble of $P'(A'B'|X'Y')$, if $\{p(e|z), P_{AB|XY}^{ez}\}$ is an ensemble of $P(AB|XY)$.

Suppose $\{p(e|z), P_{AB|XY}^{ez}\}$ is an ensemble of P , then

$$\sum_{e\lambda_1\lambda_2} p(e|z)\mu(\lambda_1)v(\lambda_2)P^{ez\lambda_1\lambda_2}(a'b'|x'y') \quad (\text{G37})$$

$$= \sum_{e\lambda_1\lambda_2} p(e|z)\mu(\lambda_1)v(\lambda_2) \sum_{xy} D^1(x|x'\lambda_1) D^2(y|y'\lambda_1) \sum_{ab} P^{ez}(ab|xy) PC^1(a'|ax\lambda_2) PC^2(b'|by\lambda_2) \quad (\text{G38})$$

$$= \sum_{xy} \sum_{\lambda_1} \mu(\lambda_1) D^1(x|x'\lambda_1) D^2(y|y'\lambda_1) \sum_{ab} \left(\sum_e p(e|z) P^{ez}(ab|xy) \right) \sum_{\lambda_2} v(\lambda_2) PC^1(a'|ax\lambda_2) PC^2(b'|by\lambda_2) \quad (\text{G39})$$

$$= \sum_{xy} \sum_{\lambda_1} \mu(\lambda_1) D^1(x|x'\lambda_1) D^2(y|y'\lambda_1) \sum_{ab} P(ab|xy) \sum_{\lambda_2} v(\lambda_2) PC^1(a'|ax\lambda_2) PC^2(b'|by\lambda_2) \quad (\text{G40})$$

$$= P'_{A'B'|X'Y'}(a'b'|x'y'), \quad (\text{G41})$$

by using Eq. (G32) and the fact that $\sum_e p(e|z) P^{ez}(ab|xy) = P(ab|xy)$.

Moreover, $\{p(e|z), P_{AB|XY}^{ez}\}$ is an arbitrary ensemble, and Eve can easily access it once she has the CE $\mathcal{E}(P)(ABE|XYZ)$:

Because $\{p(e|z)\mu(\lambda_1)\nu(\lambda_2), P^{e\lambda_1\lambda_2}(a'b'|x'y')\}$ is an ensemble of P' ,

$$\forall_{x',y'} \inf_{\{p_i, P^i\} \in S^{all}(P')} \sum_i p_i I(A' : B')_{P^i} \leq \sum_{e\lambda_1\lambda_2} p(e|z)\mu(\lambda_1)\nu(\lambda_2) I(A' : B')_{P^{e\lambda_1\lambda_2}(A'B'|X'=x',Y'=y')} \quad (G42)$$

$$\stackrel{(I)}{\leq} \sum_{e\lambda_1\lambda_2} p(e|z)\mu(\lambda_1)\nu(\lambda_2) I(AX : BY)_{P^{e\lambda_1\lambda_2}(AXBY|X'=x',Y'=y')} \quad (G43)$$

$$\stackrel{(II)}{=} \sum_{e\lambda_1} p(e|z)\mu(\lambda_1) (I(A : B|XY) + I(X : B|Y) + I(A : Y|X) + I(X : Y))_{P^{e\lambda_1}(AXBY|X'=x',Y'=y')} \quad (G44)$$

$$\stackrel{(III)}{=} \sum_{e\lambda_1} p(e|z)\mu(\lambda_1) I(A : B|XY)_{P^{e\lambda_1}(AXBY|X'=x',Y'=y')} \quad (G45)$$

$$= \sum_{exy\lambda_1} p(e|z)\mu(\lambda_1) D^1(x|x'\lambda_1) D^2(y|y'\lambda_1) I(A : B)_{P^{e\lambda_1}(AB|X=x, Y=y, X'=x', Y'=y')} \quad (G46)$$

$$\stackrel{(IV)}{\leq} \max_{xy} \sum_e p(e|z) I(A : B)_{P^{e\lambda_1}(AB|XY)}, \quad (G47)$$

where in (I) we use the data processing inequality and also use the fact that the distribution $P_{AXBY|X'Y'}^{e\lambda_1\lambda_2}(axy|x'y') = D^1(x|x'\lambda_1)D^2(y|y'\lambda_1)P_{AB|XY}^{e\lambda_1}(ab|xy) \sum_{a'b'} PC^1(a'|ax\lambda_2)PC^2(b'|by\lambda_2)$ is independent of λ_2 . The chain rule of mutual information has been used in (II) whereas in (III), we use the fact that given x', y' and λ_1 , the random variables X and Y are independent, hence $I(X : B|Y) = I(A : Y|X) = 0$, which follows from the nonsignaling condition. In (IV) we simply write $P^{e\lambda_1}(AB|X=x, Y=y, X'=x', Y'=y') = P^{e\lambda_1}(AB|X=x, Y=y)$.

The r.h.s. of (G47) is valid for an arbitrary ensemble $\{p(e|z), P^{e\lambda_1}\} \in S^{all}(P)$, so it is still valid when taking infimum over all ensembles. Hence,

$$\max_{x'y'} \inf_{\{p_i, P^i\} \in S^{all}(P)} \sum_i p_i I(A' : B')_{P^i(A'B'|X'Y')} \leq \max_{xy} \inf_{\{p_i, P^i\} \in S^{all}(P)} \sum_i p_i I(A : B)_{P^i(AB|XY)}, \quad (G48)$$

$$\Rightarrow \mathcal{N}_{sq}(\Omega_{\text{GMDLOPC}}(P)) \leq \mathcal{N}_{sq}(P). \quad (G49)$$

As MDLOPC \subset GMDLOPC, so we have

Corollary 3. The nonsignaling squashed nonlocality of any nonsignaling bipartite device P satisfies

$$\forall_{\Lambda_{\text{MDLOPC}}} \mathcal{N}_{sq}(\Lambda_{\text{MDLOPC}}(P)) \leq \mathcal{N}_{sq}(P), \quad (G50)$$

The above monotonicity property also holds for the nonsignaling squashed conditional mutual information $\widehat{I}(A : B|E)_{\mathcal{E}(P)(ABE|XYZ)}$.

5. Superadditivity and additivity

*Proposition 7.*¹⁸ If two bipartite nonsignaling devices $P(A_1B_1|X_1Y_1)$ and $Q(A_2B_2|X_2Y_2)$ are the marginals of a four partite nonsignaling device $\bar{P}(A_1A_2B_1B_2|X_1X_2Y_1Y_2)$, then the

nonsignaling squashed nonlocality \mathcal{N}_{sq} is superadditive,

$$\mathcal{N}_{sq}(\bar{P}(A_1A_2B_1B_2|X_1X_2Y_1Y_2)) \geq \mathcal{N}_{sq}(P(A_1B_1|X_1Y_1)) + \mathcal{N}_{sq}(Q(A_2B_2|X_2Y_2)), \quad (G51)$$

and additive for tensor product of devices $P(A_1B_1|X_1Y_1) \otimes Q(A_2B_2|X_2Y_2)$, that is

$$\mathcal{N}_{sq}(P(A_1B_1|X_1Y_1) \otimes Q(A_2B_2|X_2Y_2)) = \mathcal{N}_{sq}(P(A_1B_1|X_1Y_1)) + \mathcal{N}_{sq}(Q(A_2B_2|X_2Y_2)). \quad (G52)$$

Proof. Superadditivity on joint device: Let us consider two devices $P(A_1B_1|X_1Y_1)$ and $Q(A_2B_2|X_2Y_2)$, which are the marginals of a big four party nonsignaling device $\bar{P}(A_1A_2B_1B_2|X_1X_2Y_1Y_2)$, i.e.,

$$\sum_{a_2b_2} \bar{P}(a_1, a_2, b_1, b_2|x_1, x_2, y_1, y_2) = P(a_1, b_1|x_1, y_1) \forall a_1, b_1, x_1, x_2, y_1, y_2, \quad (G53)$$

$$\sum_{a_1b_1} \bar{P}(a_1, a_2, b_1, b_2|x_1, x_2, y_1, y_2) = Q(a_2, b_2|x_2, y_2), \quad \forall a_2, b_2, x_1, x_2, y_1, y_2. \quad (G54)$$

where $\bar{P}(A_1=a_1, A_2=a_2, B_1=b_1, B_2=b_2|X_1=x_1, X_2=x_2, Y_1=y_1, Y_2=y_2) \equiv \bar{P}(a_1, a_2, b_1, b_2|x_1, x_2, y_1, y_2)$, $P(A_1=a_1, B_1=b_1|X_1=x_1, Y_1=y_1) = P(a_1, b_1|x_1, y_1)$, and $Q(A_2=a_2, B_2=b_2|X_2=x_2, Y_2=y_2) = Q(a_2, b_2|x_2, y_2)$. Moreover, $\bar{P}(A_1A_2B_1B_2|X_1X_2Y_1Y_2)$ is also satisfy nonsignaling conditions among all of its parties, as defined in Eqs. (B1) and (B2).

Consider an arbitrary nonsignaling extension of $\bar{P}(A_1A_2B_1B_2|X_1X_2Y_1Y_2) \rightarrow \bar{P}(A_1A_2B_1B_2E|X_1X_2Y_1Y_2Z)$, with unary input $|Z\rangle$ in the extended part. The input is unary, so the nonsignaling condition is automatic and we can omit the Z . The conditional mutual information of the distribution after performing an arbitrary pair of direct measurements, i.e., $\mathcal{M}_{x_1, y_1}^F \otimes \mathcal{M}_{x_2, y_2}^F$ on the inputs X_1, Y_1 and X_2, Y_2

¹⁸The result of this section is partially based on Ref. [31].

reads

$$\begin{aligned} & \forall x_1, x_2, y_1, y_2 \\ & \mathbf{I}(A_1 A_2 : B_1 B_2 | E)_{(\mathcal{M}_{x_1, y_1}^F \otimes \mathcal{M}_{x_2, y_2}^F) \bar{P}(A_1 A_2 B_1 B_2 E | X_1 X_2 Y_1 Y_2)} \\ & \stackrel{(I)}{=} (\mathbf{I}(A_1 : B_1 | E) + \mathbf{I}(A_2 : B_1 | E A_1) + \mathbf{I}(A_1 : B_2 | E B_1) + \mathbf{I}(A_2 : B_2 | E A_1 B_1))_{(\mathcal{M}_{x_1, y_1}^F \otimes \mathcal{M}_{x_2, y_2}^F) \bar{P}(A_1 A_2 B_1 B_2 E | X_1 X_2 Y_1 Y_2)} \quad (\text{G55}) \end{aligned}$$

$$\stackrel{(II)}{\geq} \mathbf{I}(A_1 : B_1 | E)_{\mathcal{M}_{x_1, y_1}^F \bar{P}(A_1 B_1 E | X_1 Y_1)} + \mathbf{I}(A_2 : B_2 | E A_1 B_1)_{(\mathcal{M}_{x_1, y_1}^F \otimes \mathcal{M}_{x_2, y_2}^F) \bar{P}(A_1 A_2 B_1 B_2 E | X_1 X_2 Y_1 Y_2)}, \quad (\text{G56})$$

where we use the chain rule of mutual information in (I) and in (II), we use positivity condition of mutual information. $\mathcal{M}_{x_1, y_1}^F \bar{P}(A_1 B_1 E | X_1 Y_1)$ is the marginal of the device $(\mathcal{M}_{x_1, y_1}^F \otimes \mathcal{M}_{x_2, y_2}^F) \bar{P}(A_1 A_2 B_1 B_2 E | X_1 X_2 Y_1 Y_2)$ after the direct measurements on the inputs. Recall that

$$(\mathcal{M}_{x_1, y_1}^F \otimes \mathcal{M}_{x_2, y_2}^F) \bar{P}(A_1 A_2 B_1 B_2 E | X_1 X_2 Y_1 Y_2) = \bar{P}(A_1 A_2 B_1 B_2 E | X_1 = x_1, X_2 = x_2, Y_1 = y_1, Y_2 = y_2). \quad (\text{G57})$$

Noticing that $\bar{P}(A_1 B_1 E | X_1 Y_1)$ is an arbitrary extension of $P(A_1 B_1 | X_1 Y_1)$ and similarly $\bar{P}(A_1 A_2 B_1 B_2 E | X_1 X_2 Y_1 Y_2)$ is for the device $Q(A_1 B_1 E | X_1 Y_1)$, we can write

$$\begin{aligned} & \forall x_1, x_2, y_1, y_2 \\ & \mathbf{I}(A_1 : B_1 | E)_{\mathcal{M}_{x_1, y_1}^F \bar{P}(A_1 B_1 E | X_1 Y_1)} \geq \inf_{\bar{P}(A_1 B_1 E | X_1 Y_1)} \mathbf{I}(A_1 : B_1 | E)_{\mathcal{M}_{x_1, y_1}^F \bar{P}(A_1 B_1 E | X_1 Y_1)}, \quad (\text{G58}) \end{aligned}$$

$$\mathbf{I}(A_2 : B_2 | E A_1 B_1)_{(\mathcal{M}_{x_1, y_1}^F \otimes \mathcal{M}_{x_2, y_2}^F) \bar{P}(A_1 A_2 B_1 B_2 E | X_1 X_2 Y_1 Y_2)} \geq \inf_{\bar{Q}(A_2 B_2 E | X_2 Y_2)} \mathbf{I}(A_2 : B_2 | E)_{\mathcal{M}_{x_2, y_2}^F \bar{Q}(A_2 B_2 E | X_2 Y_2)}. \quad (\text{G59})$$

From inequalities (G56), (G58), and (G59), we have

$$\begin{aligned} & \forall x_1, x_2, y_1, y_2 \\ & \mathbf{I}(A_1 A_2 : B_1 B_2 | E)_{(\mathcal{M}_{x_1, y_1}^F \otimes \mathcal{M}_{x_2, y_2}^F) \bar{P}(A_1 A_2 B_1 B_2 E | X_1 X_2 Y_1 Y_2)} \\ & \geq \inf_{\bar{P}(A_1 B_1 E | X_1 Y_1)} \mathbf{I}(A_1 : B_1 | E)_{\mathcal{M}_{x_1, y_1}^F \bar{P}(A_1 B_1 E | X_1 Y_1)} + \inf_{\bar{Q}(A_2 B_2 E | X_2 Y_2)} \mathbf{I}(A_2 : B_2 | E)_{\mathcal{M}_{x_2, y_2}^F \bar{Q}(A_2 B_2 E | X_2 Y_2)}. \quad (\text{G60}) \end{aligned}$$

The above inequality holds for all extensions of $\bar{P}(A_1 A_2 B_1 B_2 | X_1 X_2 Y_1 Y_2)$, hence also for an optimal extension on the l.h.s., so

$$\begin{aligned} & \forall x_1, x_2, y_1, y_2 \\ & \inf_{\bar{P}(A_1 A_2 B_1 B_2 E | X_1 X_2 Y_1 Y_2)} \mathbf{I}(A_1 A_2 : B_1 B_2 | E)_{(\mathcal{M}_{x_1, y_1}^F \otimes \mathcal{M}_{x_2, y_2}^F) \bar{P}(A_1 A_2 B_1 B_2 E | X_1 X_2 Y_1 Y_2)} \\ & \geq \inf_{\bar{P}(A_1 B_1 E | X_1 Y_1)} \mathbf{I}(A_1 : B_1 | E)_{\mathcal{M}_{x_1, y_1}^F \bar{P}(A_1 B_1 E | X_1 Y_1)} + \inf_{\bar{Q}(A_2 B_2 E | X_2 Y_2)} \mathbf{I}(A_2 : B_2 | E)_{\mathcal{M}_{x_2, y_2}^F \bar{Q}(A_2 B_2 E | X_2 Y_2)}. \quad (\text{G61}) \end{aligned}$$

Suppose that \bar{x}_1, \bar{y}_1 are the optimal direct measurement choice for $\mathcal{N}_{\text{sq}}(P)$ and \bar{x}_2, \bar{y}_2 are for $\mathcal{N}_{\text{sq}}(Q)$,

$$\mathcal{N}_{\text{sq}}(P(A_1 B_1 | X_1 Y_1)) = \max_{x_1, y_1} \inf_{\bar{P}(A_1 B_1 E | X_1 Y_1)} \mathbf{I}(A_1 : B_1 | E)_{\mathcal{M}_{x_1, y_1}^F \bar{P}(A_1 B_1 E | X_1 Y_1)} = \inf_{\bar{P}(A_1 B_1 E | X_1 Y_1)} \mathbf{I}(A_1 : B_1 | E)_{\mathcal{M}_{\bar{x}_1, \bar{y}_1}^F \bar{P}(A_1 B_1 E | X_1 Y_1)}, \quad (\text{G62})$$

$$\mathcal{N}_{\text{sq}}(Q(A_2 B_2 | X_2 Y_2)) = \max_{x_2, y_2} \inf_{\bar{Q}(A_2 B_2 E | X_2 Y_2)} \mathbf{I}(A_2 : B_2 | E)_{\mathcal{M}_{x_2, y_2}^F \bar{Q}(A_2 B_2 E | X_2 Y_2)} = \inf_{\bar{Q}(A_2 B_2 E | X_2 Y_2)} \mathbf{I}(A_2 : B_2 | E)_{\mathcal{M}_{\bar{x}_2, \bar{y}_2}^F \bar{Q}(A_2 B_2 E | X_2 Y_2)}. \quad (\text{G63})$$

Finally,

$$\mathcal{N}_{\text{sq}}(\bar{P}(A_1 B_1 A_2 B_2 | X_1 X_2 Y_1 Y_2)) = \max_{x_1, y_1, x_2, y_2} \inf_{\bar{P}(A_1 B_1 A_2 B_2 E | X_1 X_2 Y_1 Y_2)} \mathbf{I}(A_1 A_2 : B_1 B_2 | E)_{(\mathcal{M}_{x_1, y_1}^F \otimes \mathcal{M}_{x_2, y_2}^F) \bar{P}(A_1 A_2 B_1 B_2 E | X_1 X_2 Y_1 Y_2)} \quad (\text{G64})$$

$$\stackrel{(I)}{\geq} \inf_{\bar{P}(A_1 B_1 A_2 B_2 E | X_1 X_2 Y_1 Y_2)} \mathbf{I}(A_1 A_2 : B_1 B_2 | E)_{(\mathcal{M}_{\bar{x}_1, \bar{y}_1}^F \otimes \mathcal{M}_{\bar{x}_2, \bar{y}_2}^F) \bar{P}(A_1 A_2 B_1 B_2 E | X_1 X_2 Y_1 Y_2)} \quad (\text{G65})$$

$$\stackrel{(II)}{\geq} \inf_{\bar{P}(A_1 B_1 E | X_1 Y_1)} \mathbf{I}(A_1 : B_1 | E)_{\mathcal{M}_{\bar{x}_1, \bar{y}_1}^F \bar{P}(A_1 B_1 E | X_1 Y_1)} + \inf_{\bar{Q}(A_2 B_2 E | X_2 Y_2)} \mathbf{I}(A_2 : B_2 | E)_{\mathcal{M}_{\bar{x}_2, \bar{y}_2}^F \bar{Q}(A_2 B_2 E | X_2 Y_2)}, \quad (\text{G66})$$

$$\stackrel{(III)}{=} \mathcal{N}_{\text{sq}}(P(A_1 B_1 | X_1 Y_1)) + \mathcal{N}_{\text{sq}}(Q(A_2 B_2 | X_2 Y_2)). \quad (\text{G67})$$

In (I), we use a specific choice of direct measurement, $\mathcal{M}_{\bar{x}_1, \bar{y}_1}^F \otimes \mathcal{M}_{\bar{x}_2, \bar{y}_2}^F$, which may not be optimal for device $\bar{P}(A_1 B_1 A_2 B_2 | X_1 X_2 Y_1 Y_2)$. We use Eq. (G61) for the direct measurements $\mathcal{M}_{\bar{x}_1, \bar{y}_1}^F \otimes \mathcal{M}_{\bar{x}_2, \bar{y}_2}^F$ in (II) and finally in (III), Eqs. (G62) and (G63) has been used.

Additivity for tensor product of devices: Let us assume that the joint nonsignaling four party device (two random variables for input and output in the honest parties' part) is the tensor product [76] of two bipartite devices,

$$\bar{P}(A_1 B_1 A_2 B_2 | X_1 X_2 Y_1 Y_2) = P(A_1 B_1 | X_1 Y_1) \otimes Q(A_2 B_2 | X_2 Y_2). \quad (\text{G68})$$

Consider the (nonsignaling) extensions with unary inputs of both the devices, $P(A_1 B_1 | X_1 Y_1) \rightarrow \bar{P}(A_1 B_1 E_1 | X_1 Y_1)$ and $Q(A_2 B_2 | X_2 Y_2) \rightarrow \bar{Q}(A_2 B_2 E_2 | X_2 Y_2)$, which are the optimal extensions for calculating \mathcal{N}_{sq} for both the devices, as given in Eq. (G3), for all x and y . Hence, their tensor product $\bar{P}(A_1 B_1 E_1 | X_1 Y_1) \otimes \bar{Q}(A_2 B_2 E_2 | X_2 Y_2)$ is an extension of $\bar{P}(A_1 B_1 A_2 B_2 | X_1 X_2 Y_1 Y_2)$, which may not be optimal one, resulting in

$$\forall x_1, x_2, y_1, y_2$$

$$\begin{aligned} & \inf_{\bar{P}(A_1 B_1 A_2 B_2 E | X_1 X_2 Y_1 Y_2)} \mathbf{I}(A_1 A_2 : B_1 B_2 | E)_{(\mathcal{M}_{x_1, y_1}^F \otimes \mathcal{M}_{x_2, y_2}^F)} \bar{P}(A_1 A_2 B_1 B_2 E | X_1 X_2 Y_1 Y_2) \\ & \leq \mathbf{I}(A_1 A_2 : B_1 B_2 | E_1 E_2)_{(\mathcal{M}_{x_1, y_1}^F \otimes \mathcal{M}_{x_2, y_2}^F)} \bar{P}(A_1 B_1 E_1 | X_1 Y_1) \otimes \bar{Q}(A_2 B_2 E_2 | X_2 Y_2) \end{aligned} \quad (\text{G69})$$

$$= \mathbf{I}(A_1 : B_1 | E_1)_{\mathcal{M}_{x_1, y_1}^F} \bar{P}(A_1 B_1 E_1 | X_1 Y_1) + \mathbf{I}(A_2 : B_2 | E_2)_{\mathcal{M}_{x_2, y_2}^F} \bar{Q}(A_2 B_2 E_2 | X_2 Y_2) \quad (\text{G70})$$

$$= \inf_{P(A_1 B_1 E_1 | X_1 Y_1)} \mathbf{I}(A_1 : B_1 | E_1)_{\mathcal{M}_{x_1, y_1}^F} P(A_1 B_1 E_1 | X_1 Y_1) + \inf_{Q(A_2 B_2 E_2 | X_2 Y_2)} \mathbf{I}(A_2 : B_2 | E_2)_{\mathcal{M}_{x_2, y_2}^F} Q(A_2 B_2 E_2 | X_2 Y_2). \quad (\text{G71})$$

Considering the optimal direct measurements $\mathcal{M}_{\bar{x}_1, \bar{y}_1}^F \otimes \mathcal{M}_{\bar{x}_2, \bar{y}_2}^F$ in the l.h.s. of the above relation, gives

$$\begin{aligned} & \mathcal{N}_{\text{sq}}(P(A_1 B_1 | X_1 Y_1) \otimes Q(A_2 B_2 | X_2 Y_2)) \\ & = \max_{x_1, x_2, y_1, y_2} \inf_{\bar{P}(A_1 B_1 A_2 B_2 E | X_1 X_2 Y_1 Y_2)} \mathbf{I}(A_1 A_2 : B_1 B_2 | E)_{(\mathcal{M}_{x_1, y_1}^F \otimes \mathcal{M}_{x_2, y_2}^F)} \bar{P}(A_1 A_2 B_1 B_2 E | X_1 X_2 Y_1 Y_2) \end{aligned} \quad (\text{G72})$$

$$= \inf_{\bar{P}(A_1 B_1 A_2 B_2 E | X_1 X_2 Y_1 Y_2)} \mathbf{I}(A_1 A_2 : B_1 B_2 | E)_{(\mathcal{M}_{\bar{x}_1, \bar{y}_1}^F \otimes \mathcal{M}_{\bar{x}_2, \bar{y}_2}^F)} \bar{P}(A_1 A_2 B_1 B_2 E | X_1 X_2 Y_1 Y_2) \quad (\text{G73})$$

$$\leq \inf_{P(A_1 B_1 E_1 | X_1 Y_1)} \mathbf{I}(A_1 : B_1 | E_1)_{\mathcal{M}_{\bar{x}_1, \bar{y}_1}^F} P(A_1 B_1 E_1 | X_1 Y_1) + \inf_{Q(A_2 B_2 E_2 | X_2 Y_2)} \mathbf{I}(A_2 : B_2 | E_2)_{\mathcal{M}_{\bar{x}_2, \bar{y}_2}^F} Q(A_2 B_2 E_2 | X_2 Y_2) \quad (\text{G74})$$

$$\leq \max_{x_1, y_1} \inf_{P(A_1 B_1 E_1 | X_1 Y_1)} \mathbf{I}(A_1 : B_1 | E_1)_{\mathcal{M}_{x_1, y_1}^F} P(A_1 B_1 E_1 | X_1 Y_1) + \max_{x_2, y_2} \inf_{Q(A_2 B_2 E_2 | X_2 Y_2)} \mathbf{I}(A_2 : B_2 | E_2)_{\mathcal{M}_{x_2, y_2}^F} Q(A_2 B_2 E_2 | X_2 Y_2) \quad (\text{G75})$$

$$= \mathcal{N}_{\text{sq}}(P(A_1 B_1 | X_1 Y_1)) + \mathcal{N}_{\text{sq}}(Q(A_2 B_2 | X_2 Y_2)). \quad (\text{G76})$$

Using relation (G67), we finish the proof with equality:

$$\mathcal{N}_{\text{sq}}(P(A_1 B_1 | X_1 Y_1) \otimes Q(A_2 B_2 | X_2 Y_2)) = \mathcal{N}_{\text{sq}}(P(A_1 B_1 | X_1 Y_1)) + \mathcal{N}_{\text{sq}}(Q(A_2 B_2 | X_2 Y_2)). \quad (\text{G77})$$

6. Subextensivity

Proposition 8. Nonsignaling squashed nonlocality is bounded by $\log_2(\min\{d_A, d_B\})$.

Proof. From the definition of nonsignaling squashed nonlocality given in Eq. (G5), we have

$$\begin{aligned} & \mathcal{N}_{\text{sq}}(P(AB|XY)) \\ & = \max_{x, y} \inf_{\{p_i, P^i(AB|XY)\} \in \mathcal{S}^{\text{all}}} \sum_i p_i \mathbf{I}(A : B)_{\mathcal{M}_{x, y}^F, P^i(AB|XY)} \end{aligned} \quad (\text{G78})$$

$$\stackrel{(I)}{\leq} \max_{x, y} \inf_{\{p_i, P^i(AB|XY)\} \in \mathcal{S}^{\text{all}}} \sum_i p_i \log_2(\min\{d_A^x, d_B^y\}) \quad (\text{G79})$$

$$\leq \log_2(\min\{d_A, d_B\}). \quad (\text{G80})$$

where in (I), we use the fact that $\mathbf{I}(A : B)_{\mathcal{M}_{x, y}^F, P^i(AB|XY)} \leq \log_2(\min\{d_A^x, d_B^y\})$ for all i , and $d_A^x = \text{supp}P(A|X = x)$ and $d_B^y = \text{supp}P(B|Y = y)$ and $d_A = \max_x \text{supp}P(A|X = x)$ and $d_B = \max_y \text{supp}P(B|Y = y)$.

APPENDIX H: NONLOCALITY COST AS AN UPPER BOUND

Definition 16. The nonlocality cost of bipartite nonsignaling device is

$$\mathcal{N}_C(P) := C(P) \log_2(\min\{d_A, d_B\}), \quad (\text{H1})$$

where $d_A = \max_x(\text{supp}\mathcal{M}_x^F(P(A|X)))$ and $d_B = \max_y(\text{supp}\mathcal{M}_y^F(P(B|Y)))$ are dimensions of the outputs, and $C(P)$ is the nonlocality fraction of P [93,94].

Proposition 9. The secret-key rate $K_{DI}^{(\text{iid})}(P)$ of a device is upper bounded by

$$\mathcal{N}_C(P) \geq K_{DI}^{(\text{iid})}(P). \quad (\text{H2})$$

Proof. Suppose Alice and Bob share a nonsignaling device $P \equiv P(AB|XY)$, and Eve has access to its complete extension [56]. The device P can be decomposed into a nonlocal vertex and a local device,

$$P = \alpha P_{\text{NL}}^V + (1 - \alpha) P_L, \quad (\text{H3})$$

where P_{NL}^V is the nonlocal vertex and P_L is any local device. Let us denote the nonlocality fraction

$$C(P) := \min_{\text{All decompositions as in Eq. (H3)}} \alpha. \quad (\text{H4})$$

Eve can always get access to this ensemble, $\{(C(P), \bar{P}_{\text{NL}}^V), (1 - C(P), \bar{P}_L)\}$, in part of the honest parties.

We assume that Eve works in favor of Alice and Bob, and informs them about her output when she obtains the above ensemble. The key rate \tilde{K} , in this scenario, must be greater than in NSDI-iid scenario, since in the latter case Eve does not work on account of Alice and Bob,

$$K_{DI}^{(\text{iid})}(P) \leq \tilde{K}(P). \quad (\text{H5})$$

With a probability $C(P)$ the honest parties share the nonlocal correlations, useful for secret-key agreement and with probability $1 - C(P)$, they share a local device with zero key rates. Since the key satisfying Maurer's security definition is upper bounded by mutual information function, and both of them are nonincreasing under the LOPC operations, we obtain

$$\tilde{K}(P) \leq C(P) (\max_{x,y} I(A : B)_{\mathcal{M}_{x,y}^F P_{NL}^V(AB|XY)}). \quad (\text{H6})$$

Furthermore,

$$I(A : B)_{\mathcal{M}_{x,y}^F(P(AB|XY))} \leq \log_2(\min\{d_A^x, d_B^y\}), \quad (\text{H7})$$

where $d_A^x = \text{supp}P(A|X=x)$ and $d_B^y = \text{supp}P(B|Y=y)$. Employing Eq. (H5), we finally obtain

$$K_{DI}^{(\text{iid})}(P) \leq C(P) \left(\sup_{\mathcal{M}_{x,y}^F} \log_2(\min\{d_A^x, d_B^y\}) \right) \quad (\text{H8})$$

$$= C(P) \log_2(\min\{d_A, d_B\}) = \mathcal{N}_C(P), \quad (\text{H9})$$

by Definition 16, with $d_A = \max_x \text{supp}P(A|X=x)$ and $d_B = \max_y \text{supp}P(B|Y=y)$. ■

APPENDIX I: EXAMPLES OF SECRECY MONOTONES, CONVEXIFICATION OF $\widehat{I}(A : B \downarrow E)$ AND A NONTRIVIAL BOUND

Monotones, based on mutual information functions, are used to upper bound the secret-key rate on the SKA scenario. However, the only one amongst them, which is easily computable, is the mutual information itself. All of them can be "squashed" and used to generate the upper bounds for $K_{DI}^{(\text{iid})}$.

Fact 1. The secrecy quantifiers and monotones [34] (and the mutual information function) are the upper bounds on $S(A : B||E)$:

$$I(A : B)_{P(ABE)} \geq S(A : B||E)_{P(ABE)}, \quad (\text{I1})$$

$$I(A : B|E)_{P(ABE)} \geq S(A : B||E)_{P(ABE)}, \quad (\text{I2})$$

$$\min\{I(A : B)_{P(ABE)}, I(A : B|E)_{P(ABE)}\} \geq S(A : B||E)_{P(ABE)}, \quad (\text{I3})$$

$$I(A : B \downarrow E)_{P(ABE)} \geq I(A : B \downarrow \downarrow E)_{P(ABE)} \geq S(A : B||E)_{P(ABE)}. \quad (\text{I4})$$

We can use all of the functions displayed in Fact 1 to construct the nonsignaling squashed secrecy quantifiers and monotones for the devices. See Appendix A for the proper definition of the above functions.

Corollary 4. The following upper bounds on $K_{DI}^{(\text{iid})}(P)$ hold

$$\widehat{I}(A : B)_{\mathcal{E}(P)(ABE|XYZ)} \geq K_{DI}^{(\text{iid})}(P), \quad (\text{I5})$$

$$\widehat{I}(A : B|E)_{\mathcal{E}(P)(ABE|XYZ)} \geq K_{DI}^{(\text{iid})}(P), \quad (\text{I6})$$

$$\min\{\widehat{I}(A : B)_{\mathcal{E}(P)(ABE|XYZ)}, \widehat{I}(A : B|E)_{\mathcal{E}(P)(ABE|XYZ)}\} \geq K_{DI}^{(\text{iid})}(P), \quad (\text{I7})$$

$$\widehat{I}(A : B \downarrow E)_{\mathcal{E}(P)(ABE|XYZ)} \geq \widehat{I}(A : B \downarrow \downarrow E)_{\mathcal{E}(P)(ABE|XYZ)} \geq K_{DI}^{(\text{iid})}(P). \quad (\text{I8})$$

The proof of the above Corollary is straightforward from Theorem 1. It is important to note that, the complete extension of a device, $P(AB|XY)$, has been denoted as $\mathcal{E}(P)(ABE|XYZ)$, where the extended systems are in full control of Eve.

The intrinsic information $\widehat{I}(A : B \downarrow E)$ and the reduced intrinsic information $\widehat{I}(A : B \downarrow \downarrow E)$ are functions without closed-form expression, and hence they cannot be computed straightforwardly. We present a technique for finding a nontrivial bound using the properties of one of them. First, we notice that for any fixed bipartite device and its complete extension, the following is true.

Observation 2 (Hierarchy between different mutual information functions).

$$\mathcal{N}_{\text{sq}}(P) = \widehat{I}(A : B \downarrow E)_{\mathcal{E}(P)(ABE|XYZ)} \leq \widehat{I}(A : B)_{\mathcal{E}(P)(ABE|XYZ)}, \quad (\text{I9})$$

$$\mathcal{N}_{\text{sq}}(P) = \widehat{I}(A : B \downarrow E)_{\mathcal{E}(P)(ABE|XYZ)} \leq \widehat{I}(A : B|E)_{\mathcal{E}(P)(ABE|XYZ)}. \quad (\text{I10})$$

The squashed nonlocality is upper bounded by the squashed conditional mutual information $\widehat{I}(A : B|E)_{\mathcal{E}(P)(ABE|XYZ)}$, and also with squashed mutual information $\widehat{I}(A : B)_{\mathcal{E}(P)(ABE|XYZ)}$, hence:

Observation 3. Nonsignaling squashed nonlocality is upper-bounded by the following expression.

$$\mathcal{N}_{\text{sq}}(P) \leq \min\{\widehat{I}(A : B)_{\mathcal{E}(P)(ABE|XYZ)}, \widehat{I}(A : B|E)_{\mathcal{E}(P)(ABE|XYZ)}\}. \quad (\text{I11})$$

Unfortunately, the squashed nonlocality lacks a closed-form expression for an arbitrary nonsignaling device. It involves optimization over general measurement and post-processing channels in the eavesdropper side. This makes it hard to compute for a generic nonsignaling device. Moreover, we obtained the squashed nonlocality to be a convex function over the mixture of devices, see Appendix G3, whereas the intrinsic information is not a convex function. This might be due to the fact that it was constructed in the same way as the nonsignaling squashed entanglement, and the latter is a convex function of quantum states [62]. In this section, we will show how convexity of squashed nonlocality can be used not only to calculate nontrivial upper bounds on $K_{DI}^{(\text{iid})}$, but also how it can be used to define new nonsignaling squashed secrecy quantifiers.

Observation 3, brings the idea of how to use the convexity property of squashed nonlocality. Since the squashed nonlocality is an upper bound on $K_{DI}^{(\text{iid})}$, hence, the r.h.s. of Eq. (I11) must also be an upper bound on secret-key rate as well. Together with the convexity property, it implies that a lower convex hull of $\widehat{I}(A : B|E)$ and $\widehat{I}(A : B)$ also bounds $K_{DI}^{(\text{iid})}$ from above.

Theorem 4. Within a family of functions $\{F_i\}$, which are convex with respect to mixtures of devices, and

$$F_i(P) \leq \widehat{I}(A : B)_{\mathcal{E}(P)(ABE|XYZ)}, \quad (\text{I12})$$

$$F_i(P) \leq \widehat{I}(A : B|E)_{\mathcal{E}(P)(ABE|XYZ)}, \quad (\text{I13})$$

there exists a function F that upper bounds any function in $\{F_i\}$ and for which the following relation holds

$$F(P) \geq K_{DI}^{(iid)}(P). \tag{I14}$$

Proof. Since $\widehat{I}(A : B \downarrow E) \in \{F_i\}$ because of Proposition 4 and $\widehat{I}(A : B \downarrow E)_{\mathcal{E}(P)} \geq K_{DI}^{(iid)}(P)$, then, for a function F which lies above the values of the squashed intrinsic mutual information, satisfies $F(P) \geq K_{DI}^{(iid)}(P)$.

Theorem 4, can be easily generalized by imposing different constraints than Eqs. (I12) and (I13), for example, by using other upper bounds on the squashed nonlocality and also an arbitrary number of them.

Remark 3. The lower convex hull of plots of an arbitrary number of functions, each being an upper bound on a convex function which upper bounds $K_{DI}^{(iid)}$, is an upper bound on the key rate itself.

This observation automatically yields a recipe on how to construct nontrivial upper bounds on $K_{DI}^{(iid)}$. We come up with the following Corollary, being a direct consequence of Theorem 4 and Remark 3.

Corollary 5. A nontrivial upper bound is given by the lower convex hull (LCH) of plots of nonsignaling squashed secrecy quantifiers.

$$\begin{aligned} \mathcal{N}_{sq}(P) &\leq F(P) \\ &:= \text{LCH}\{\widehat{I}(A : B)_{\mathcal{E}(P)(ABE|XYZ)}, \widehat{I}(A : B|E)_{\mathcal{E}(P)(ABE|XYZ)}\}. \end{aligned} \tag{I15}$$

Proof. We prove by contradiction. If there would be a function which at any point is greater than the lower convex hull of $\widehat{I}(A : B)$ and $\widehat{I}(A : B|E)$, either it would not be convex or it would be greater (at least at a single point) than at least one from the above nonsignaling squashed nonlocality quantifiers. Therefore it is not in the set $\{F_i\}$. ■

The upper bound on $K_{DI}^{(iid)}$ introduced in the above corollary can be computed much more easily than the nonsignaling squashed nonlocality. We will refer to the procedure of calculating upper bounds via this technique as *convexification*. Observation 2 and Proposition 9 provide a collection of functions which are upper bounds for \mathcal{N}_{sq} . Hence, there exists a convex (in the same sense) function, which is an upper bound on the squashed nonlocality, but at the same time, it is a lower bound on any function in this group, which is very clear from the proof of Theorem 4.

APPENDIX J: NUMERICAL UPPER BOUNDS ON SQUASHED NONLOCALITY

In this section, we will provide the upper bound on the \mathcal{N}_{sq} , for some exemplary family of devices, namely, two binary input and two binary output devices (2,2,2,2) and for a device which has ternary input for one subsystem and binary input for the other subsystem but all the outputs are binary (3,2,2,2). We have obtained that there exist some devices that are not MD-LOPC key distillable, although they are nonlocal. Describing the procedure of convexification, we focused on obtaining upper bounds by employing a lower convex hull of the upper bounds on \mathcal{N}_{sq} . The reason behind such an approach is to simplify our calculations. In this Section, we present a specific example of upper bounds on \mathcal{N}_{sq} , which we have obtained via this technique for some bipartite binary input output nonlocal devices. Let us recall here that \mathcal{N}_{sq} is defined as

$$\mathcal{N}_{sq}(P) = \max_{x,y} \min_z \inf_{\Theta_{E|E'}} I(A : B|E')_{(\mathcal{M}_{x,y}^F \otimes \mathcal{M}_z^G)_{\mathcal{E}(P)}}. \tag{J1}$$

The core strategy is based on the observation that the definition of nonsignaling squashed nonlocality involves two minimizations: one in the measurement process and another one in applying suitable post-processing channel, in part of the eavesdropper. We notice that one can obtain upper bounds also in the case in which used measurement and channels are not optimal, which follows from the property of infimum. Knowing this, we can run a three-step strategy to obtain an upper bound on $K_{DI}^{(iid)}$ for the desired set of devices.

- (1) Choose an (arbitrary, possibly continuous) set of devices, for which an upper bound is to be calculated.
- (2) Calculate the values of upper bounds on nonsignaling squashed nonlocality employing different devices, different measurement choices, and different post-processing channels. These can be obtained either via educated guess, some heuristic method or with computer aid, including a random search over the space.
- (3) Construct lower convex hull of all previously generated plots, and the result is the convex hull of the chosen set of points.

1. Upper bound for the nonsignalling device used by Hänggi, Renner and Wolf

We will now employ the above technique to bound the $K_{DI}^{(iid)}$. As we have argued, the notion of security employed by us is equivalent to that used by Hänggi, Renner, and Wolf [17]. The protocol proposed by them yields a positive key rate for devices exhibiting quantum correlations, we compare our upper bounds with the lower bound presented by them [17,63], in Fig. 8. The nonsignaling device we consider, as in Ref. [17], is given by

$$P_{HRW}(ab|xy) = \begin{array}{c|cc|cc|cc} & & x & & 0 & 1 & & 1 \\ & & \swarrow a & & & & & \\ y & b & & & & & & \\ \hline 0 & 0 & & \frac{1}{2} - \frac{\delta}{2} & \frac{\delta}{2} & \frac{3}{8} - \frac{\epsilon}{2} & \frac{1}{8} + \frac{\epsilon}{2} \\ & 1 & & \frac{\delta}{2} & \frac{1}{2} - \frac{\delta}{2} & \frac{1}{8} + \frac{\epsilon}{2} & \frac{3}{8} - \frac{\epsilon}{2} \\ \hline 1 & 0 & & \frac{3}{8} - \frac{\epsilon}{2} & \frac{1}{8} + \frac{\epsilon}{2} & \frac{1}{8} + \frac{\epsilon}{2} & \frac{3}{8} - \frac{\epsilon}{2} \\ & 1 & & \frac{1}{8} + \frac{\epsilon}{2} & \frac{3}{8} - \frac{\epsilon}{2} & \frac{3}{8} - \frac{\epsilon}{2} & \frac{1}{8} + \frac{\epsilon}{2} \end{array}. \tag{J2}$$

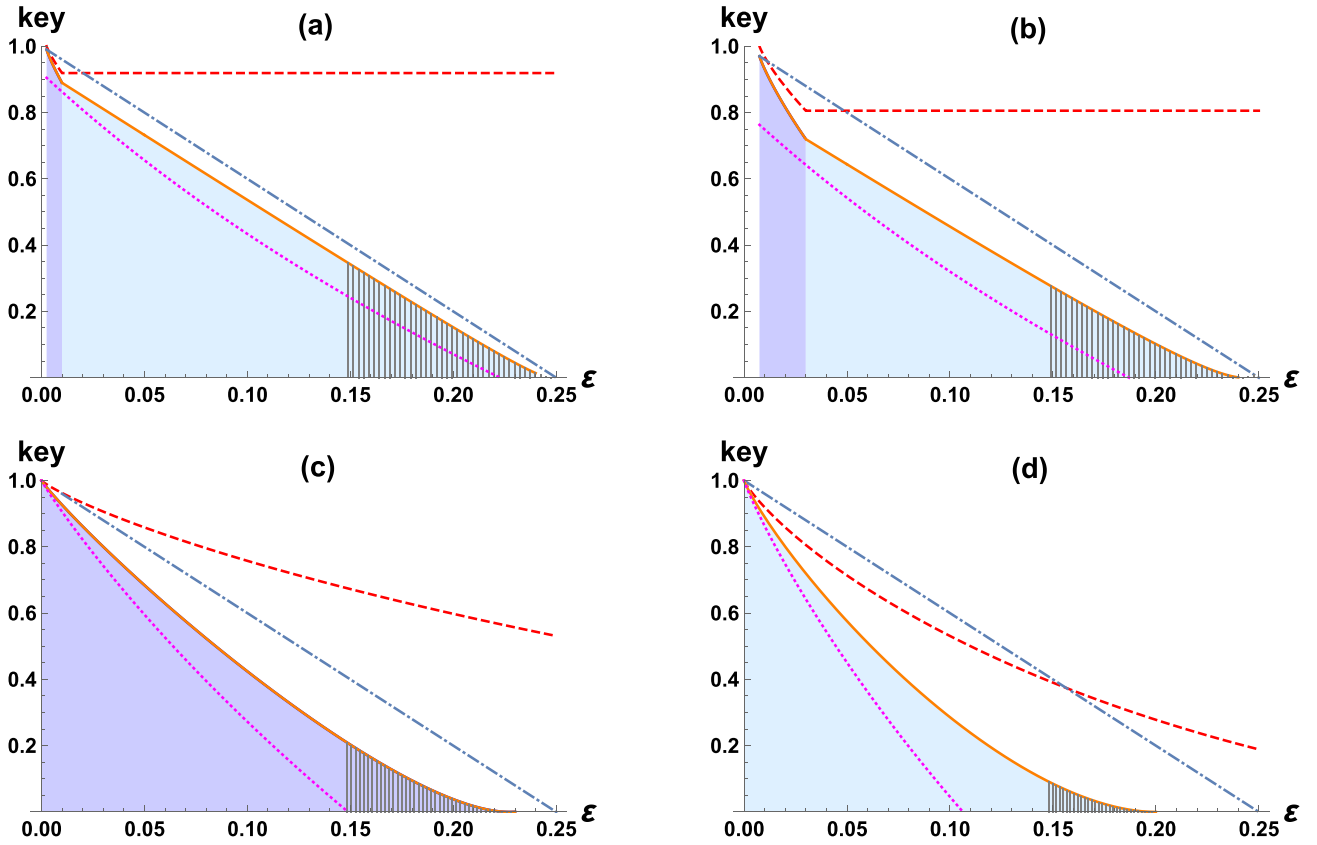


FIG. 8. Plot of several nonsignaling secrecy quantifiers $\widehat{M}(A : B|E)$, as an upper bound on secure key rate $K_{DI}^{(iid)}$, for the bipartite binary input output device P_{HRW} given in Eq. (J2) (also in Ref. [17]). The parameters chosen for drawing these figures are provided in Table II. The dashed red line corresponds to the nonsignaling squashed mutual information $\widehat{I}(A : B)_{P_{HRW}}$. The blue dashed-dotted straight line represents the nonlocality cost, as well as the nonsignaling squashed conditional mutual information $\widehat{I}(A : B|E)_{\mathcal{E}(P_{HRW})}$ over the complete extension $\mathcal{E}(P_{HRW})$ of the given device P . The solid orange line represents the upper bound on the nonsignaling squashed nonlocality \mathcal{N}_{sq} which is in fact the lower convex hull of the several other upper bounds on \mathcal{N}_{sq} . The magenta dotted line is the key rate $\mathcal{R}(P|_{P_{HRW}})$ of the protocol design by Hänggi, Renner, and Wolf [17].

It remains a valid nonsignaling probability distribution in the parameter range $0 \leq \delta \leq 1$ and $-\frac{1}{4} \leq \epsilon \leq \frac{3}{4}$. It exhibits nonlocal correlation for a very small range of parameters, quantified by the parameter ϵ , probability of not winning the CHSH game [60], which is

$$\epsilon = \Pr(a \oplus b \neq x \cdot y) = \frac{1}{4} \left(\frac{3}{4} + \delta + 3\epsilon \right). \quad (J3)$$

The device is nonlocal when the error $\epsilon \in [0, \frac{1}{4})$, and there are multiple choices of δ and ϵ to attain this. Without loss of generality, we choose $0 \leq \delta \leq 1$ and $-\frac{1}{4} \leq \epsilon \leq \frac{1}{12} - \frac{\delta}{3}$.

TABLE II. Table of the different values of the parameters δ and ϵ , for the sub-figures as given in Fig. 8. δ and ϵ are the parameters of bipartite nonsignaling device P_{HRW} given in Eq. (J2).

Figure	δ	ϵ
(a)	0.01	$\frac{1}{16}(3.04 + 12\epsilon)$
(b)	0.03	$\frac{1}{16}(3.12 + 12\epsilon)$
(c)	$\frac{2}{5}\epsilon$	$\frac{6}{5}\epsilon - \frac{1}{4}$
(d)	ϵ	$\epsilon - \frac{1}{4}$

The nonlocality fraction of these devices in the above range of parameters is $C(P) = \frac{1}{4} - \delta - 3\epsilon$.

The polytope of P_{HRW} , bipartite binary input-output devices, consists of 24 extremal devices [95], among which 16 are local or deterministic devices, and the remaining 8 are nonlocal. The local devices are given by

$$L_{\alpha\beta\gamma\sigma}(ab|xy) = \begin{cases} 1 & \text{if } a = \alpha x \oplus \beta, b = \gamma y \oplus \sigma, \\ 0 & \text{otherwise,} \end{cases} \quad (J4)$$

where $\alpha, \beta, \gamma, \sigma \in \{0, 1\}$. And the nonlocal devices are

$$B_{rst}(ab|xy) = \begin{cases} 1/2 & \text{if } a \oplus b = xy \oplus rx \oplus sy \oplus t, \\ 0 & \text{otherwise,} \end{cases} \quad (J5)$$

where $r, s, t \in \{0, 1\}$.

In Fig. 8, we plot several nonsignaling squashed secrecy quantifiers and monotones $\widehat{M}(A : B|E)$ for different choices of the parameters δ and ϵ , with respect to the ϵ , which forms the upper bound on $K_{DI}^{(iid)}$. Different plots correspond to different choices of the parameters ϵ and δ , as given in Table II. The last row of Table II, give rise to the isotropic device, i.e., $P_{iso} = (1 - \epsilon)PR + \epsilon\overline{PR}$, described in the main text.

In all the four figures, the red dashed line represents the squashed mutual information $\widehat{I}(A : B)_P$ between Alice and Bob. The optimal choices of the measurements by Alice and Bob in the squashing process varies with δ and ϵ . For Figs. 8(a) and 8(b), the optimum direct measurement choice is $(x = 0, y = 0)$ for $\delta < \epsilon$, and any one of the other three input choices for $\delta \geq \epsilon$. The measurement choice $(x = 0, y = 0)$ is optimal in the entire range of ϵ for Fig. 8(c), and all measurements choices give the same mutual information for the choice of δ and ϵ in Fig. 8(d).

The nonlocality cost $\mathcal{N}_C(\mathcal{P}_{\text{HRW}})$ is plotted with the dashed-dot blue line in all the figures.

Figure 8(d) clearly shows that our measure, nonsignaling squashed nonlocality \mathcal{N}_{sq} is *not* a faithful measure of nonlocality. The orange curve is the upper bound on \mathcal{N}_{sq} , and we have found that the bound reaches to 0 for $\epsilon = 0.2$ (it remains equal 0 for $\epsilon \in (0.2, 0.25]$ due to the convexity of the measure). It strongly suggests that there exists nonlocality which can not be turned into security. Indeed, for these devices, no protocol of distribution is known. Using wirings that is necessary for the key to be nonzero, imply that we enter to some extent the general scenario of K_{DI} for which there is a wide class of attacks [26]. Since our scenario is restricted, we can not postulate

nonequivalence between nonlocality and secrecy in NSDI paradigm.

a. Method to obtain the upper bound on \mathcal{N}_{sq}

The nonsignaling squashed nonlocality defined in Eq. (J1), is the optimal conditional mutual information $I(A : B|E')_{\mathcal{E}(\mathcal{P}_{\text{HRW}})}$, between Alice and Bob, when Eve holds the complete extension of the device \mathcal{P}_{HRW} . It involves a maximization over the measurement (input) choices of Alice and Bob. In our cryptographic protocol, we assume that Eve will perform an adaptive choice of measurements after learning Alice and Bob's measurements, followed by a post-processing channel. We also observed that an arbitrary adaptive measurement by Eve, direct or general, with any post-processing channel, provides an upper bound on \mathcal{N}_{sq} , which remains convex over ϵ , in the entire range of ϵ .

We calculate the CE [56] of \mathcal{P}_{HRW} numerically in the entire range of δ and ϵ , where the device is nonlocal. The most tighter upper bound we have obtained numerically, involve a direct measurement by Eve. This direct measurement is no doubt is a function of Alice and Bob's input choice, which is intended to reduce the correlation shared by them. This measurement on Eve's system creates the following minimal ensembles in part of Alice and Bob,

$$v = \left[\frac{1}{4} - \delta - 3\epsilon, \frac{1+4\epsilon}{8}, \frac{1+4\epsilon}{8}, \frac{1+4\epsilon}{8}, \frac{1+4\epsilon}{8}, \frac{1+4\epsilon}{8}, \frac{1+4\epsilon}{8}, \frac{1+4\epsilon}{8}, \frac{\delta}{2}, \frac{\delta}{2} \right], \quad (\text{J6})$$

$$\mathcal{E}_{z_0} = [\mathbf{B}_{000}, \mathbf{L}_{0000}, \mathbf{L}_{0010}, \mathbf{L}_{0101}, \mathbf{L}_{0111}, \mathbf{L}_{1000}, \mathbf{L}_{1101}, \mathbf{L}_{1011}, \mathbf{L}_{1110}]. \quad (\text{J7})$$

The same measurement leads us to the nonsignaling squashed conditional mutual information $\widehat{I}(A : B|E)_{\mathcal{E}(P)}$ for all input choices of Alice and Bob, which we have plotted by the dashed-dotted blue line in all the figures of Fig. 8. We have obtained that nonlocality cost of the shared device is $\mathcal{N}_C(\mathcal{P}_{\text{HRW}}) = \widehat{I}(A : B|E)_{\mathcal{E}(P)}$.

The classical discrete post-processing channel $\Theta_{E'|E}$, that we have obtained is different for different input choice of Alice and Bob. And they are

Device	B ₀₀₀	L ₀₀₀₀	L ₀₀₁₀	L ₀₁₀₁	L ₀₁₁₁	L ₁₀₀₀	L ₁₁₀₁	L ₁₀₁₁	L ₁₁₁₀
$e' \backslash e$	0	1	2	3	4	5	6	7	8
0	1	0	0	0	0	0	0	1	1
1	0	1	0	0	0	0	0	0	0
2	0	0	1	0	0	0	0	0	0
3	0	0	0	1	0	0	0	0	0
4	0	0	0	0	1	0	0	0	0
5	0	0	0	0	0	1	0	0	0
6	0	0	0	0	0	0	1	0	0

$$\Theta_{E'|E}^{0,0} = \quad (\text{J8})$$

Box	B ₀₀₀	L ₀₀₀₀	L ₀₀₁₀	L ₀₁₀₁	L ₀₁₁₁	L ₁₀₀₀	L ₁₁₀₁	L ₁₀₁₁	L ₁₁₁₀
$e' \backslash e$	0	1	2	3	4	5	6	7	8
0	1	0	0	0	0	1	1	0	0
1	0	1	0	0	0	0	0	0	0
2	0	0	1	0	0	0	0	0	0
3	0	0	0	1	0	0	0	0	0
4	0	0	0	0	1	0	0	0	0
5	0	0	0	0	0	0	0	1	0
6	0	0	0	0	0	0	0	0	1

$$\Theta_{E'|E}^{0,1} = \quad (\text{J9})$$

$$\Theta_{E'|E}^{1,0} = \begin{array}{c|cccccccccc} \text{Device} & \mathbf{B}_{000} & \mathbf{L}_{0000} & \mathbf{L}_{0010} & \mathbf{L}_{0101} & \mathbf{L}_{0111} & \mathbf{L}_{1000} & \mathbf{L}_{1101} & \mathbf{L}_{1011} & \mathbf{L}_{1110} \\ \hline e \backslash e & 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ \hline 0 & \mathbf{1} & 0 & \mathbf{1} & 0 & \mathbf{1} & 0 & 0 & 0 & 0 \\ \hline 1 & 0 & \mathbf{1} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \hline 2 & 0 & 0 & 0 & \mathbf{1} & 0 & 0 & 0 & 0 & 0 \\ \hline 3 & 0 & 0 & 0 & 0 & 0 & \mathbf{1} & 0 & 0 & 0 \\ \hline 4 & 0 & 0 & 0 & 0 & 0 & 0 & \mathbf{1} & 0 & 0 \\ \hline 5 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \mathbf{1} & 0 \\ \hline 6 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \mathbf{1} \end{array}, \tag{J10}$$

$$\Theta_{E'|E}^{1,1} = \begin{array}{c|cccccccccc} \text{Device} & \mathbf{B}_{000} & \mathbf{L}_{0000} & \mathbf{L}_{0010} & \mathbf{L}_{0101} & \mathbf{L}_{0111} & \mathbf{L}_{1000} & \mathbf{L}_{1101} & \mathbf{L}_{1011} & \mathbf{L}_{1110} \\ \hline e \backslash e & 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ \hline 0 & \mathbf{1} & \mathbf{1} & 0 & \mathbf{1} & 0 & 0 & 0 & 0 & 0 \\ \hline 1 & 0 & 0 & \mathbf{1} & 0 & 0 & 0 & 0 & 0 & 0 \\ \hline 2 & 0 & 0 & 0 & 0 & \mathbf{1} & 0 & 0 & 0 & 0 \\ \hline 3 & 0 & 0 & 0 & 0 & 0 & \mathbf{1} & 0 & 0 & 0 \\ \hline 4 & 0 & 0 & 0 & 0 & 0 & 0 & \mathbf{1} & 0 & 0 \\ \hline 5 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \mathbf{1} & 0 \\ \hline 6 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \mathbf{1} \end{array}. \tag{J11}$$

Hence, the upper bound on the key, according to our numerical findings is

$$K_{DI}^{(\text{iid})} \leq \mathcal{N}_{\text{sq}}(P) \leq \text{LCH}\{\hat{\mathbb{I}}(A : B|E)_{\mathcal{E}(P_{\text{RH}})(ABE|XYZ)}, \hat{\mathbb{I}}(A : B|E)_{Q(ABE|XYZ)}\}, \tag{J12}$$

where $Q(ABE|XYZ) = \Theta_{E'|E}^{X,Y}(\mathcal{E}(P_{\text{HRW}})(ABE'|XYZ))$, an arbitrary optimal extension, which is obtained from CE by applying the above post-processing channel.

The plot of the r.h.s. of the above inequality is given by the solid orange curve in Fig. 8. The color shade is used to separate the two regions, where the optimal measurement choices of the honest parties are coming from two different inputs. The light blue shade in Figs. 8(a) and 8(b) represents the choices of optimal inputs to be $(x = 0, y = 0)$, whereas the dark blue shade is for the other choices of input (all of them give rise to the same value). In Fig. 8(c), the optimal input by the honest parties is $(x = 0, y = 0)$, and in Fig. 8(d) all the other set of inputs are equally likely, and the color shed has been chosen to light blue.

We compare our upper bound with the key rate $\mathcal{R}(P|_{P_{\text{HRW}}})$, generated by Hänggi, Renner, and Wolf [17], which is the magenta dotted line in all the figures in Fig. 8. It lies below the solid orange line, as it represents the NSDI key rate for a particular protocol, and we provide the upper bound over all possible protocols.

Moreover, if we compare the bounds among the sub-figures of Fig. 8, we observe that for a fixed ε , the bound is almost decreasing if one goes from Fig. 8(a) to 8(d). This is because in Fig. 8(a), the choices of the parameters δ and ϵ are such that the probability of not winning the CHSH game is smaller for one choice of the input compared to the other input choices of the honest parties. In Fig. 8(d), all the distribution has the same error ε , depicting the lowest bound, i.e., all the inputs give rise to the same error, which leads to no specific choice of inputs.

The nonfaithfulness of our measure is visible from Fig. 8(d). We have found that the bound reaches to 0 for $\varepsilon = 0.2$ (it remains equal 0 for $\varepsilon \in (0.2, 0.25]$ due to the convexity of the measure). It strongly suggests that there exists nonlocality which can not be turned into security. Indeed, for these devices, no protocol of distribution is known. Using wirings that is necessary for the key to be nonzero, imply that we enter to some extent the general scenario of K_{DI} for which there is a wide class of attacks [26].

2. Upper bound for the nonsignaling device used by Acín, Massar, and Pironio

In this section, we will find an upper bound on the nonsignaling squashed nonlocality, for a device, which the honest parties Alice and Bob can obtain by performing quantum measurements on a shared bipartite quantum state, given in Ref. [16]. The shared quantum state is the Werner state $\rho_{AB} = p|\psi_+\rangle\langle\psi_+|_{AB} + \frac{1-p}{4}I_{AB}$, where $|\psi_+\rangle_{AB} = \frac{1}{\sqrt{2}}(|0\rangle_A|0\rangle_B + |1\rangle_A|1\rangle_A)$, and $p \in [0, 1]$. One of the honest parties, Alice consider three possible measurement choices $x \in \{0, 1, 2\}$, whereas Bob chooses only two possible measurements $y \in \{0, 1\}$. Among those set of measurements when both the measurement settings are $x = 0$ and $y = 0$, the measurement bases coincides and only that choice of measurement has been used for the key distribution run. The other two measurements $x \in \{1, 2\}$, for Alice and two measurements $y \in \{0, 1\}$, for Bob, are for the test of nonlocal correlation present in the system, i.e., for the violation of Bell inequality, of the shared state.

The shared probability distribution by both the parties, or the device obtained after the possible set of measurements is given by

$$P_{\text{AMP}}(ab|xy) = \begin{array}{c|cc|cc|cc|cc} & \begin{array}{c} x \\ b \backslash a \end{array} & \begin{array}{c} 0 \\ 0 \end{array} & \begin{array}{c} 1 \\ 1 \end{array} & \begin{array}{c} 0 \\ 0 \end{array} & \begin{array}{c} 1 \\ 1 \end{array} & \begin{array}{c} 0 \\ 0 \end{array} & \begin{array}{c} 1 \\ 1 \end{array} \\ \hline 0 & & \frac{1+p}{4} & \frac{1-p}{4} & \frac{2+\sqrt{2}p}{8} & \frac{2-\sqrt{2}p}{8} & \frac{2+\sqrt{2}p}{8} & \frac{2-\sqrt{2}p}{8} \\ & & \frac{1-p}{4} & \frac{1+p}{4} & \frac{2-\sqrt{2}p}{8} & \frac{2+\sqrt{2}p}{8} & \frac{2-\sqrt{2}p}{8} & \frac{2+\sqrt{2}p}{8} \\ \hline 1 & & \frac{1}{4} & \frac{1}{4} & \frac{2+\sqrt{2}p}{8} & \frac{2-\sqrt{2}p}{8} & \frac{2-\sqrt{2}p}{8} & \frac{2+\sqrt{2}p}{8} \\ & & \frac{1}{4} & \frac{1}{4} & \frac{2-\sqrt{2}p}{8} & \frac{2+\sqrt{2}p}{8} & \frac{2+\sqrt{2}p}{8} & \frac{2-\sqrt{2}p}{8} \end{array} \quad (J13)$$

In the entire range of p , the device is a valid probability distribution but it exhibit nonlocal correlation only for an small range of p . To compute the range of p , where let us quantify the probability of not winning the CHSH game [60], by the parameter ε , which is

$$\varepsilon(P_{\text{AMP}}) = \Pr(a \oplus b \neq (x - 1) \cdot y)_{P_{\text{AMP}}} = \frac{1}{4}(2 - \sqrt{2}p). \quad (J14)$$

Note that for P_{AMP} , Alice will use her inputs $x \in \{1, 2\}$ for the detection of nonlocality. Now the device is nonlocal when $0 \leq \varepsilon < \frac{1}{4}$, hence the device may be useful for secure key agreement protocol in presence of nonsignalling Eve in the range of $\frac{1}{\sqrt{2}} < p \leq 1$.

To make a rough estimation on the upper bound of \mathcal{N}_{sq} , of $P_{\text{AMP}}(ab|xy)$, we first focus on the squashed conditional mutual information $\widehat{I}(A : B|E)_{\mathcal{E}(P_{\text{AMP}})}$, where $\mathcal{E}(P_{\text{AMP}})$ is the complete extension of the given quantum device. In general, obtaining the complete extension of a given box, in this new (3,2,2,2) polytope is an extremely difficult task and hence, we have found here only one exemplary minimal ensemble which up to our numerical search is an optimal eavesdropping strategy, i.e., achieving the \min_z [see Eqs. (6) and Appendix A for the definition of $\widehat{I}(A : B|E)_{\mathcal{E}(P_{\text{AMP}})}$], for the chosen values of the measurement setup by the honest parties for key sharing $x = y = 0$. The minimal ensemble is

$$v = \left[\frac{p}{\sqrt{2}} - \frac{1}{2}, \frac{p}{\sqrt{2}} - \frac{1}{2}, \frac{2 - \sqrt{2}p}{8}, \frac{2 - \sqrt{2}p}{8}, \frac{2 - \sqrt{2}p}{8}, \frac{2 - \sqrt{2}p}{8}, \frac{2 - \sqrt{2}p}{8}, \frac{2 - \sqrt{2}p}{8}, \frac{1-p}{4}, \frac{1-p}{4}, \frac{(2 - \sqrt{2})p}{8}, \frac{(2 - \sqrt{2})p}{8} \right], \quad (J15)$$

$$\mathcal{E}_{z_0} = [B_0, B_1, L_0, L_1, L_2, L_3, L_4, L_5, L_6, L_7, L_8, L_9], \quad (J16)$$

where B_0 and B_1 are the two nonlocal extremal devices and L_0, \dots, L_9 are the local deterministic devices (extremal), in the polytope of the devices where P_{AMP} lies, and they are given by [96]

$$\begin{array}{c} B_0(ab|xy) = \begin{array}{c|cc|cc|cc} & \begin{array}{c} x \\ b \backslash a \end{array} & \begin{array}{c} 0 \\ 0 \end{array} & \begin{array}{c} 1 \\ 1 \end{array} & \begin{array}{c} 0 \\ 0 \end{array} & \begin{array}{c} 1 \\ 1 \end{array} \\ \hline 0 & & \frac{1}{2} & 0 & \frac{1}{2} & 0 & \frac{1}{2} & 0 \\ & & 0 & \frac{1}{2} & 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ \hline 1 & & \frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 & \frac{1}{2} \\ & & 0 & \frac{1}{2} & 0 & \frac{1}{2} & \frac{1}{2} & 0 \end{array}, \quad B_1(ab|xy) = \begin{array}{c|cc|cc|cc} & \begin{array}{c} x \\ b \backslash a \end{array} & \begin{array}{c} 0 \\ 0 \end{array} & \begin{array}{c} 1 \\ 1 \end{array} & \begin{array}{c} 0 \\ 0 \end{array} & \begin{array}{c} 1 \\ 1 \end{array} \\ \hline 0 & & \frac{1}{2} & 0 & \frac{1}{2} & 0 & \frac{1}{2} & 0 \\ & & 0 & \frac{1}{2} & 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ \hline 1 & & 0 & \frac{1}{2} & \frac{1}{2} & 0 & 0 & \frac{1}{2} \\ & & \frac{1}{2} & 0 & 0 & \frac{1}{2} & \frac{1}{2} & 0 \end{array}; \quad (J17) \\ \\ L_0 = \begin{array}{c|cc|cc|cc} & \begin{array}{c} x \\ b \backslash a \end{array} & \begin{array}{c} 0 \\ 0 \end{array} & \begin{array}{c} 1 \\ 1 \end{array} & \begin{array}{c} 0 \\ 0 \end{array} & \begin{array}{c} 1 \\ 1 \end{array} \\ \hline 0 & & 1 & 0 & 1 & 0 & 1 & 0 \\ & & 0 & 0 & 0 & 0 & 0 & 0 \\ \hline 1 & & 1 & 0 & 1 & 0 & 1 & 0 \\ & & 0 & 0 & 0 & 0 & 0 & 0 \end{array}, \quad L_1 = \begin{array}{c|cc|cc|cc} & \begin{array}{c} x \\ b \backslash a \end{array} & \begin{array}{c} 0 \\ 0 \end{array} & \begin{array}{c} 1 \\ 1 \end{array} & \begin{array}{c} 0 \\ 0 \end{array} & \begin{array}{c} 1 \\ 1 \end{array} \\ \hline 0 & & 1 & 0 & 1 & 0 & 0 & 1 \\ & & 0 & 0 & 0 & 0 & 0 & 0 \\ \hline 1 & & 0 & 1 & 0 & 1 & 0 & 1 \\ & & 1 & 0 & 0 & 0 & 0 & 0 \end{array} \end{array}$$

$$L_2 = \begin{array}{c|ccc|cc} & x & 0 & 1 & 2 & & \\ & y & b \backslash a & 0 & 1 & 0 & 1 \\ \hline 0 & & 0 & 0 & 0 & 0 & 0 \\ & & 1 & 0 & 1 & 0 & 1 \\ \hline 1 & & 0 & 0 & 0 & 0 & 0 \\ & & 1 & 0 & 1 & 0 & 1 \end{array}, \tag{J18}$$

$$L_3 = \begin{array}{c|ccc|cc} & x & 0 & 1 & 2 & & \\ & y & b \backslash a & 0 & 1 & 0 & 1 \\ \hline 0 & & 0 & 0 & 0 & 0 & 0 \\ & & 1 & 0 & 1 & 0 & 1 \\ \hline 1 & & 0 & 0 & 0 & 0 & 0 \\ & & 1 & 0 & 1 & 0 & 1 \end{array}, \quad L_4 = \begin{array}{c|ccc|cc} & x & 0 & 1 & 2 & & \\ & y & b \backslash a & 0 & 1 & 0 & 1 \\ \hline 0 & & 0 & 0 & 0 & 0 & 0 \\ & & 1 & 0 & 1 & 1 & 0 \\ \hline 1 & & 0 & 0 & 0 & 1 & 0 \\ & & 1 & 0 & 0 & 0 & 0 \end{array},$$

$$L_5 = \begin{array}{c|ccc|cc} & x & 0 & 1 & 2 & & \\ & y & b \backslash a & 0 & 1 & 0 & 1 \\ \hline 0 & & 0 & 1 & 0 & 0 & 1 \\ & & 1 & 0 & 0 & 0 & 0 \\ \hline 1 & & 0 & 0 & 0 & 0 & 0 \\ & & 1 & 1 & 0 & 0 & 1 \end{array}, \tag{J19}$$

$$L_6 = \begin{array}{c|ccc|cc} & x & 0 & 1 & 2 & & \\ & y & b \backslash a & 0 & 1 & 0 & 1 \\ \hline 0 & & 0 & 0 & 0 & 0 & 0 \\ & & 1 & 1 & 0 & 0 & 1 \\ \hline 1 & & 0 & 0 & 0 & 0 & 0 \\ & & 1 & 1 & 0 & 0 & 1 \end{array}, \quad L_7 = \begin{array}{c|ccc|cc} & x & 0 & 1 & 2 & & \\ & y & b \backslash a & 0 & 1 & 0 & 1 \\ \hline 0 & & 0 & 0 & 1 & 0 & 1 \\ & & 1 & 0 & 0 & 0 & 0 \\ \hline 1 & & 0 & 0 & 0 & 0 & 0 \\ & & 1 & 0 & 1 & 0 & 1 \end{array},$$

$$L_8 = \begin{array}{c|ccc|cc} & x & 0 & 1 & 2 & & \\ & y & b \backslash a & 0 & 1 & 0 & 1 \\ \hline 0 & & 0 & 1 & 0 & 1 & 0 \\ & & 1 & 0 & 0 & 0 & 0 \\ \hline 1 & & 0 & 0 & 0 & 0 & 0 \\ & & 1 & 1 & 0 & 1 & 0 \end{array}, \tag{J20}$$

$$L_9(ab|xy) = \begin{array}{c|ccc|cc} & x & 0 & 1 & 2 & & \\ & y & b \backslash a & 0 & 1 & 0 & 1 \\ \hline 0 & & 0 & 0 & 0 & 0 & 0 \\ & & 1 & 0 & 1 & 1 & 0 \\ \hline 1 & & 0 & 0 & 0 & 0 & 0 \\ & & 1 & 0 & 1 & 1 & 0 \end{array}. \tag{J21}$$

For the given decomposition of the device $P_{AMP}(ab|xy)$, the squashed conditional mutual information reduces to $\widehat{I}(A : B|E)_{\mathcal{E}(P_{AMP})} = \sqrt{2}p - 1$, which is equal to the nonlocality cost of the shared device i.e., $\mathcal{N}_C(P_{AMP})$. It reaches to $\sqrt{2} - 1$, for $p = 1$, i.e., when the Bell state is shared.

To obtain the upper bound on $\mathcal{N}_{sq}(P_{AMP})$, we will again apply some post-processing channel $\Theta_{E|E'}$, on the output of Eve E , and apply the procedure of getting the lower convex hull, by the relation

$$\mathcal{N}_{sq}(P_{AMP}) \leq \text{LCH}\{\widehat{I}(A : B|E)_{\mathcal{E}(P_{AMP})(ABE|XYZ)}, \widehat{I}(A : B|E)_{Q_{AMP}(ABE|XYZ)}\}, \tag{J22}$$

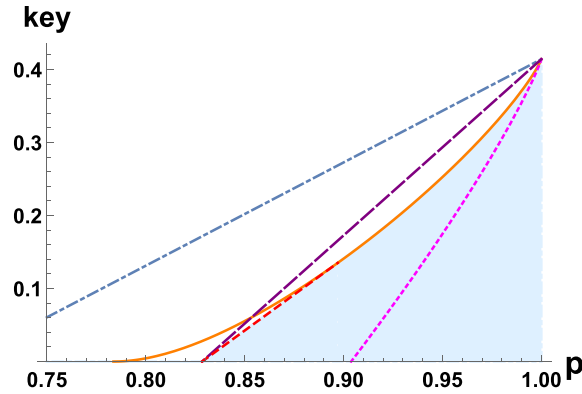


FIG. 9. Plot of nontrivial upper bound on the nonsignaling squashed nonlocality \mathcal{N}_{sq} , of $P_{\text{AMP}}(ab|xy)$ given in Eq. (J13), by the blue shaded region under the orange solid line and a red dashed line. The red dashed line is the (segment of) lower convex hull of the orange solid curve and the purple long-dashed straight line. The solid orange line is obtained by the lower convex hull of several upper bounds of \mathcal{N}_{sq} , with the help of Eq. (339). Blue dashed-dotted line is the squashed conditional mutual information $\hat{I}(A : B|E)_{\mathcal{E}(P_{\text{AMP}})}$. The magenta dotted line is the lower bound on the key rate, whereas the purple long-dashed line is the upper bound on intrinsic information of the eavesdropping strategy used in Ref. [16]. We observe that the convexification technique resulting in the convex-hull bound allows to obtain tighter upper bound on \mathcal{N}_{sq} , and therefore the tightest known upper bound on the secret-key rate in the nonsignaling scenario.

where $Q_{\text{AMP}}(ABE|XYZ) = \Theta_{E|E'}(\mathcal{E}(P_{\text{AMP}})(ABE'|XYZ))$ is an arbitrary extension of P_{AMP} , upon applying the post-processing channel $\Theta_{E|E'}$, given by

Device	B ₀	B ₁	L ₀	L ₁	L ₂	L ₃	L ₄	L ₅	L ₆	L ₇	L ₈	L ₉
$e \setminus e'$	0	1	2	3	4	5	6	7	8	9	10	11
0	1	1	0	0	0	0	1	1	0	0	0	0
1	0	0	1	0	0	0	0	0	0	0	0	0
2	0	0	0	1	0	0	0	0	0	0	0	0
3	0	0	0	0	1	0	0	0	0	0	0	0
4	0	0	0	0	0	1	0	0	0	0	0	0
5	0	0	0	0	0	0	0	0	1	0	0	0
6	0	0	0	0	0	0	0	0	0	1	0	0
7	0	0	0	0	0	0	0	0	0	0	1	0
8	0	0	0	0	0	0	0	0	0	0	0	1

(J23)

Note that here we need only one post-processing channel, because in the squashing procedure unlike Appendix J1, Eve’s know which outcomes of Alice and Bob are used for the key generation run.

The upper bound on $\mathcal{N}_{\text{sq}}(P_{\text{AMP}})$, i.e., the right-hand side of (339), has been plotted in Fig. 9, by the orange line, which vanishes for $p \approx 0.783$, and from the procedure of lower

convex hull we will consider it 0, for all $p < 0.783$. The magenta dotted line is the lower bound on the key rate of Ref. [16], whereas the violate dashed line is the upper bound on the intrinsic information $I(A : B \downarrow E)$, of Ref. [16], for a particular eavesdropping strategy. We have found that our bound on \mathcal{N}_{sq} is better than the bound on $I(A : B \downarrow E)$, by [16], for $p > 0.853$.

[1] I. Csiszár and J. Körner, Broadcast channels with confidential messages, *IEEE* **24**, 339 (1978).
 [2] U. M. Maurer, Secret key agreement by public discussion from common information, *IEEE Trans. Inf. Theory* **39**, 773 (1993).
 [3] N. Gisin, G. Ribordy, W. Tittel, and H. Zbinden, Quantum cryptography, *Rev. Mod. Phys.* **74**, 145 (2002).
 [4] C. H. Bennett and G. Brassard, Quantum cryptography: Public key distribution and coin tossing, in *Proceedings of the IEEE International Conference on Computers, Systems and Signal Processing*, Bangalore, India, December, 1984 (IEEE Computer Society Press, New York, 1984), pp. 175–179.

[5] A. K. Ekert, Quantum Cryptography Based on Bell’s Theorem, *Phys. Rev. Lett.* **67**, 661 (1991).
 [6] C. H. Bennett, Quantum Cryptography using Any Two Nonorthogonal States, *Phys. Rev. Lett.* **68**, 3121 (1992).
 [7] A. Acín, J. Bae, E. Bagan, M. Baig, L. Masanes, and R. Muñoz-Tapia, Secrecy properties of quantum channels, *Phys. Rev. A* **73**, 012327 (2006).
 [8] N. Brunner, D. Cavalcanti, S. Pironio, V. Scarani, and S. Wehner, Bell nonlocality, *Rev. Mod. Phys.* **86**, 839(E) (2014).
 [9] D. Mayers and A. Yao, Self testing quantum apparatus, *Quantum Inf. Comp.* **4**, 273 (2004).

- [10] A. Acín, N. Brunner, N. Gisin, S. Massar, S. Pironio, and V. Scarani, Device-Independent Security of Quantum Cryptography against Collective Attacks, *Phys. Rev. Lett.* **98**, 230501 (2007).
- [11] L. Masanes, S. Pironio, and A. Acín, Secure device-independent quantum key distribution with causally independent measurement devices, *Nat. Commun.* **2**, 238 (2011).
- [12] R. Arnon-Friedman, R. Renner, and T. Vidick, Simple and tight device-independent security proofs, *SIAM J. Comput.* **48**, 181 (2019).
- [13] J. Barrett, L. Hardy, and A. Kent, No Signaling and Quantum Key Distribution, *Phys. Rev. Lett.* **95**, 010503 (2005).
- [14] A. Acín, N. Gisin, and L. Masanes, From Bell's Theorem to Secure Quantum Key Distribution, *Phys. Rev. Lett.* **97**, 120405 (2006).
- [15] L. Masanes, R. Renner, M. Christandl, A. Winter, and J. Barrett, Full security of quantum key distribution from no-signaling constraints, *IEEE Trans. Inf. Theory* **60**, 4973 (2014).
- [16] A. Acín, S. Massar, and S. Pironio, Efficient quantum key distribution secure against no-signaling eavesdroppers, *New J. Phys.* **8**, 126 (2006).
- [17] E. Hänggi, R. Renner, and S. Wolf, *Efficient Quantum Key Distribution Based Solely on Bell's Theorem* (EUROCRYPT, 2010), pp. 216–234.
- [18] I. Devetak and A. Winter, Relating Quantum Privacy and Quantum Coherence: An Operational Approach, *Phys. Rev. Lett.* **93**, 080501 (2004).
- [19] K. Horodecki, M. Horodecki, P. Horodecki, and J. Oppenheim, General paradigm for distilling classical key from quantum states, *IEEE Trans. Inf. Theory* **55**, 1898 (2009).
- [20] K. Horodecki, M. Horodecki, P. Horodecki, and J. Oppenheim, Secure Key from Bound Entanglement, *Phys. Rev. Lett.* **94**, 160502 (2005).
- [21] J. Barrett, R. Colbeck, and A. Kent, Unconditionally secure device-independent quantum key distribution with only two devices, *Phys. Rev. A* **86**, 062326 (2012).
- [22] V. Scarani, N. Gisin, N. Brunner, L. Masanes, S. Pino, and A. Acín, Secrecy extraction from no-signaling correlations, *Phys. Rev. A* **74**, 042339 (2006).
- [23] L. Masanes, Universally Composable Privacy Amplification from Causality Constraints, *Phys. Rev. Lett.* **102**, 140501 (2009).
- [24] E. Hänggi, R. Renner, and S. Wolf, The impossibility of non-signaling privacy amplification, *Theor. Comp. Sci.* **486**, 27 (2013).
- [25] R. Arnon-Friedman and A. Ta-Shma, Limits of privacy amplification against nonsignaling memory attacks, *Phys. Rev. A* **86**, 062333 (2012).
- [26] B. Salwey and S. Wolf, Stronger attacks on causality-based key agreement, in *2016 IEEE International Symposium on Information Theory (ISIT)* (IEEE, New York, 2016), pp. 2254–2258.
- [27] R. Arnon-Friedman, E. Hänggi, and A. Ta-Shma, Towards the impossibility of nonsignaling privacy amplification from time-like ordering constraints, [arXiv:1205.3736](https://arxiv.org/abs/1205.3736).
- [28] P. Rastall, Locality, Bell's theorem, and quantum mechanics, *Found. Phys.* **15**, 963 (1985).
- [29] L. A. Khalfin and B. S. Tsirelson, *Quantum and Quasi-Classical Analogs of Bell Inequalities*, In *Symposium on the Foundations of Modern Physics*, edited by P. Lahti and P. Mittelstaedt (World Scientific Publishing, 1985), pp. 441–460.
- [30] S. Popescu and D. Rohrlich, Quantum nonlocality as an axiom, *Found. Phys.* **24**, 379 (1994).
- [31] E. Kaur, M. M. Wilde, and A. Winter, Fundamental limits on key rates in device-independent quantum key distribution, *New J. Phys.* **22**, 023039, (2020).
- [32] B. Kraus, N. Gisin, and R. Renner, Lower and Upper Bounds on the Secret-Key Rate for Quantum Key Distribution Protocols Using One-Way Classical Communication, *Phys. Rev. Lett.* **95**, 080501 (2005).
- [33] I. Devetak and A. Winter, Distillation of secret key and entanglement from quantum states, *Proc. R. Soc. London A* **461**, 207 (2005).
- [34] M. Christandl, A. Ekert, M. Horodecki, P. Horodecki, J. Oppenheim, and R. Renner, Unifying classical and quantum key distillation, *Proceedings of the 4th Theory of Cryptography Conference*, *Lect. Notes Comput. Sci.* **4392**, 456 (2007).
- [35] R. Augusiak and P. Horodecki, Multipartite secret-key distillation and bound entanglement, *Phys. Rev. A* **80**, 042307 (2009).
- [36] M. Christandl, The quantum analog to intrinsic information, Diploma Thesis, Institute for Theoretical Computer Science, ETH Zurich, 2002.
- [37] D. Yang, K. Horodecki, M. Horodecki, P. Horodecki, J. Oppenheim, and W. Song, Squashed entanglement for multipartite states and entanglement measures based on the mixed convex roof, *IEEE Trans. Inf. Theory* **55**, 3375 (2009).
- [38] M. M. Wilde, Squashed entanglement and approximate private states, *Quantum Inf. Proc.* **15**, 4563 (2016).
- [39] M. Takeoka, S. Guha, and M. M. Wilde, The squashed entanglement of a quantum channel, *IEEE Trans. Inf. Theory* **60**, 4987 (2014).
- [40] M. Takeoka, S. Guha, and M. M. Wilde, Fundamental rate-loss tradeoff for optical quantum key distribution, *Nat. Commun.* **5**, 5235 (2014).
- [41] S. Pirandola, R. Laurenza, C. Ottaviani, and L. Banchi, Fundamental limits of repeaterless quantum communications, *Nat. Commun.* **8**, 15043 (2017).
- [42] M. Takeoka, K. P. Seshadreesan, and M. M. Wilde, Unconstrained distillation capacities of a pure-loss bosonic broadcast channel, in *2016 IEEE International Symposium on Information Theory (ISIT) Theory (ISIT)* (IEEE, New York, 2016).
- [43] M. M. Wilde, M. Tomamichel, and M. Berta, Converse bounds for private communication over quantum channels, *IEEE Trans. Inf. Theory* **63**, 1792 (2017).
- [44] R. Laurenza and S. Pirandola, General bounds for sender-receiver capacities in multipoint quantum communications, *Phys. Rev. A* **96**, 032318 (2017).
- [45] S. Pirandola, S. L. Braunstein, R. Laurenza, C. Ottaviani, T. P. W. Cope, G. Spedalieri, and L. Banchi, Theory of channel simulation and bounds for private communication, *Quantum Sci. Technol.* **3**, 035009 (2018).
- [46] S. Pirandola, U. L. Andersen, L. Banchi, M. Berta, D. Bunandar, R. Colbeck, D. Englund, T. Gehring, C. Lupo, C. Ottaviani, J. L. Pereira, M. Razavi, J. Shamsul Shaari, M. Tomamichel, V. C. Usenko, G. Vallone, P. Villoresi, and P. Wallden, Advances in quantum cryptography, *Adv. Opt. Photonics* **12**, 1012 (2020).
- [47] E. Kaur, *Limitations on protecting information against quantum adversaries*, Ph.D. thesis, LSU Doctoral Dissertations, Louisiana State University, 2020.

- [48] M. Christandl, R. Ferrara, and K. Horodecki, Upper Bounds on Device-Independent Quantum Key Distribution, *Phys. Rev. Lett.* **126**, 160501 (2021).
- [49] R. Arnon-Friedman and F. Leditzky, Upper bounds on device-independent quantum key distribution rates and a revised peres conjecture, *IEEE Trans. Inf. Theory* **67**, 6606 (2021).
- [50] M. Farkas, M. Balanzó-Juandó, K. Łukanowski, J. Kołodyński, and A. Acín, Bell Nonlocality Is Not Sufficient for the Security of Standard Device-Independent Quantum Key Distribution Protocols, *Phys. Rev. Lett.* **127**, 050503 (2021).
- [51] E. Kaur, K. Horodecki, and S. Das, Upper bounds on device-independent quantum key distribution rates in static and dynamic scenarios, [arXiv:2107.06411](https://arxiv.org/abs/2107.06411).
- [52] K. Horodecki, M. Winczewski, and S. Das, Fundamental limitations on the device-independent quantum conference key agreement, *Phys. Rev. A* **105**, 022604 (2022).
- [53] K. Audenaert, J. Eisert, E. Jané, M. B. Plenio, S. Virmani, and B. De Moor, Asymptotic Relative Entropy of Entanglement, *Phys. Rev. Lett.* **87**, 217902 (2001).
- [54] W. van Dam, P. Grunwald, and R. Gill, The statistical strength of nonlocality proofs, *IEEE Trans. Inf. Theory* **51**, 2812 (2005).
- [55] A. Grudka, K. Horodecki, M. Horodecki, P. Horodecki, R. Horodecki, P. Joshi, W. Kłobus, and A. Wójcik, Quantifying Contextuality, *Phys. Rev. Lett.* **112**, 120401 (2014).
- [56] M. Winczewski, T. Das, K. Horodecki, P. Horodecki, M. Piani, Ł. Pankowski, and R. Ramanathan, Complete extension: the nonsignaling analog of quantum purification, [arXiv:1810.02222](https://arxiv.org/abs/1810.02222).
- [57] R. Schwonnek, K. T. Goh, I. W. Primaatmaja, E. Y.-Z. Tan, R. Wolf, V. Scarani, and C. C.-W. Lim, Device-independent quantum key distribution with random key basis, *Nat. Commun.* **12**, 2880 (2021).
- [58] U. Maurer and S. Wolf, Unconditionally secure key agreement and the intrinsic conditional information, *IEEE Trans. Inf. Theory* **45**, 499 (1999).
- [59] U. Maurer and S. Wolf, The intrinsic conditional mutual information and perfect secrecy, in *Proc. 1997 IEEE Symposium on Information Theory (Abstracts)* (IEEE, New York, 1997), p. 88.
- [60] J. F. Clauser, M. A. Horne, A. Shimony, and R. A. Holt, Proposed Experiment to Test Local Hidden-Variable Theories, *Phys. Rev. Lett.* **23**, 880 (1969).
- [61] R. R. Tucci, Entanglement of distillation and conditional mutual information, [arXiv:quant-ph/0202144](https://arxiv.org/abs/quant-ph/0202144).
- [62] M. Christandl and A. Winter, “squashed entanglement”: An additive entanglement measure, *J. Math. Phys.* **45**, 829 (2004).
- [63] E. Hänggi, *Device-independent quantum key distribution*, Ph.D. thesis, ETH Zurich, 2010.
- [64] E. Hänggi and R. Renner, Device-independent quantum key distribution with commuting measurements, [arXiv:1009.1833](https://arxiv.org/abs/1009.1833).
- [65] G. Chiribella, G. M. D’Ariano, and P. Perinotti, Probabilistic theories with purification, *Phys. Rev. A* **81**, 062348 (2010).
- [66] G. Chiribella, G. M. D’Ariano, and P. Perinotti, Informational derivation of quantum theory, *Phys. Rev. A* **84**, 012311 (2011).
- [67] M. Christandl and B. Toner, Finite de Finetti theorem for conditional probability distributions describing physical theories, *J. Math. Phys.* **50**, 042104 (2009).
- [68] J. Barrett, R. Colbeck, and A. Kent, Memory Attacks on Device-Independent Quantum Cryptography, *Phys. Rev. Lett.* **110**, 010503 (2013).
- [69] R. Arnon-Friedman, *Reductions to IID in Device-independent quantum information processing*, Ph.D. thesis, ETH Zurich, 2018.
- [70] R. Renner and S. Wolf, New bounds in secret-key agreement: The gap between formation and secrecy extraction, in *Advances in Cryptology — EUROCRYPT 2003*, edited by Eli Biham (Springer Berlin, Heidelberg, 2003), pp. 562–577.
- [71] R. Renner and R. König, Universally composable privacy amplification against quantum adversaries, in *TCC* (Springer, 2005), Vol. 3378.
- [72] B. S. Cirel’son, Quantum generalizations of Bell’s inequality, *Lett. Math. Phys.* **4**, 93 (1980).
- [73] J. Barrett, Information processing in generalized probabilistic theories, *Phys. Rev. A* **75**, 032304 (2007).
- [74] Ll. Masanes, A. Acin, and N. Gisin, General properties of nonsignaling theories, *Phys. Rev. A* **73**, 012112 (2006).
- [75] U. Maurer and S. Wolf, Information-theoretic key agreement: from weak to strong secrecy for free, *Lect. Notes Comput. Sci.* **1807**, 351 (2000).
- [76] L. Lami, C. Palazuelos, and A. Winter, Ultimate data hiding in quantum mechanics and beyond, *Commun. Math. Phys.* **361**, 661 (2018).
- [77] C. Portmann and R. Renner, Cryptographic security of quantum key distribution, [arXiv:1409.3525](https://arxiv.org/abs/1409.3525).
- [78] N. J. Beaudry, *Assumptions in quantum cryptography*, Ph.D. thesis, ETH Zurich, 2015.
- [79] M. Ben-Or and D. Mayers, General security definition and compossibility for quantum & classical protocols, 2004, [arXiv:quant-ph/0409062](https://arxiv.org/abs/quant-ph/0409062).
- [80] M. Ben-Or, M. Horodecki, D. W. Leung, D. Mayers, and J. Oppenheim, The universal composable security of quantum key distribution, in *Theory of Cryptography*, edited by J. Kilian, Lecture Notes in Computer Science, Vol. 3378 (Springer, Berlin, Heidelberg, 2005), pp. 386–406.
- [81] R. Canetti, Universally composable security: A new paradigm for cryptographic protocols, in *Proceedings 2001 IEEE International Conference on Cluster Computing* (IEEE, New York, 2001), pp. 136–145.
- [82] K. T. Goh, J. Kaniewski, E. Wolfe, T. Vértesi, X. Wu, Y. Cai, Y.-C. Liang, and V. Scarani, Geometry of the set of quantum correlations, *Phys. Rev. A* **97**, 022104 (2018).
- [83] C. A. Miller and Y. Shi, Universal security for randomness expansion from the spot-checking protocol, *SIAM J. Comput.* **46**, 1304 (2017).
- [84] M. Winczewski, T. Das, and K. Horodecki, Analogue of the Uhlmann’s theorem in post-quantum theory and asymptotic continuity of squashed non-locality (unpublished).
- [85] K. Horodecki, A. Grudka, P. Joshi, W. Kłobus, and J. Łodyga, Axiomatic approach to contextuality and nonlocality, *Phys. Rev. A* **92**, 032104 (2015).
- [86] A. Karimi, Z. Huang, and M. R. Paul, Erratum: Exploring spiral defect chaos in generalized swift-hohenberg models with mean flow [Phys. Rev. E 84, 046215 (2011)], *Phys. Rev. E* **99**, 039901 (2019).
- [87] L. Hardy, Quantum theory from five reasonable axioms, 2001, [arXiv:quant-ph/0101012](https://arxiv.org/abs/quant-ph/0101012).
- [88] R. Renner, J. Skripsky, and S. Wolf, A new measure for conditional mutual information and its properties, in *IEEE International Symposium on Information Theory, 2003. Proceedings* (IEEE, New York, NY, 2003), pp. 259–259.

- [89] J. Tuziemski and K. Horodecki, On the non-locality of tripartite non-singaling boxes emerging from wirings, *Quantum Info. Comput.* **15**, 1081 (2015).
- [90] M. E. Shirokov, Tight continuity bounds for the quantum conditional mutual information, for the Holevo quantity and for capacities of quantum channels, *J. Math. Phys.* **58**, 102202 (2017).
- [91] R. Alicki and M. Fannes, Continuity of quantum conditional information, *J. Phys. A: Math. Gen.* **37**, L55 (2004).
- [92] M. Horodecki R. Horodecki, P. Horodecki, and K. Horodecki, Quantum entanglement, *Rev. Mod. Phys.* **81**, 865 (2009).
- [93] A. C. Elitzur, S. Popescu, and D. Rohrlich, Quantum nonlocality for each pair in an ensemble, *Phys. Lett. A* **162**, 25 (1992).
- [94] N. Brunner, D. Cavalcanti, A. Salles, and P. Skrzypczyk, Bound Nonlocality and Activation, *Phys. Rev. Lett.* **106**, 020402 (2011).
- [95] J. Barrett, N. Linden, S. Massar, S. Pironio, S. Popescu, and D. Roberts, Non-local correlations as an information theoretic resource, *Phys. Rev. A* **71**, 022101 (2005).
- [96] N. S. Jones and L. Masanes, Interconversion of nonlocal correlations, *Phys. Rev. A* **72**, 052312 (2005).

Complete extension: the non-signalling analog of quantum purification

Marek Winczewski^{1,2}, Tamoghna Das^{2,3}, John H. Selby², Karol Horodecki^{2,3}, Paweł Horodecki^{2,4}, Łukasz Pankowski⁵, Marco Piani⁶, and Ravishankar Ramanathan⁷

¹Institute of Theoretical Physics and Astrophysics and National Quantum Information Centre in Gdańsk, University of Gdańsk, 80–952 Gdańsk, Poland

²International Centre for Theory of Quantum Technologies, University of Gdańsk, Wita Stwosza 63, 80-308 Gdańsk, Poland

³Institute of Informatics and National Quantum Information Centre in Gdańsk, Faculty of Mathematics, Physics and Informatics, University of Gdańsk, 80–952 Gdańsk, Poland

⁴Faculty of Applied Physics and Mathematics and the National Quantum Information Centre, Gdańsk University of Technology, 80–233 Gdańsk, Poland

⁵VOICELAB.AI, Al. Grunwaldzka 135A; 80-264 Gdańsk, Poland,

⁶*SUPA* and Department of Physics, University of Strathclyde, Glasgow, G4 0NG, UK

⁷Department of Computer Science, The University of Hong Kong, Pokfulam Road, Hong Kong

October 20, 2022

Deriving quantum mechanics from information-theoretic postulates is a recent research direction taken, in part, with the view of finding a beyond-quantum theory; once the postulates are clear, we can consider modifications to them. A key postulate is the purification postulate, which we propose to replace by a more generally applicable postulate that we call the complete extension postulate (CEP), i.e., the existence of an extension of a physical system from which one can generate any other extension. This new concept leads to a plethora of open questions and research directions in the study of general theories satisfying the CEP (which may include a theory that hyper-decoheres to quantum theory). For example, we show that the CEP implies the impossibility of bit-commitment. This is exemplified by a case study of the theory of non-signalling behaviors which we show satisfies the CEP. We moreover show that in certain cases the complete extension will not be pure, highlighting the key divergence from the purification postulate.

Contents

1	Introduction	2
1.1	Generalised probabilistic theories	4
2	The purification postulate	6
2.1	No-go theorem for purification in discrete theories	6
3	The complete extension postulate	8
3.1	Impossibility of bit-commitment	10
3.2	No no-go for hyperdecoherence	12
4	Case study - theory of non-signalling behaviours	13
4.1	The complete extension in the theory of non-signalling behaviours	16
4.2	On the dimensionality of the complete extension	21
5	Conclusions	22

A	Proofs	30
A.1	Proof of Proposition 13	30
A.2	Proof of Theorem 16	31
A.3	Proof of Corollary 20	33
A.4	Proof of Theorem 23	33
B	Explicit examples in the theory of no-signaling behaviours	35
B.1	Complete extensions of binary input-output behaviours	35
B.1.1	Example: the non-signaling complete extension that is not a vertex	38
B.1.2	Example: NSCE of P_A gives ACCESS to any PME of P_A	40
B.1.3	Example: NSCE of P_A gives ACCESS to any mixed ensemble of P_A	40
B.1.4	NSCE can generate any extension	42
B.1.5	Quantifying non-locality introduced in NSCE	42
B.1.6	Complete extension of the conjugate box	43
B.2	Dimensionality of the No-signaling complete extension	44
B.3	Complete extensions of three-cycle contextual behaviour	46
B.4	Complete extension of the behaviours lying on the isotropic line	50
B.4.1	Minimal ensembles for isotropic behaviours	50
B.4.2	Dimension of the extending party for the behaviours lying on the isotropic line	51
B.4.3	Minimal ensembles for non-local isotropic behaviours	53

1 Introduction

Quantum mechanics appears to be one of the best validated physical theories and, at the same time, lacking a single agreed-on interpretation [1]. It can be expressed via several mathematical axioms [2–4] and leads to numerous phenomena, which have been confirmed experimentally, and now form the basis for many of our everyday technologies. However, it cannot explain some of the visible (or detectable) parts of the universe – the key example being the realm of the theory of gravity. For this reason, an important scientific effort is underway to find a beyond-quantum theory which would explain phenomena beyond quantum theory [5–10]. However, it is not clear which axiom of quantum mechanics should be replaced or modified in order to keep the theory as expressive as quantum mechanics whilst explaining phenomena, such as gravity, which seem to be beyond quantum theory.

It is known that the standard axioms of quantum mechanics are not independent from one another [11–13], i.e., they can be used to derive one another. It is therefore difficult to modify the standard postulates of quantum theory as they cannot be independently modified. In order to examine modifications of quantum theory, it is therefore useful to first re-express quantum mechanics in the form of a new set of postulates which can be independently modified. Recently, there has been a great deal of interest in such *reformulations*, particularly, via information theoretic postulates within the framework of so-called Generalised Probabilistic Theories (GPTs). This idea led to quite a number of reformulations of quantum theory [8, 14–27] in which sets of axioms are proposed which single out quantum mechanics from the space of all GPTs.

One of the key insights in the reconstruction programme came from [28] which introduced the purification postulate. This was an essential postulate, in that it was the sole postulate that distinguished quantum from classical theory. Since its introduction the purification postulate has been used extensively in the literature to prove many results, for example, pertaining to computation [29–31], cryptography [32, 33], thermodynamics [34, 35], and interference [36]. Essentially the purification postulate is a generalisation of the notion of purification within quantum theory to arbitrary GPTs:

Quantum Purifications *In quantum theory, for any state ρ_A there exists a system B and a bipartite pure state $|\psi\rangle_{AB}$, which satisfies $\text{tr}_B(|\psi\rangle\langle\psi|) = \rho_A$. The bipartite state $|\psi\rangle_{AB}$ is known as a purification of ρ_A , and these are essentially unique, in that there is always an isometry mapping between any pair of purifications.*

We formally introduce purifications in Sec. 2 and the formalism of GPTs necessary to talk about it in Sec. 1.1. At this point, however, let us observe that there are various issues with the purification postulate that may lead one to believe that it will have to be modified in order to find a more fundamental theory of nature.

1. There is a no-go theorem proven in Ref. [37] which shows that if one believes that a more fundamental theory should be causal, and should reduce to quantum theory by a decoherence-like mechanism, then one must give up on the purification postulate.
2. We prove a theorem in Sec. 2.1 which shows that the purification postulate must fail in any discrete theory, whilst in Refs. [38–40] ideas coming from quantum gravity are used to argue that on a fundamental level the quantum state space will become discrete.
3. The purification postulate also fails to hold in super-selected quantum systems [24, 27, 41], and so if we believe in any fundamental superselection rules then the postulate must be modified in some way.
4. There are also questions about whether the purification postulate is suitable in a theory which permits indefinite causal structure [42–45], as it is shown in Ref. [46] that a purification-like postulate fails to hold for all quantum process matrices. That a theory should permit indefinite causal structure can again be motivated by ideas coming from quantum gravity.

For all of these reasons one can then ask if the purification postulate can be weakened or replaced by another postulate, resulting in a physical theory with more explanatory power. In this manuscript, we propose such a replacement for the purification postulate called the Complete Extension Postulate (CEP) and study its consequences in various contexts.

In order to understand the complete extension postulate we begin by observing that, within quantum mechanics, purifications are examples of complete extensions [47, Exercise 2.8.2].

Quantum Complete Extensions *A complete extension for a state ρ_A in quantum theory, is an extension σ_{AB}^* of ρ_A , that is it satisfies $\text{tr}_B(\sigma_{AB}^*) = \rho_A$, where, moreover, any other extension σ can be reached via local operations on B , and any ensemble for ρ_A can be reached via a measurement on B .*

It can be straightforwardly seen that, within quantum theory, all quantum purifications are quantum complete extensions but not vice versa. It is precisely these complete extensions which we generalise to GPTs in order to define the complete extension postulate in Sec. 3. As this is a weaker notion within quantum theory it is therefore plausible that the associated postulate will apply more broadly and avoid some of the aforementioned issues with the purification postulate. Indeed, we show that this is the case for the first three issues raised above:

1. In Sec. 3.2 we show the proof of the no-go theorem for hyperdecoherence provided in [37] no longer holds if we use the complete extension postulate.
2. We show that the complete extension postulate can hold in discrete theories, in particular, in Sec. 4 we demonstrate this for the widely studied theory of non-signalling behaviours [5, 48, 49] colloquially known as *boxworld*.
3. In Sec. 3 we show that the complete extension postulate holds in super-selected quantum systems and within classical systems.

This may therefore lead one to wonder whether perhaps the complete extension postulate is too weak and does not have any meaningful consequences. We show that this is not the case. In particular, we show that there are theories which have been studied in the literature, such as those found in [50], which fail to satisfy the CEP. This means that it is not a trivial postulate and does offer explanatory power. Indeed, there are two key reasons that a given GPT may fail to satisfy the CEP. On the one hand, it may not have sufficient bipartite states, in which case it would not admit of an extension which enables access of all ensembles, on the other hand, it may not have sufficient dynamics as would be

the case for highly asymmetric state spaces, in which case even if there is an extension which enables access it may not be generating. Moreover, in Sec. 3.1 We show that any theory which satisfies the CEP cannot allow for the cryptographic primitive of bit-commitment, and, moreover, following the techniques of [33], we find a non-trivial lower bound on the success probability for any protocol. As an added benefit, this unifies the proof of impossibility of bit-commitment for the quantum and classical cases which previously were treated separately.

We also study the complete extensions within the theory of non-signalling behaviours in some depth in Sec. 4 and in App. B. In particular, we show how a complete extension can be constructed within this theory, and that the dimensions of these complete extensions are always finite. In this way, we show that CEP is not an empty postulate in general theories, as the complete extensions can actually be constructed. Furthermore, we characterise the minimal ensembles of the non-local isotropic behaviours, which was already proved useful in the context of the non-signalling device independent secure key agreement [51].

The introduction of a new postulate opens up many doors for future exploration. We touch on just a few of them in this paper and point out many others as future research directions as we go. In particular, we collect together many of these in Sec. 5.

1.1 Generalised probabilistic theories

The purification postulate and complete extension postulate that we introduce are expressed in the language of generalised probabilistic theories (GPTs), hence, in order to understand these we first provide a brief introduction to this formalism. For a more in-depth introduction see, for example, Refs. [52–54]

The formalism of GPTs [8, 48] is a formalism for describing the operational predictions of essentially arbitrary conceivable theories of physics. In the case of quantum theory, this operational description is isomorphic to the standard quantum informatic approach to quantum theory (that is, of completely positive trace preserving maps, density matrices, and positive operator value measures).

The primitive building blocks of any GPT, \mathcal{G} are the *systems*, denoted $A, B, \dots \in \text{Syst}[\mathcal{G}]$, these come equipped with an associative binary composition rule $\otimes : \text{Syst}[\mathcal{G}] \times \text{Syst}[\mathcal{G}] \rightarrow \text{Syst}[\mathcal{G}]$ which allow us to build composite systems out of simpler system. In quantum theory these systems can be labeled by natural numbers, i.e., by their dimension, and then these labels compose simply by multiplication, as the dimension of a tensor product of two Hilbert spaces is the product of the individual dimensions.

Each system A corresponds to a finite dimensional real vector space V_A . The fact that it is finite can either be viewed as simply a choice for technical convenience, or can be motivated operationally as the idea that in practice we can only ever perform a finite number of measurements to do tomography on a system (even if in principle there are an infinite number of degrees of freedom). In quantum theory these real vector spaces are the spaces of Hermitian operators on some finite dimensional Hilbert space, whilst in classical theory these real vector spaces are the function spaces \mathbb{R}^Λ for some finite sample space Λ . It is important to note at this point that $V_{A \otimes B}$ is not necessarily equal to $V_A \otimes V_B$. These are equal only under the assumption of tomographic locality [8] – that every multipartite state can be characterised by means of local measurements on its parts. Note that both quantum and classical theory are tomographically local theories.

Within these vector spaces V_A there is a specified convex set $\Omega_A \subset V_A$ which describes the *state space* for the system. This convex set must be compact and closed, and, moreover, must have an affine dimension one less than the linear dimension of the vector space such that the affine span of Ω_A does not intersect the origin. We can also define the convex state cone K_A which includes also subnormalised and supernormalised states as $K_A := \{rs | r \in \mathbb{R}^+, s \in \Omega_A\}$. In quantum theory this state space corresponds to the space of density matrices and the state cone to positive semidefinite operators.

Within the dual vector space V_A^* – that is, the vector space of linear functionals on V_A – there is also a specified convex set $\mathcal{E}_A \subset V_A^*$ which describes the *effect space* for the system. This convex set must also be compact and closed, however, it will be full dimensional and contain the origin. The effects $e \in \mathcal{E}_A$ assign probabilities to measurement outcomes when the measurement is performed on an arbitrary state in Ω_A , they therefore must satisfy $e(\Omega_A) \subseteq [0, 1]$. The effect space \mathcal{E}_A must also

contain the unique unit effect $u_A \in \mathcal{E}_A$ which is defined as $u_A(\Omega_A) = \{1\}$. The uniqueness of this effect is built into the setup of our framework, however, there are modifications to the framework in which this is instead an additional principle which captures the notion of causality [15]. In quantum theory this effect space corresponds to $\text{tr}(\sigma \cdot)$ where σ is a POVM element and the unit corresponds to $\sigma = \mathbb{1}$.

Like the systems themselves, these state and effect spaces have an associative composition rule, $\Omega_A \otimes \Omega_B := \Omega_{A \otimes B}$, $\mathcal{E}_A \otimes \mathcal{E}_B := \mathcal{E}_{A \otimes B}$ and $u_A \otimes u_B := u_{A \otimes B}$. Note that in general \otimes is not related to the tensor product of vector spaces, the symbol is used in analogy to the role of the tensor product in composing quantum systems. This composition must be bilinear, and satisfy $e \circ f(s \otimes t) = e(s)f(t)$ for all $e \in \mathcal{E}_A$, $f \in \mathcal{E}_B$, $s \in \Omega_A$ and $t \in \Omega_B$. These conditions, in particular, ensure that the unit effect u_B gives a way to uniquely define a kind of “partial trace” and hence marginal states – this in turn ensures that the theory does not permit signalling without sending a physical system [15, 55, 56].

Next we consider transformations between systems. For a pair of systems A and B there is a space of transformations \mathcal{T}_A^B from system A to system B . These again form a closed compact and convex set. In this case, there are two relevant associative composition rules. The first, parallel composition, $\mathcal{T}_A^B \otimes \mathcal{T}_C^D := \mathcal{T}_{A \otimes C}^{B \otimes D}$ and the second, sequential composition, $\mathcal{T}_B^C \circ \mathcal{T}_A^B := \mathcal{T}_A^C$. These are both bilinear and must satisfy the condition that $(T_1 \otimes T_2) \circ (T_3 \otimes T_4) = (T_1 \circ T_3) \otimes (T_2 \circ T_4)$. Note that states (and effects) can be viewed as particular kinds of transformations, namely those that have a trivial input (resp. output) system. Denoting this trivial system by \star and noting that \star is a unit of system composition ($A \otimes \star = A = \star \otimes A$) we can therefore write that $\Omega_A = \mathcal{T}_\star^A$ and $\mathcal{E}_A = \mathcal{T}_A^\star$. Note that for every system, A , there is an identity transformation $\mathbb{1}_A$ which for every $T : A \rightarrow B$ must satisfy $\mathbb{1}_B \circ T = T = T \circ \mathbb{1}_A$. One can then observe that what we are defining here is nothing but a monoidal category where the objects are the GPT systems and the morphisms are the GPT transformations.

It is convenient to work with the convention whereby every GPT contains classical systems, denoted Δ_I , in which measurement outcomes can be encoded, and which can act as control variables in descriptions of experiments. See, for example, [24, 57]. In particular, a demolition measurement of system A with outcome set I can then be viewed as a transformation in the set $\mathcal{T}_A^{\Delta_I \mathbb{1}}$. In order to see the connection between this and the description of measurements in terms of effects, we first set up some notation for describing classical systems.

Consider a classical system Δ_I which is a system corresponding to, for example, some outcome degree of freedom on some measurement device where I labels the set of possible outcomes. The vector space V_{Δ_I} will correspond to the real vector space of real functions from $I \rightarrow \mathbb{R}$, denoted \mathbb{R}^I . The state space Ω_{Δ_I} is the space of probability distributions over I , that is, real functions $p : I \rightarrow \mathbb{R}$ such that $p(i) \in [0, 1]$ and $\sum_i p(i) = 1$ for all $i \in I$. Note that geometrically this is a simplex with vertices labelled by the elements of I . These vertices correspond to delta function distributions which we denote as δ_i . The effect space \mathcal{E}_{Δ_I} lives in the dual vector space, however, here we make use of the Riesz representation theorem via the inner product $\sum_{i \in I} f(i)g(i)$ in order to view the effects as living in \mathbb{R}^I . Using this representation then the effects correspond to functions $e : I \rightarrow \mathbb{R}$ such that $e(i) \in [0, 1]$. Geometrically these form a hypercube which contains the simplex of states. The vertices of the simplex, when interpreted as effects via the Riesz representation will be denoted as ϵ_i such that $\epsilon_i(\delta_j) = \sum_{k \in I} \delta_i(k)\delta_j(k) = \delta_{ij}$. These systems compose via $\Delta_I \otimes \Delta_J := \Delta_{I \times J}$. Transformations between classical systems correspond to stochastic linear maps.

Now, note an important property of classical theory, that identity transformations $\mathbb{1}_{\Delta_I}$ can be decomposed as $\sum_{i \in I} \delta_i \circ \epsilon_i$. Hence, any measurement $M : A \rightarrow \Delta_I$ satisfies $M = \mathbb{1}_{\Delta_I} \circ M = \sum_{i \in I} \delta_i \circ \epsilon_i \circ M$ then, noting that $e_i := \epsilon_i \circ M \in \mathcal{T}_A^\star = \mathcal{E}_A$ we can rewrite this as $\sum_{i \in I} \delta_i \circ e_i$. One can then, recalling that $\epsilon_i(\delta_j) = \delta_{ij}$, see that it is possible to construct an isomorphism between measurements as transformations to a classical system and measurements as a collection of effects.

¹Nondemolition measurements would live in the space $\mathcal{T}_A^{A \otimes \Delta_I}$

2 The purification postulate

The purification postulate [15] has been widely studied within the literature on generalised probabilistic theories, from the study of thermodynamics [34, 35], cryptography [32, 33], and computation [29–31] through to higher-order interference [36] and reconstructions of quantum theory [28]. One of the reasons for the wide use of this principle is the fact that it is mathematically very powerful. For instance, one immediate corollary is that the state space is transitive [15] and there is a continuum of pure states (see Theorem 7 below).

Despite the utility of the purification postulate, as we discussed in Sec. 1, there are several criticisms which can be made of it. This motivates the search for alternative postulates such as the complete extension postulate that we propose here. Nonetheless, the purification postulate served as the inspiration behind the complete extension so it is useful to introduce it in some detail here.

The purification postulate takes the key ideas behind quantum purification and recasts them in the language of GPTs such that it can be taken as a postulate. We will slowly build up the relevant GPT concepts here before presenting the postulate itself.

Definition 1 (Pure states). *A state s of a system A is said to be pure if and only if it is extremal in Ω_A .*

In the case of quantum theory this coincides with the usual notion of purity, namely that of rank-1 density matrices. More generally, pure states are thought of as states of maximal knowledge whilst mixed states can be thought of as describing classical uncertainty of knowledge about which pure state has been prepared.

We can then define the notion of a purification of a given state:

Definition 2 (Purifications of a state). *A purification of a state ω_A of system A is a non-signalling extension to system B . That is, a state ε_{AB} of the composite system $A \otimes B$, satisfying*

$$(a) \quad [\mathbb{1}_A \otimes u_B](\varepsilon_{AB}) = \omega_A.$$

$$(b) \quad \varepsilon_{AB} \text{ is a pure state.}$$

If s is pure by itself, we assume that it is one of its own purifications, by taking B system to be trivial. We denote the set of purifications of a given state ω_A as $\mathbf{Purif}[\omega_A]$.

According to Definition 2, a state may possess more than one state that purifies it. That is, $\mathbf{Purif}[s]$ typically is not a singleton set. In quantum mechanics, for example $|\psi_{AB}\rangle$ and $|\psi_{AB}\rangle \otimes |\phi_{B'}\rangle$ purify the same state, i.e., $\text{tr}_B(|\psi_{AB}\rangle\langle\psi_{AB}|)$. In quantum mechanics, however, purifications are unique up to isometries on the purifying system. This motivates the following:

Definition 3 (Essential uniqueness of purifications). *The purifications of a given state s are said to be essentially unique, if and only if the elements in $\mathbf{Purif}[\omega_A]$ with the same purifying system can be related by reversible transformations on the purifying systems.*

With these ideas in place we are then in the position that we can succinctly define the purification postulate:

Definition 4 (Purification Postulate). *A GPT \mathcal{G} satisfies the Purification Postulate if and only if for all systems A and all states $\omega_A \in \Omega_A$, there exists purifications (Def. 2) which are essentially unique (Def. 3).*

2.1 No-go theorem for purification in discrete theories

In this subsection, we will prove that there cannot exist purifications of arbitrary states in any discrete theory. Later on, we will see that this implies that in such theories the complete extensions that we have introduced will generically not be pure. This fact has been proven in case of classical theory and for the theory of non-signalling behaviours in [15, 28]. We show below a different, direct argument, which holds in general for all non-signalling, convex discrete theories. In fact, the result developed

here is valid in any convex, non-signalling theory, provided that in each dimension number of different systems' types (state spaces with distinct shapes) is countable. We show that there is no single finite dimensional discrete theory [58–60], which vertices can purify all the states from a theory of a smaller dimension. We recall below the formal definition of a discrete theory:

Definition 5 (Discrete theory). *A GPT \mathcal{G} is said to be discrete, if and only if, for each system $T \in \text{Syst}[\mathcal{G}]$ there is a discrete number of pure states, that is, where each state space, Ω_T , is a polytope.*

The demand for the theory to be convex presumes that the state space lives inside a vector space, the dimension of which is also the dimension of a theory. These assumptions together with the non-signalling condition allow for the well-definiteness of the partial trace as a linear map $\text{Tr}_B(\cdot) : \Omega_A \otimes \Omega_B \equiv \Omega_{A \otimes B} \mapsto \Omega_A$ ², which for normalized states (in a sense of probabilistic outcomes) ρ_A, σ_B satisfies $\text{Tr}_B(\rho_A \otimes \sigma_B) = (\mathbb{1}_A \otimes u_B)(\rho_A \otimes \sigma_B) = \rho_A$, where u_B is the unit effect on the system B.

Lemma 6. *In any convex (nontrivial) theory, the cardinality of the set of states is at least of power of the continuum \mathfrak{c} .*

Proof. Let us take $\omega_A \neq \omega'_A$ in Ω_A . Since the theory is convex, any state of the form $p\omega_A + (1-p)\omega'_A$, for $p \in [0, 1]$ is still a state in Ω_A . The set of states $\{p\omega_A + (1-p)\omega'_A\}_{p=0}^{p=1}$, forms an interval in Ω_A . Since there is a bijection between any interval and the set of real numbers that has cardinality of continuum, the state space has at least cardinality of continuum. \square

We are ready to state the main theorem of this section now.

Theorem 7. *In any discrete theory (Def. 5) there are no (non-trivial) systems that have purifications for all states. Hence, theories with purifications (Def. 4) cannot be discrete.*

Proof. As the theory is a discrete theory, any system, T within the theory contains only finite number of v , vertices, and so the vertices can only purify a finite number of states f (as each pure state purifies a single mixed state). On the other hand according to Lemma 6, for all (nontrivial) systems, T' , in the theory, the cardinality of the set of states is at least \mathfrak{c} . So according to set-theoretic fact [63] $f < \aleph_0 < \mathfrak{c}$, in the system T there are not enough vertices to purify all states from the system T' . As it is true for all systems T in the theory, there exist no discrete theory, that can purify all states. \square

Note that the proof of theorem 7 does not strictly rely on convexity, but just that the cardinality of the state space is greater than \aleph_0 . One could therefore replace the assumption of convexity simply with the assumption that the number of mixed states in a considered discrete theory is greater than \aleph_0 .

As a consequence of Theorem 7, we have the following Proposition. Note that this result does not rely on the essential uniqueness of purifications, only their existence.

Proposition 8. *For any theory (with a countable number of system types) where purifications exist for all mixed states, there must be at least one system with a continuum of pure states.*

Proof. In any non-trivial theory there are state spaces which correspond to convex sets other than the singleton set. Such convex sets necessarily contain a continuum of mixed states, i.e., a continuum of non-extremal states. If purifications exist then each of these mixed states must purify to a distinct pure state in the theory (as marginalisation is described by a mapping from the bipartite states to the local states).

Now suppose, for the sake of contradiction, that there is a countable number of system types, that the state spaces of these systems live in finite dimensional vector spaces (as per the definition of a GPT), and, that these state spaces all have a discrete number of vertices (i.e., pure states). Then, the theory as a whole has a countable number of pure states, and hence there are not enough pure states to purify the continuum of mixed states – in contradiction to our initial assumption.

If we want to hold on to the countable number of system types, the finite dimensional vector spaces, and the existence of purifications, it must therefore be the case that there exists a system with a continuum of pure states. \square

²The theory of composition in GPTs was developed in various places, for example, Refs. [15, 61, 62].

3 The complete extension postulate

The idea of the complete extension is that it should be taken as a competitor to the purification postulate and is, *prima facie*, a mathematically weaker postulate. Indeed, one cannot derive either transitivity of the state space or the necessity for a continuum of pure states from this postulate whereas these are consequences of the purification postulate. Consequently, the complete extension postulate has the potential to be much more broadly applicable in scope and, hence, any results which can be derived using this postulate will hold in a broader class of theories than those derived using the purification postulate.

In this section we will formalise the definition of the complete extension postulate in the language of GPTs introduced in Sec. 2. We will work towards this by first defining ensembles, extensions and the principles of ACCESS and GENERATION within this language.

To begin, suppose we have some state, s , a vector in the convex set of states Ω_A of a GPT system A . An *ensemble* for s , is some set of pairs $\{(p_i, s_i)\}_{i \in I}$ such that $s_i \in \Omega_A$ and $\{p_i\}$ define a probability distribution ($p_i \in \mathbb{R}$, $p_i \geq 0$, and $\sum_i p_i = 1$), satisfying:

$$s = \sum_{i \in I} p_i s_i \quad (1)$$

in other words, it is some convex decomposition of s within the state space. We will denote the set of all possible ensembles for a state s as $\mathbf{Ens}[s]$.

The set of states which appear in some ensemble in $\mathbf{Ens}[s]$ is, will be, a face of the state space Ω_A , we denote this as $\mathbf{Face}[s]$. A state is *pure* if and only if $\mathbf{Face}[s] = \{s\}$ and is said to be an *interior* state if and only if $\mathbf{Face}[s] = \Omega_A$.

We can then define a *pure ensemble* for s as an ensemble $\{(p_i, s_i)\}_{i \in I} \in \mathbf{Ens}[s]$ in which all of the s_i are pure, that is, $\mathbf{Face}[s_i] = \{s_i\} \forall i \in I$. Let us denote the set of such pure ensembles as $\mathbf{Ens}_P[s] \subseteq \mathbf{Ens}[s]$, this is necessarily non-empty as we can always decompose a point inside a compact and closed convex set in terms of the vertices of the convex set.

An *extension* of a state $s \in \Omega_A$ of system A is a state $\sigma \in \Omega_{A \otimes E}$ of a bipartite system $A \otimes E$ for which we have:

$$s = [\mathbf{1}_A \otimes u_E](\sigma) \quad (2)$$

we will denote the set of extensions, σ , of a state s as $\mathbf{Ext}[s]$, note that, in general, there will be extensions with many different extending systems E . A *purification*, should it exist, is simply some $\sigma \in \mathbf{Ext}[s]$ which is pure, that is, $\mathbf{Purif}[s] \subseteq \mathbf{Ext}[s]$.

Now we can define what we mean for an extension to satisfy GENERATION. An extension $\sigma^* \in \mathbf{Ext}[s]$ with extending system E^* is said to be *generating* if and only if for every $\sigma \in \mathbf{Ext}[s]$ with arbitrary extending system E there is some transformation $T_\sigma : E^* \rightarrow E$ such that:

$$\mathbf{1}_A \otimes T_\sigma(\sigma^*) = \sigma \quad (3)$$

Next we show an equivalence between the notion of an ensemble and a particular class of extensions. This particular class of extensions are extensions in which the extending system, E , is taken to be a *classical* system, Δ_d for some $d \in \mathbb{N}$. We will denote this class of extensions for a state s as $\mathbf{Ext}_{class}[s]$ and show that there is an isomorphism between $\mathbf{Ext}_{class}[s]$ and $\mathbf{Ens}[s]$.

Proposition 9. *The space of extensions with classical extending system is isomorphic to the space of ensembles, i.e., $\mathbf{Ext}_{class}[s] \cong \mathbf{Ens}[s]$*

Proof. Starting with an ensemble:

$$\{(p_i, s_i)\}_{i \in I} \quad (4)$$

we can construct an extension:

$$\sum_{i \in I} p_i s_i \otimes \delta_i \in \Omega_A \otimes \Delta_I \quad (5)$$

where $\delta_i \in \text{Vert}[\Delta_I]$.

Whilst, on the other hand, if we start with an extension

$$\sigma \in \Omega_A \otimes \Delta_J \quad (6)$$

we can construct an ensemble:

$$\left\{ \left(u_A \otimes \epsilon_j(\sigma) \ , \ \frac{1}{u_A \otimes \epsilon_j(\sigma)} \mathbb{1}_A \otimes \epsilon_j(\sigma) \right) \right\}_{j \in J} \quad (7)$$

It is then simple to verify that these two constructions are the inverse of one another and hence establishes an isomorphism between the two sets. \square

Using this isomorphism we can easily express the property of ACCESS in a way which makes it manifestly a special case of GENERATION. That is, an extension $\sigma^* \in \mathbf{Ext}[s]$ with extending system E^* is an *extension with access* if and only if, for any ensemble $\sigma \in \mathbf{Ext}_{class}[s]$ with extending system Δ_I we can find a measurement $M_\sigma : E^* \rightarrow \Delta_I$ such that:

$$\mathbb{1}_A \otimes M_\sigma(\sigma^*) = \sigma \quad (8)$$

Then it is clear that:

Proposition 10. *If an extension σ has the property of GENERATION then it necessarily has the property of ACCESS.*

Proof. This immediately follows from inspection of the two definitions, and noting that $\mathbf{Ens}[s] \cong \mathbf{Ext}_{class}[s] \subseteq \mathbf{Ext}[s]$. \square

The converse however will not always be true, that is, from ACCESS we cannot generally derive GENERATION.

Proposition 11. *If an extension σ has the property of ACCESS then it does not necessarily have the property of GENERATION.*

Proof. We prove this by providing an example of a GPT which has extensions which allow for ACCESS but do not allow for generation.

Specifically, consider the GPT which has the same states and measurements as quantum theory. Then, the static part of the purification postulate, namely the existence of purifications, is satisfied. Hence, we have that any purification of a state allows us to access the ensembles of that state – that is they have the property ACCESS.

On the other hand, let us assume that the set of possible dynamics is restricted relative to quantum theory, such that transformations between non-trivial systems are necessarily noisy, that is, mixed with some non-zero parameter ϵ of the totally depolarising channel. Then we fail the dynamical, “essential uniqueness” part of the purification postulate and so these extensions will not have the property of GENERATION. \square

We are now in a position such that we can formally define the complete extension postulate within the language of GPTs.

Definition 12 (Complete Extension Postulate). *A GPT \mathcal{G} satisfies the Complete Extension Postulate (CEP) iff: for all systems A and all states $s \in \Omega_A$, there exists an extension $\sigma^* \in \mathbf{Ext}[s]$ which is generating, that is, which has the GENERATION property.*

With this definition in place, we can now compare this to the Purification Postulate [15], Def. 4. The essential uniqueness property in the purification postulate is intuitively very closely related to the GENERATION property of the complete extension postulate. However, note an important subtlety – essential uniqueness is a property for extensions on the same extending system, whilst GENERATION is a property for extensions on arbitrary extending systems. In order to derive GENERATION from essential uniqueness we must therefore invoke one additional, relatively innocuous looking, assumption about the GPT. Specifically, that the parallel composition of pure states is necessarily pure.

Proposition 13. *For the set of GPTs in which the product of pure states is pure, the purification postulate implies the complete extension postulate.*

Proof. See App. A.1. □

An immediate corollary of this is that quantum theory satisfies the CEP, as it is a GPT satisfying the purification postulate and in which the product of pure states is pure. More specifically, in the case of quantum theory, a purification $|\psi\rangle$ of a state ρ is a complete extension. In particular, it allows for GENERATION of arbitrary extensions and ACCESS of arbitrary ensembles.

It can also easily be shown that classical probability theory satisfies the CEP. That is, the complete extension of a probability distribution with finite support over some set X has a complete extension given by simply copying the variable X . See, for example, Ex. 3.4 in [64] for details. Specifically, a general extension of a probability distribution $p(x)$ is simply a distribution $q(x, y)$ which marginalises to $p(x)$, that is, such that $\sum_y q(x, y) = p(x)$. It is then straightforward to verify that the bipartite probability distribution $p_{ce}(x, x') := p(x)\delta(x, x')$ (obtained by copying x) is a complete extension, as any other extension can be generated by applying a stochastic map to x' . This is in stark contrast to the purification postulate, which does not hold in the case of classical theory. In this sense, CEP is a far more natural postulate to consider; indeed, it can be thought of as salvaging the most intuitive part of the purification postulate, which is common to both quantum and classical theory (even though the full purification postulate does not hold in classical theory). We also show in Sec. 4, that the complete extension postulate holds in the GPT of non-signalling boxes (colloquially known as Boxworld) which is also known to fail the purification postulate.

Another proposal to modify the purification postulate was put forward in Ref. [24] which provides a time-symmetric version of the purification postulate. It can be shown that this time-symmetric purification postulate, together with another postulate from Ref. [24], namely, the existence of cups & caps, suffices to derive the CEP. More specifically, if we take a time-symmetric purification of some state and then turn the input via a cup, this defines a complete extension of the state. Since both quantum and classical theory satisfy all of the postulates of Ref. [24], it therefore immediately follows that they both have complete extensions. Moreover, superselected quantum systems also satisfy the time-symmetric purification postulate and have cups & caps, hence they also satisfy CEP.

We now consider two results which have been proven using the purification postulate, and consider whether or not they can be reproven using the complete extension postulate instead.

3.1 Impossibility of bit-commitment

The task of *bit-commitment* is a two-party cryptographic task that can be used as a primitive in building up other important cryptographic protocols such as coin-flipping [65] and zero-knowledge proofs [66]. It is well known that bit-commitment and its generalisation to *integer-commitment* is impossible in both quantum and classical theory [67, 68]. However, this is not true in all GPTs, indeed in [50] they show that any GPT which is non-classical but does not have entanglement allows for bit-commitment. It is therefore interesting to study what are the features of quantum theory which make this task impossible.

There have been various works on this subject including [15, Corollary 45] and [33]. The former demonstrates the impossibility of bit-commitment under a set of physical postulates including the purification postulate. The latter goes beyond a strict impossibility proof and provides an analytic lower bound on the product of the two relevant cheating probabilities, again, however, relying on the purification postulate in the process. In this section we will show that we can derive exactly the same analytic lower bound but using the complete extension rather than the purification postulate.

Before getting to this result, however, let us first introduce the protocol. For honest parties this task consists of two phases: i) the commit phase, in which Alice chooses an integer $j \in \{1, \dots, n\}$ via a uniform random distribution and “commits” it to Bob by passing him a “token”. After doing so it should be impossible for Alice to change the value j but it should also be impossible for Bob to learn the value j . ii) the reveal phase, in which Alice communicates the integer j to Bob and sends him a second token to verify the integer. Bob must be able to check that the integer that Alice revealed was indeed the integer that she committed to in the first phase. Succinctly, for a protocol to be secure,

it must be the case that Bob cannot learn anything about the integer j prior to the reveal phase and that Alice cannot change j after the commit phase. The extent to which they can deviate from this ideal is characterised by the cheating probabilities p_B^* and p_A^* , respectively.

More formally, the protocol can be described as follows: (i) In the commit phase, Alice chooses an integer $j \in \{1, \dots, n\}$ via a uniform random distribution and creates a corresponding bipartite state $s^j \in \Omega_{A \otimes B}$ of the bipartite system $A \otimes B$. This is her ‘‘commitment’’ to the integer j . She then sends the system B to Bob to complete this phase. The idea behind this being that is now having access to only one part of the state will prevent Alice from being able to change her commitment, whilst at the same time, the fact that Bob only has access to one part of the state will prevent him from being able to learn the integer value. (ii) In the reveal phase, Alice communicates the integer j to Bob and sends him the system A , Bob applies a two-outcome measurement $(e_{\text{accept}}^j, e_{\text{reject}}^j)$ to the bipartite system AB to check that it is in fact in state s^j to verify that the integer that Alice reveals is indeed the integer that she originally committed to.

The cheating probabilities can then be described as follows. p_B^* is the maximum probability with which Bob can learn the value j prior to the reveal phase, that is, by performing some measurement on the reduced states on system B alone. In the case of quantum theory, this can be computed as a semi-definite program, whilst, as shown in [33], in the case of generic GPTs, they are computed via cone programs (a natural generalisation of SDPs). p_A^* is the maximum probability with which Alice can reveal an integer other than the value j that she committed to in the commit phase. There are many ways in which Alice can attempt to cheat but in order to find a lower bound on p_A^* we focus on the following particular cheating strategy: she applies some transformation to her part of the shared state between the two phases of the task thereby transforming s_j into some other state $s'_{j'}$. Alice’s success probability with this strategy is then the probability which Bob’s measurement $(e_{\text{accept}}^{j'}, e_{\text{reject}}^{j'})$ will give the accept outcome on state $s'_{j'}$. A perfect protocol would be one such that $p_B^* = p_A^* = 0$, that is, in which there is no probability that Bob can learn anything about j prior to the reveal phase, and there is no probability that Alice can influence the value after the commit phase. In the case that the GPT satisfies the complete extension postulate and an additional assumption called the no-restriction hypothesis [15], then it can be shown that such a perfect protocol is impossible. The following theorem encapsulates the result:

Theorem 14. *In any GPT satisfying the complete extension postulate and the no-restriction hypothesis [15], and in any integer-commitment protocol within that GPT, Alice and Bob’s cheating probabilities satisfy*

$$p_A^* \cdot p_B^* \geq \frac{\alpha}{n} > \frac{1}{2n} \quad (9)$$

Proof. (Adapted from [33]). The proof is identical up to the paragraph containing eq. 14. At which point we must rewrite the proof as follows:

Since ρ^j is the marginal of s^j , we call s^j a *extension* let t^j be an extension of r^j . Then

$$\chi^j := \frac{1}{u_{\mathcal{B}}[x] \cdot n} (s^j + t^j) \in K_{\mathcal{A} \otimes \mathcal{B}} \quad (10)$$

is an extension of x' for each j .

Now, consider a complete extension of x' , denoted $\tilde{\chi} \in K_{B \otimes E^*}$. Then, the GENERATION property implies that, for each j there exists $T_j : E^* \rightarrow A$ such that $\mathbb{1}_A \otimes T_j(\tilde{\chi}) = \chi^j$.

We now have the following cheating strategy for Alice: Alice prepares the state $\tilde{\chi}$ and passes system B to Bob keeping hold of the system E^* . Then to reveal j , she ‘steers’ the state $\tilde{\chi}$ to χ^j using T_j before sending j and the system A to Bob.

The remainder of the proof is identical relying on the use of the tool from convex-optimisation known as *cone programming* [69]. This is the natural generalisation of the notion of semi-definite programs, which are ubiquitous in quantum information theory, to the setting of GPTs. To date there have only been a handful of papers which utilise this tool either in quantum theory [70–74] or in GPTs [33, 75–80]. \square

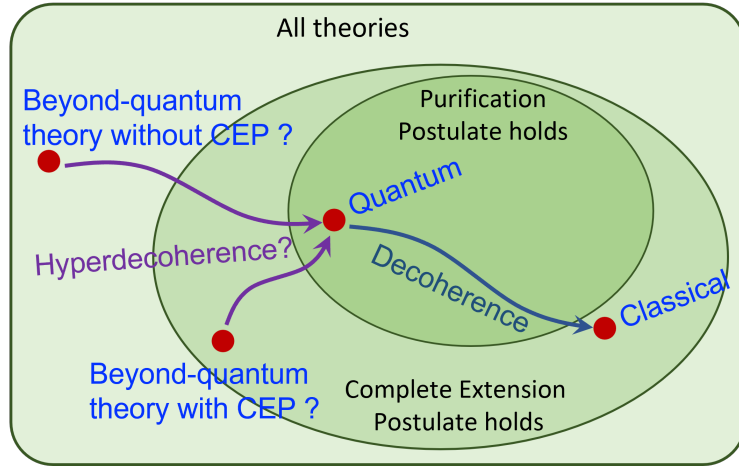


Figure 1: Schematic depiction of the relationship between the set of all theories, those are satisfying the complete extension postulate and those are satisfying the purification postulate. Also shown is the hyperdecoherence mechanism (purple arrows) relating hypothetical beyond quantum theories and quantum theory in analogy to the relationship between quantum and classical theory by a decoherence mechanism (blue arrow).

In the above, the no-restriction hypothesis [15] means that any logically possible measurement compatible with the state space of theory is physically realisable within the GPT. This is satisfied by both quantum and classical theory and is well motivated in adversarial scenarios when one does not want to make assumptions on the capabilities of the adversaries. Formally this means that, for every system A in the GPT, that the effect space \mathcal{E}_A is the dual of the state space Ω . That is:

$$\mathcal{E}_A = \Omega_A^* := \{e \in V_A^* | e(s) \in [0, 1] \forall s \in \Omega_A \subset V_A\} \quad (11)$$

That is, any convex-linear functional which gives valid probabilities for all states corresponds to a physically realisable effect.

3.2 No no-go for hyperdecoherence

The question that any beyond-quantum theory must answer is the following: why do we not observe beyond-quantum phenomena in all of our current experimental tests of quantum theory? That is, within the beyond-quantum theory there must be some mechanism which explains the emergence of the quantum world. We can gain intuition for this by considering the beyond-classical theory that is quantum theory. Within quantum theory, the mechanism that best explains the emergence of the classical world is decoherence.

It has recently been understood [81, 82] how decoherence is not simply a way in which quantum states can be made effectively classical (as diagonal density matrices are isomorphic to classical probability distributions) but that this can be lifted to a mechanism which shows how the entirety of quantum theory (i.e., including states, measurements, transformations, and so on) can be made effectively classical (i.e., the theory of classical stochastic dynamics). In [37] this idea was generalised to define the notion of hyperdecoherence (a term coined in [9]) which similarly describes how the entirety of a beyond-quantum theory can be made effectively quantum.

The basic idea, is that for every beyond-quantum system, A , there exists some hyperdecoherence process,

$$\mathbf{H}_A : A \rightarrow A, \quad (12)$$

which will cause the system A to behave essentially as quantum systems. These hyperdecoherence processes must satisfy three basic properties:

1. they must be unit-effect preserving,

$$u_A \circ \mathbf{H}_A = u_A, \quad (13)$$

which is analogous to the trace-preservation condition for quantum decoherence processes; and

2. that they must be idempotent,

$$\mathbf{H}_A \circ \mathbf{H}_A = \mathbf{H}_A, \quad (14)$$

which means that once the beyond-quantum features have been lost, that hyperdecohering again does nothing.

3. Moreover, they must be chosen compositionally, that is, such that

$$\mathbf{H}_{A \otimes B} = \mathbf{H}_A \otimes \mathbf{H}_B, \quad (15)$$

such that if two systems hyperdecohere independently then the global system will behave as a quantum system.

We then can model the hyperdecoherence of a beyond-quantum theory by replacing the identity processes $\mathbb{1}_A$ with hyperdecoherence processes \mathbf{H}_A , the intuition being that we are trying to describe a situation in which the hyperdecoherence happens on a time scale which is much shorter than any we can currently experimentally probe such that when we “do nothing” we are actually letting the system hyperdecohere. The conditions that we have imposed on the hyperdecoherence processes ensure that this can be described in consistent manner and that the result of this is a valid physical theory.

In order for such a procedure to describe the emergence of the quantum world from some beyond-quantum theory, it must be the case that the physical theory that we obtain by replacing the identities by hyperdecoherence processes in the beyond-quantum theory, is (isomorphic to) quantum theory. In [37], they demonstrate that any beyond-quantum theory which can hyperdecohere to quantum theory in this way, must violate either the causality principle – that information can only flow forwards in time, or the purification postulate. Letting go of purification seems like the more palatable option, and so this motivates the question of whether or not the complete extension postulate could be satisfied by a beyond-quantum theory which hyperdecoheres to quantum theory.

By examination of the proof of [37], however, it becomes clear that one cannot obtain the same result using the complete extension postulate – at least, not with the same proof technique. One may have suspected that the proof would still go through as it is clear that at no point does it crucially rely on the fact that a purification is actually pure. Instead, what is needed is the fact, which can be derived from the purification postulate, that pure extensions are generating. Or, in other words, that the purification postulate tells us that pure extensions are necessarily complete extensions. This fact, however, does not follow from the complete extension postulate alone, and so the proof does not go through. It therefore remains a possibility that there is a beyond-quantum theory satisfying the CEP, which moreover satisfies the causality principle and can hyperdecohere to quantum theory. An important direction for future research is therefore to try to formulate such a beyond-quantum theory.

4 Case study - theory of non-signalling behaviours

As a case study, in this section, we focus on the theory of non-signalling behaviours [5]. It exhibits stronger correlations between its subsystem than quantum theory, for example, it violates the Information Causality principle [83] and reaches the algebraic maximum for the CHSH inequality. It is also proved that, such behaviours can not be used for teleportation, and it does not allow for swapping of non-separable correlations [84,85] between two parties which were initially completely uncorrelated. This is not surprising as we already know that the theory of non-signalling behaviours violates the purification postulate (as it is a discrete theory), whilst in case of quantum theory, the ability to teleport an unknown state or entanglement swapping directly related to the existence of purifications [86,87].

In this section, we consider the lack of existence of purifications for the theory of non-signalling behaviours and try to bypass it with other entity fitted for the theory of non-signalling behaviours, namely, the complete extension that we introduced in Sec. 3. For a given behaviour P_A , we want to construct a non-signalling extension P_{AE} , which, whilst not being a purification, still satisfies the two key properties of purifications that we identified, namely:

ACCESS: The extension gives access to all ensembles of the extended system.

That is one can generate an arbitrary ensemble $\{(p_i, P_A^i)\}$ of P_A , (such that $\sum_i p_i P_A^i = P_A$ and $\sum_i p_i = 1$) from the extension P_{AE} .

GENERATION: The extension can be transformed to any other extension of P_A .

That is, upon applying pre(post)-processing of inputs(outputs) (a classical channel) \mathcal{P}_E on system E of P_{AE} , it should be transformed to an arbitrary extension $P'_{AE} = \mathcal{P}_E(P_{AE})$, e.g., $\text{tr}_E(P'_{AE}) = P_A$.

In the following, we show that this is indeed possible, and, hence, the theory of non-signalling behaviours satisfies the Complete Extension Postulate. In this way we show that CEP is not an empty postulate and complete extensions can be actually constructed in certain theories. Moreover, we explore the properties of such extensions. In particular, we show that there are infinitely many ensembles of each behaviour, therefore each behaviour can be extended to infinite number of different extensions. Each ensemble corresponds then to certain choice of input in the extending system. As a consequence, these extensions might possess an arbitrarily large number of inputs. But in the theory of non-signalling behaviours, one can have behaviours lying only in a finite-dimensional polytope. This shows that some of the ensembles (inputs) are redundant, in the sense that they can be obtained from others via adequate operations of pre(post)-processing of inputs(outputs), and others can be equivalent under reversible operations, e.g., relabelling of inputs and outputs (see Fig. 2). In this section, we will show that corresponding to each behaviour, there exist many finite-dimensional non-signalling extensions with the property of ACCESS and GENERATION. In fact, what we show, the above properties are equivalent, access to all possible ensembles is equivalent to generation of arbitrary extensions.

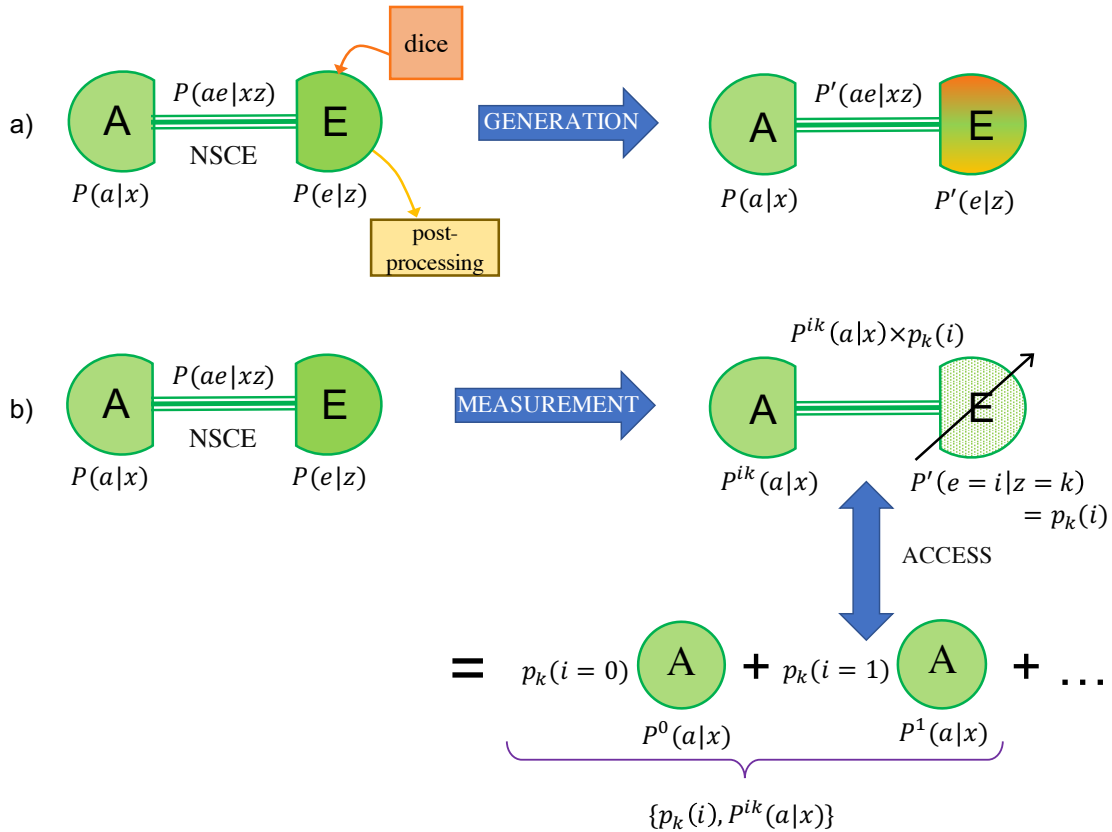


Figure 2: Schematic diagram. In any discrete convex theory, generic states do not have an extension which is a vertex (aka purification). This is different to the quantum theory that is not discrete. However, in non-signalling theory, the Complete Extension postulate is satisfied. a) From the Non-Signalling Complete Extension (NSCE) of a system, A one can generate any other extension by means of the input randomizer and output post-processing. b) The extending system E of the NSCE provides access to any ensemble $\{p_k(i), P^{ik}(a|x)\}$ of the system A , which can be generated upon measurement $z = k$ performed on the system E .

We also show that there is always a finite-dimensional complete extension, usually of a large dimension. We give an upper bound on this dimension as a function of the dimension of an extended

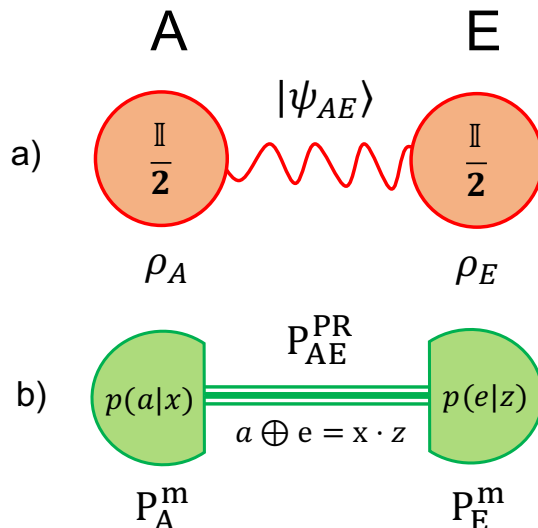


Figure 3: Schematic diagram of purifications of system A to an extended system E . a) Purification of any quantum state ρ_A , to $|\psi_{AE}\rangle$ in the extended state space. If the quantum state is a single qubit maximally mixed state $\frac{\mathbb{I}}{2}$, then its purification is maximally entangled Bell state. b) Complete extension of a maximally mixed behaviour P_A^m , (given in Eq. (50)) is a pure behaviour - the well known PR box, which is however generically not the case.

system and the number of inputs and outputs. Among these possible complete extensions, we give a prescription of how to construct (up to relabeling) the minimal one, which we call the non-signalling extension with access (NSEA). This lies in the lowest dimensional polytope among all possible extensions with the above properties. The fact that access to all possible ensembles of a non-signalling system, has been considered operationally for the worst case extension someone might have, in context of device independent cryptography [88, 89] against non-signalling adversary by the Authors of Ref. [89–92], and more recently in private randomness [93–98]. The NSEA which we define is the structure responsible for this fact: access to this special extension, gives the non-signalling eavesdropper an ultimate operational power. That is, the extending system of NSEA naturally represents the minimal memory of a non-signalling adversary with maximal power in cryptography based on no-faster than light communication [51].

Interestingly, whilst we know that not all NSEAs can be pure, we show that some of them are. In particular, the aforementioned NSEA of a maximally mixed binary input binary output behaviour, i.e., the PR box [5] is a pure non-signalling behaviour. Observing that a PR box is an extension with access for a maximally mixed system can be viewed as a *derivation* of the PR box without referring to the notion of a Bell inequality – this is in stark contrast to the original approach of Popescu and Rohrlich [5]. The sense in which we derive the PR box is the following. We assume only a) the structure of single-partite systems, and b) that NSCE of any single-partite system is a valid state of the theory. In particular, we do *not* presuppose a) that all non-signalling behaviours belong to the given theory, or b) a particular composition rule, or c) the no-restriction hypothesis. This means that if the PR box was not the NSCE of any single-partite behaviour, then it would be “excluded” from the state space built with the above prescription – i.e., a priori, we do not know that the PR box is an allowed physical behaviour. What we show is that the PR box is the complete extension of the maximally mixed single-partite behaviour, and hence, by our assumptions, it is a valid bipartite state.

Since an NSEA P_{AE} is generically not extreme in the set of behaviours, it can be non-trivially extended to a new system E' giving a new NSEA $P_{AE'E'}$. Note that an interesting corollary of Theorem 7 is that for particular behaviours this process can be iterated indefinitely. Thus, identifying or constructing such behaviours and understanding their properties is an important direction for future research.

From our proof technique (heavily based on convex geometry) we expect that the complete extensions exemplified using the theory of non-signalling behaviours can also be defined in a similar

way in any discrete convex theory. Proving this, however, needs further work.

In Sec. 4.1, we provide the basic definition of the *Non-signalling extension with access* (NSEA) and prove that it is a complete extension as it satisfies ACCESS and GENERATION, and moreover, in Sec. 4.2 that it is a minimal complete extension. In App. B we construct many explicit examples and study their properties. In particular in App. B.1, we prove that the NSEA of the maximally mixed single party binary input and binary output behaviour is the Popescu-Rohrlich behaviour (See Fig. 3). We also give an explicit example showing that the non-signalling complete extension has the property of ACCESS and GENERATION. In App. B.4, we find the NSCE for two party binary input and binary output behaviour in the isotropic line, which enclose the Bell-Tsirelsen box and the NSCE for Specker's triangle has been found in Sec. B.3.

4.1 The complete extension in the theory of non-signalling behaviours

In this section, we provide a definition of the *non-signalling extension with access* (NSEA) in the theory of non-signalling behaviours (NS), and show that NSEA satisfies the property ACCESS. We then show that NSEA also satisfies the GENERATION property, and, moreover, that in this theory these two properties are actually equivalent. This will allow us to refer to NSEA as a non-signalling complete extension (NSCE), and demonstrates that NS satisfies the complete extension postulate (CEP). What is interesting here is that it provides a constructive proof that NS satisfies the CEP, which proves useful for introducing explicitly the state of the system of the eavesdropper in the non-signalling device independent secure key agreement protocol [51].

Let us first fix the notation. We call a conditional probability distribution, $P_A = p_{A|\mathcal{X}}(a|x)$ a box or behaviour, representing the state of system A , where \mathcal{X} stands for all possible measurement choices (inputs) and the index $x \in \mathcal{X}$ is a particular choice. \mathcal{A}_x is the set of measurement outcomes (outputs), corresponding to input x and the index $a \equiv a_x \in \mathcal{A}_x$ represents one particular instance. Note that if all of the outputs have the same outcome set then we will drop the subscript x and write \mathcal{A} instead. As P_A is a conditional probability distribution it must satisfy the positivity conditions, $0 \leq p_{A|\mathcal{X}}(a|x) \leq 1$, $\forall a \in \mathcal{A}_x$, $x \in \mathcal{X}$, and the normalization condition $\sum_{a \in \mathcal{A}_x} p_{A|\mathcal{X}}(a|x) = 1$, $\forall x \in \mathcal{X}$ ³. Let $P_{AE} = p_{AE|\mathcal{X}\mathcal{Z}}(ae|xz)$ be a bipartite conditional probability distribution, where the system E has associated input and output sets \mathcal{Z} and \mathcal{E} respectively. We say that it is a non-signalling extension of the behaviour P_A if P_A is its marginal distribution

$$\sum_{e \in \mathcal{E}_z} p_{AE|\mathcal{X}\mathcal{Z}}(ae|xz) = p_{A|\mathcal{X}}(a|x), \quad \forall a \in \mathcal{A}_x, x \in \mathcal{X}, z \in \mathcal{Z}, \quad (16)$$

and it satisfies non-signalling conditions

$$\sum_{e \in \mathcal{E}_z} p_{AE|\mathcal{X}\mathcal{Z}}(ae|xz) = \sum_{e \in \mathcal{E}_{z'}} p_{AE|\mathcal{X}\mathcal{Z}}(ae|xz'), \quad \forall a \in \mathcal{A}_x, x \in \mathcal{X}, z, z' \in \mathcal{Z}, \quad (17)$$

$$\sum_{a \in \mathcal{A}_x} p_{AE|\mathcal{X}\mathcal{Z}}(ae|xz) = \sum_{a \in \mathcal{A}_{x'}} p_{AE|\mathcal{X}\mathcal{Z}}(ae|x'z), \quad \forall e \in \mathcal{E}_z, z \in \mathcal{Z}, x, x' \in \mathcal{X}, \quad (18)$$

For the sake of simplicity we omit the subscript, and we assign $p_{A|\mathcal{X}}(a|x) \equiv P_A(a|x)$. The set of all non-signalling behaviours of system A is denoted by Ω_A , mutatis mutandis for multipartite systems. The P_A can have inner non-signalling structure [5], i.e., system A can be compound (e.g., $A = A_1 A_2$) itself, where the system A_1 can not signal to system A_2 and vice versa. We will encounter this later when we consider the extension of the binary two input two output behaviour.

All n -partite non-signalling behaviours, with the same cardinalities of inputs and outputs satisfy a set of linear equations and inequalities (constraints). These constraints define a convex, bounded polyhedron (a polytope) that is a subset of \mathbb{R}^N for some $N \in \mathbb{N}$, with a finite number D of vertices. Each point within this polytope represents a different non-signalling behaviour. In such a polytope of behaviours, we will distinguish those which are *extremal* (vertices), i.e., the behaviours which can not be expressed as a convex combination of other non-signalling behaviours, distinguishing them by subscript \mathbf{E} : $P_{\mathbf{E}}$.

³The summation over the output here is an analogue of partial trace.

An ensemble $\{(p_i, P^i)\}_i$, of a behaviour P , where $P = \sum_i p_i P^i$, we will denote by $\mathcal{E}(P)$. In this paper, we will consider only those ensembles which consists of finite number of elements. We will also need the notion of the set of *members* of an ensemble $\mathcal{E}(P) = \{(p_i, P^i)\}_i$, this is defined by $V(\mathcal{E}(P)) = V(\{(p_i, P^i)\}_i) := \{P^i : p_i > 0\}$, and its *distribution*, which is $\{p_i\}$. If *all* the members, $P^i \in V(\mathcal{E})$, are extremal (pure), P_E^i , then we call the ensemble a *pure members ensemble* (PME) and denoted as $\mathcal{E}_{pure}(P)$. We say that an ensemble $\mathcal{E}_{pure}(P)$ is generated on system A by measurement \mathcal{M} on system E , if upon this measurement on the extending system, the outcome $e = i$ is obtained on E with probability p_i and conditionally upon it, the state of the system A is described by behaviour P_E^i . By S_P we will denote the set of all PME of a behaviour P .

Our aim is to obtain the minimal extension with properties ACCESS and GENERATION. To begin with, however, as an intermediate step used to simplify further proofs, we will define an extension, satisfying properties ACCESS and GENERATION yet of a larger dimension hence called *overcomplete non-signalling extension with access* (ONSEA), defined as:

Definition 15 (Overcomplete non-signalling extension with access - ONSEA). *Given a behaviour $P_A : P_A(a|x)$, we say that a behaviour $P_{AE} : P_{AE}(ae|xz)$ is its overcomplete non-signalling extension with access (extension to system E with access to system A), if for any input choice $z = k$ and outcome $e = i$ obtained in the extending part, there holds*

$$P_{AE}(a, e = i|x, z = k) = P_E^{i,k}(a|x)p(e = i|z = k), \quad (19)$$

such, that, for each k , the ensemble $\left\{ \left(p(e = i|z = k), P_E^{i,k}(a|x) \right) \right\}_i$ is a pure members ensemble of the behaviour P_A , and corresponding to each pure members ensemble of P_A , there is exactly one input $z = k$, in the extending system which generates it.

It is simple to see that the overcomplete extension ONSEA of an arbitrary behaviour P_A exists. Indeed, P_A belongs to a polytope. That is it belongs to a set with a finite number of pure behaviours. Any pure member ensemble of P_A is a subset of the set of pure behaviours. Hence, there is finite number of the latter ensembles and we can construct extension where for each of such ensemble we have k such that $z = k$ generates it. This extension is precisely the ONSEA. We will denote it as $\tilde{\mathcal{E}}(P)_{AE}$. Notice that when a particular input $z = k$ is chosen in the extending part, an outcome $e = i$ occurs with the probability $p(e = i|z = k)$, resulting in a pure behaviour $P_E^{i,k}(a|x)$ in part of A . Eq. (19), expresses the partially measured behaviour with the probability times the conditional pure behaviour. Moreover, $\left\{ \left(p(e = i|z = k), P_E^{i,k}(a|x) \right) \right\}_i$ is a pure members ensemble of P_A , for each k , as

$$\sum_i p(e = i|z = k) P_E^{i,k}(a|x) = \sum_i \tilde{\mathcal{E}}(P)_{AE}(a, e = i|x, z = k) = P_A(a|x), \quad (20)$$

and given an overcomplete extension (ONSEA) there are $|\mathcal{Z}|$ pure members ensembles, where $\mathcal{Z} = \{z\}$, is the set of all input choices of the extending system.

The above definition of ONSEA satisfies the non-signalling condition for its both subsystems. For system A , it is by construction, and for system E it holds due to the fact that for each input output pair $(z = k, e = i)$ of E , system A holds a behaviour $P_E^{i,k}(a|x)$ according to Eq. (19), which gives 1 when summed over a .

We will see later that the ONSEA is actually a complete extension as it allows for ACCESS and GENERATION. However, the dimension of the extending system is larger than necessary, we therefore seek to construct a minimal complete extension⁴. To do so we will now introduce a representative subset of all PMEs. An ensemble $\mathcal{E}_{pure}(P) = \{p_i, P_E^i\}_i$, of a behaviour $P = \sum_i p_i P_E^i$, is called *minimal ensemble*, if any proper subset $V' \subset V(\mathcal{E}_{pure})$, with another choices of probabilities $\{p'_j\}$ is not an ensemble of the behaviour P , i.e., $P \neq \sum_{j: P_E^j \in V'} p'_j P_E^j$. Any minimal ensemble will be denoted as $\mathcal{M}(P)$, to distinguish it from an arbitrary PME that is not necessarily minimal.

Next, we will eliminate the redundant ensembles from the ONSEA, to obtain the NSEA that has the lowest number of inputs sufficient to satisfy ACCESS property. To achieve this we will show

⁴Minimality will be proven in Prop. 27 in the following subsection.

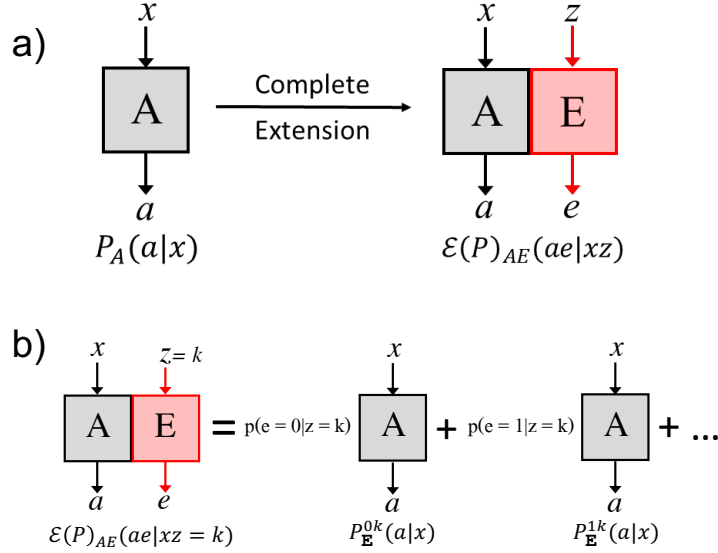


Figure 4: Panel *a*), schematic diagram of a NSEA of an arbitrary behaviour P_A to $\mathcal{E}(P)_{AE}$. Panel *b*), Corresponding to each input $z = k$ in part of the extending system, the composite behaviour partitioned to a minimal ensemble $\{p(e|z = k), P_E^{ek}(a|x)\}$ in part of A .

that having access to all minimal ensembles $\{\mathcal{M}(P)\}$ of P , along with arbitrary randomness, one can generate any PME, $\mathcal{E}_{pure}^{\mathcal{E}}(P)$.

Theorem 16. *ACCESS to PMEs is equivalent to ACCESS to minimal ensembles.*

Proof. See App. A.2. □

Remark 17. *This theorem (Thm. 16) can be generalized to any GPT that is a discrete theory (see Definition 5).*

The above theorem motivates the following definition of *non-signalling extension with access* (NSEA), which, like the extension defined in Definition 15, also satisfies the properties of access and generation:

Definition 18 (Non-signalling Extension with Access). *Given a behaviour $P_A : P_A(a|x)$, we say that a behaviour $P_{AE} : P_{AE}(ae|xz)$ is its non-signalling extension with access extended to system E if for any input choice $z = k$ an outcome $e = i$ occurred in the extending system, there holds*

$$P_{AE}(a, e = i|x, z = k) = P_E^{i,k}(a|x)p(e = i|z = k) \quad (21)$$

such, that, for each k , the ensemble $\left\{ \left(p(e = i|z = k), P_E^{i,k}(a|x) \right) \right\}_i$ is a minimal ensemble of the behaviour P_A . Moreover, corresponding to each minimal ensemble of P_A , there is exactly one input $z = k$, in part of the extending system which generates it.

We will represent the NSEA of an arbitrary behaviour P_A as $\mathcal{E}(P)_{AE}$. If there is more than one subsystem in the system A satisfying non-signalling constraints, i.e., if $A \equiv A_1 A_2 \dots A_N$, and none of the k systems $\{A_{i_1}, \dots, A_{i_k}\}$ with $I = \{i_1, \dots, i_k\}$ can signal to the remaining $N - k$ systems $\{A_{j_1}, \dots, A_{j_{N-k}}\}$ with $j_l \in \{1, \dots, N\} \setminus I$, for all $k \in \{1, 2, \dots, N-1\}$, then also our method of constructing NSEA holds, and it will be denoted by $\mathcal{E}(P)_{A_1 A_2 \dots A_N E}$. A schematic diagram of NSEA has been depicted in Fig. 4. For NSEA defined above, the inputs of the extending system correspond to the minimal ensembles of the given behaviour P_A . If there are N number of minimal ensembles, then the total number of input choices in the extending part is $|\mathcal{Z}| = N$.

Proposition 19. *For each behavior P_A of system A , its NSEA $\mathcal{E}(P)_{AE}$ is unique up to the local relabeling on the extending system E .*

Proof. Let P_{AE} and P'_{AE} be two different NSEA of a behaviour P_A . First note, that P_A determines the number of inputs $|E|$. However the inputs of P_{AE} can be labelled in different way than that of P'_{AE} . However, due to Definition 18, each input of P_{AE} corresponds to a unique minimal ensemble of P_A . The same holds for P'_{AE} . Thus for each input e of P_{AE} there exists input e' of P'_{AE} which corresponds to the same unique minimal ensemble. Hence for this pair of inputs the distribution of outputs of P_{AE} and P'_{AE} are the same up to labeling of outputs. For this reason, the two extensions differ only by relabelling of inputs and outputs of the extending system. Thus, relabellings establish an equivalence class over different behaviors being NSEA of behavior P_A . \square

Note that one can easily show that such an extension necessarily exists for any behaviour P_A . For example, one can check (see Section B.1 of the Appendix) that the NSEA of the behaviour P_A , the maximally mixed behaviour with a single binary input and single binary output, extended to system E , has the following structure:

$$P_{AE}(ae|xz) = \begin{array}{c|cc|cc} & x & 0 & 1 & & \\ & \swarrow e \searrow a & & & & \\ z & & 0 & 1 & | & 0 & 1 \\ \hline 0 & 0 & 1/2 & 0 & | & 1/2 & 0 \\ & 1 & 0 & 1/2 & | & 0 & 1/2 \\ \hline 1 & 0 & 1/2 & 0 & | & 0 & 1/2 \\ & 1 & 0 & 1/2 & | & 1/2 & 0 \end{array}, \quad (22)$$

which is nothing but the famous Popescu-Rohrlich (PR) behaviour satisfying conditions $x.z = a \oplus e$ with \oplus being addition modulo 2. We have thus arrived at this structure without referring to the CHSH inequality [99] (in contrast to the way in which it was done in [5]).

We are in a position to show, that having access to NSEA we can generate any PME.

Corollary 20. *The non-signalling extension with access (NSEA) of a behaviour P given in Definition 18, together with access to arbitrary local randomness, gives access to any pure members ensemble of a behaviour P .*

Proof. See App. A.3. \square

From the above corollary it is clear that an arbitrary PME can be accessed from the NSEA by using a randomness generator $\{p(k)\}$, i.e., by using a dice(coin). To access all possible PME, one needs an access to arbitrary randomness. This can be done by setting the output (k) of a dice with a distribution, $p(k|z')$, where z' is the tuning parameter, as the input of the extending party of NSEA. Here $|\{k\}| = |\mathcal{Z}|$, and $|\{z'\}|$ will be equal to the possible number of PME one wants to generate. The dice can be thought of as a local behaviour with z' being the input and k as the output. Different choices of the z' can be considered as dices with different probabilities of outcome, actually led to different PMEs. Accessing all possible PME has been pictorially depicted in Fig. 5.

An explicit example of constructing an arbitrary pure members ensembles has been given in Sec B.1.2. Where we have chosen an arbitrary behaviour containing single binary input and single binary output.

Theorem 21. *ACCESS to PMEs is equivalent to ACCESS to all ensembles.*

Proof. Let us consider the access of the mixed ensemble $\mathcal{E}_{mix}(P) = \{(p_m, P^m)\}_m$, which is the most general ensemble, from the set of all ensembles of a given behavior P . Now each P^m lies in the same polytope as P , hence, all of them has a pure behaviour decomposition,

$$P^m = \sum_i q_i^m P_{\mathbf{E}}^i, \quad (23)$$

where $\sum_i q_i^m = 1$, $\forall m$ and $0 \leq q_i^m \leq 1$, $\forall i, m$. Note that this decomposition is not unique, unless it is a minimal decomposition. Now $P = \sum_m p_m P^m = \sum_{m,i} p_m q_i^m P_{\mathbf{E}}^i = \sum_i r_i P_{\mathbf{E}}^i$, where $r_i = \sum_m p_m q_i^m$, implies that $\{(r_i, P_{\mathbf{E}}^i)\}_i$ is also a PME of P . From Theorem 20, we know that by using an appropriate

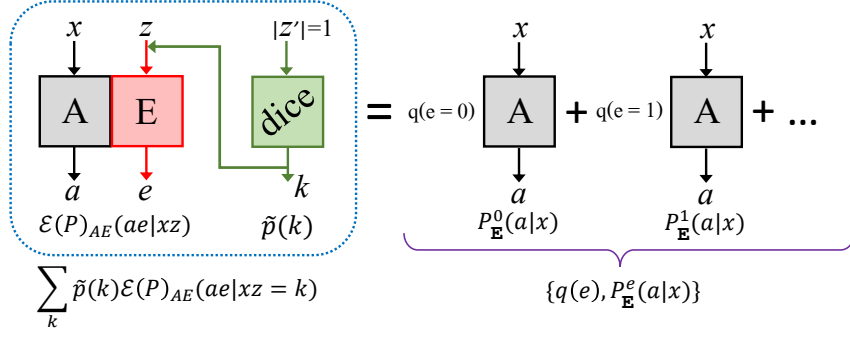


Figure 5: Schematic diagram visualizing the mixing the minimal ensembles $\{\mathcal{M}_k\}$ of A with arbitrary randomness $p(k)$ in part of the extending system on the NSEA of the behaviour, which is obtained from the output of a dice (a local behaviour with unary input), results an arbitrary pure members ensemble $\{p(k), \{(p(e|z=k), P_E^{e|z=k}(a|x))\}\} = \{q(e), P_E^e(ab|xy)\}$, where $q(e) = \sum_k p(k)p(e|z=k)$.

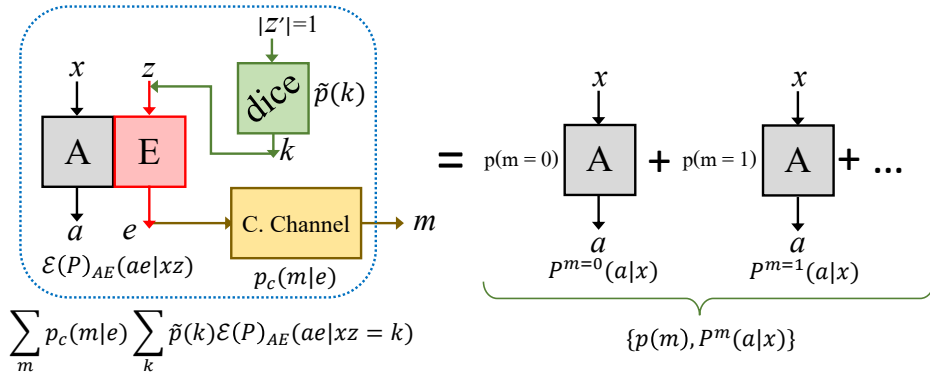


Figure 6: Explanation of Theorem 22 of accessing all possible ensembles even mixed in part of the extended system. By passing the output e of the extending party's (Eve) behaviour, through a post-processing channel PC, Eve is able to interpret the behaviour shared by Alice and Bob as an ensemble of mixed bipartite behaviours $\{p(m), P^m(ab|xy)\}$, where $P^m(ab|xy)$, are mixed behaviours and can be expanded as $P^m(ab|xy) = \sum_e q_e^m P_E^e(ab|xy)$. Now the post-processing channel $p_c(m|e)$, helps Eve to interpret the mixed behaviours as $\sum_e p_c(m|e)r(e)P_E^e(ab|xy) = p(m)P^m(ab|xy)$.

randomness generator in the input of the extending system E (input randomizer), such PME can be realized. To interpret the mixed ensembles, from $\{(r_i, P_E^i)\}_i$, the extending system send her output ($e = i$, of the NSEA along with an appropriate input randomizer) through a post-processing channel $p_c(m|e)$. The existence of this channel (24), $p_c(m|e)$ with $\sum_m p(m|e) = 1$, is guaranteed by the initial decomposition of the members of the mixed ensemble into pure behaviours.

$$\sum_i p_c(m|e=i)r_i P_E^i = \sum_i \frac{p_m q_i^m}{r_i} r_i P_E^i = p_m P^m, \quad (24)$$

where $p_c(m|e)$ take the form $\frac{p_m q_e^m}{r_e}$. □

Corollary 22. (NSEA satisfies ACCESS) *The extending system of the NSEA gives access to any possible (even mixed) ensemble of the extended behaviour.*

Proof. Suppose the extending system wants to access an arbitrary mixed ensemble, once she obtained the NSEA. From Corollary 20, NSEA gives access to any PME, and from Theorem 21, it is clear that access to PMEs is equivalent to access to all ensembles. Hence, NSEA gives access to any possible (even mixed) ensemble of the extended behaviour. □

From the above theorem, we observe that, corresponding to each mixed ensemble of P , there exists at least one PME of P , from which one can obtain the mixed ensemble by using an appropriate

post-processing channel. In this construction the NSEA plays an important role due to the fact that Theorem 20 certifies the existence of that particular PME. In Sec. B.1.3 of the Appendix, we exemplify Theorem 22 for a single party binary input binary output behaviour, by providing the explicit form of input randomizer and the post-processing channel. A visualization of Theorem 22, has been given in Fig. 6. Note that the input randomizer $\tilde{p}(k)$, which generates the required PME of P, is a dice with a unary input (trivial input), $\tilde{p}(k|z' = \text{fix})$. One can generate all possible ensembles by appropriately tuning the dice with the help of the dice input z' and a post-processing channel (also dependent on z') on part of the extending system.

We have now seen that we can construct the NSEA which is an extension with ACCESS. However, in order for this to be a complete extension we need that it also satisfies GENERATION. One can show that this is indeed the case by demonstrating that ACCESS and GENERATION are equivalent conditions in the theory of non-signalling behaviours.

Theorem 23. *In the theory of non-signalling behaviours ACCESS is equivalent to GENERATION.*

Proof. See App. A.4. □

From the above theorem, Theorem 22, and the fact that NSEAs exist for all behaviours, we obtain the main result of this section:

Corollary 24. *The Non-signalling Extension with Access (NSEA) has the properties of ACCESS and GENERATION, that is, it is a Non-signalling Complete Extension (NSCE). Hence, the theory of non-signalling behaviours satisfies the Complete Extension Postulate (CEP).*

4.2 On the dimensionality of the complete extension

In this subsection we are going to discuss the size of the non-signalling complete extension NSEA, and, what is most important, is that we are going to show that it is finite. The restriction on the size of the non-signalling complete extension (NSEA) is particularly important from point of view of non-signalling Device Independent cryptography [89, 90, 100], and interesting on its own. We finish this section with another observation, namely we show that for any behaviour P , its NSEA $\mathcal{E}(P)$ has the lowest dimension (is minimal) among all non-signalling extension of P having the property of ACCESS.

The following Theorem holds.

Theorem 25. *Let be \mathcal{B} a polytope of n -partite non-signalling behaviours, with m_i inputs for parties and v_{ij} outputs respectively. Then for each $P \in \mathcal{B}$, there exists a non-signalling polytope $\tilde{\mathcal{B}} \ni \mathcal{E}(P)$ which the NSEA lives in, such that:*

$$\dim \tilde{\mathcal{B}} < (\dim \mathcal{B} + 1) \times \left(\left(\binom{2t - \lfloor t/2 \rfloor - \dim \mathcal{B}}{\lfloor t/2 \rfloor} + \binom{3t - \lfloor t/2 \rfloor - (\dim \mathcal{B} + 1)}{t - \lfloor t/2 \rfloor - 1} \right) \dim \mathcal{B} + 1 \right), \quad (25)$$

where:

$$\dim \mathcal{B} = \prod_{i=1}^n \left(\sum_{j=1}^{m_i} (v_{ij} - 1) + 1 \right) - 1, \quad t = \prod_{i=1}^n \sum_{j=1}^{m_i} v_{ij}. \quad (26)$$

Proof. The proof is the content of Section B.2 of the Appendix. □

In Theorem 25 we were interested only in showing that the dimension of the complete extension is always finite and therefore the upper bound is very loose. For instance, the expression in equation (25) in the case of $n = 1$, $m_i = 2$, $v_{ij} = 2$ yields the dimension of 339, for $n = 2$, $m_i = 2$, $v_{ij} = 2$ we have c.a. 1.2×10^{54} , and for $n = 2$, $m_i = 3$, $v_{ij} = 3$ we obtain c.a. 1.14×10^{1762} . All exemplified results are far above any numerical predictions. The discussion about the possible improvements in the upper bound is left to the Appendix. Importantly for us, however, is the following simple corollary of the existence of any upper bound, namely:

Corollary 26. *The dimension of the NSCE can always be finite.*

A tempting question to ask about NSEA is whether it has the lowest dimension amongst all NSCEs, and therefore whether it is the minimal one. Similarly to earlier in this section, we refer to the dimension of the extension as the dimension of the behaviour polytope to which it belongs. The answer to this query is positive and therefore we finish this section with the following proposition.

Proposition 27. *Among all non-signalling extensions of a behaviour P , having the property of ACCESS, NSEA $\mathcal{E}(P)$ is a minimal one.*

Proof. Let P be a non-signalling behaviour in \mathcal{B} , and $\mathcal{E}(P) \in \tilde{\mathcal{B}}$ be its NSEA. The dimensions of non-signalling polytopes can be determined with equation (89) as before. Suppose now, there exists another non-signalling extension $\hat{\mathcal{E}}(P) \in \hat{\mathcal{B}}$, having a property of ACCESS, such that $\dim \hat{\mathcal{B}} < \dim \tilde{\mathcal{B}}$. The fact that $\hat{\mathcal{E}}(P)$ has the property of ACCESS implies that upon processing (possibly trivial) of inputs and outputs of the extending system with local randomness, $\hat{\mathcal{E}}(P)$ also has access also to all minimal ensembles of P . However, the minimal ensembles of P are extremal points in the polytope of all possible ensembles (see Theorem 30) of the behaviour P , and so cannot be created via probabilistic processing of inputs and outputs in the extending system (class of operations considered in ACCESS). Therefore, $\hat{\mathcal{E}}(P)$ having a property of ACCESS, must have for each minimal ensemble of P , an input that generates it, like $\mathcal{E}(P)$ does. This implies that $\dim \hat{\mathcal{B}} \geq \dim \tilde{\mathcal{B}}$, and so proves by contradiction that NSEA $\mathcal{E}(P)$ is the minimal extension of P having the property of ACCESS. \square

5 Conclusions

To summarize, our main contribution is a new concept, which is the *complete extension*. We show that complete extensions are present in classical theory, quantum theory, super-selected quantum theory, the theory of non-signalling behaviours, and, moreover, any theory satisfying the purification postulate in which the product of pure states is pure. We also postulate that it may exist hypothetical beyond-quantum theories which could hyperdecohere to quantum theory. In the case of quantum and classical theory, as well as the theory of non-signalling behaviours, we can explicitly construct these complete extensions. This notion implies a number of paths for research, some of which we have exemplified in our case study on the theory of non-signalling behaviors.

The idea of the CEP sets a demarcation line in the set of results obtained on the basis of the purification postulate. It divides them into those that really require all of the purification postulate and those for which the CEP suffices. We exemplify this by considering the possibility of bit commitment showing that the no-go for it is not specific to theories with the purification postulate. This has the added benefit of giving a unified proof for the quantum and classical cases.

The CEP may also be viewed as a razor for excluding theories that can not substitute quantum theory in the future. Indeed, a theory not satisfying CEP may not be physical, as (see Theorem 24) even the theory of non-signalling behaviours satisfies it. The easiest way to obtain theories that do not satisfy CEP is to restrict state space or dynamics so that it is not possible to generate all extensions of some state. An example of a theory that does not satisfy CEP for this reason is given in [50].

We also show that an interesting “mirror” property of quantum purifications no longer holds for the case of non-signalling behaviours. That is, suppose that we have a purification $|\psi_{AB}\rangle$ of a quantum state ρ_A , and let us define $\sigma_B := \text{Tr}_A(|\psi_{AB}\rangle\langle\psi_{AB}|)$. Then, any purification of σ_B on system A , (which necessarily exists as $|\psi_{AB}\rangle$ is such a purification) we denote as $|\phi_{AB}\rangle$ and call the “mirror” purification, is equal to $|\psi_{AB}\rangle$ up to a local unitary on A . That is $|\psi_{AB}\rangle = U_A \otimes \mathbb{1}_B |\phi_{AB}\rangle$ for some unitary U_A . In general, beyond-quantum theory, it need not be the case, as we have exemplified in the theory of non-signalling behaviours (see Sec. B.1.6 of the Appendix). A properly defined minimal distance between the complete extension and its “mirror” one, overall systems whose complete extension is not pure, can characterise to what extent a given theory departs from quantum mechanics.

Following the development of quantum cryptography, one can ask if a post-quantum theory should lead to secure communication. As it is shown in [101], this need not always be the case. In the quantum case, the system E of the purification ψ_{AE} describes the worst-case state of the knowledge of the quantum adversary that may have about a given system A in many information processing protocols

such as QKD. Similarly, in the theory of non-signalling behaviours, the system E of the NSCE τ_{AE} represents the worst case knowledge of an adversary about system A . The minimal dimension of the system E , therefore, represents the maximal memory needed by the adversary. We, therefore, enable the study of the worst-case adversary’s capabilities in other post-quantum theories, see Ref. [51] for an application of this idea.

Significantly, one can use CEP to define composed systems in a post-quantum theory. Namely given a set of states of a single system A , one can define as valid states of a joint system AB only those that are either (i) complete extensions of the states of A , or (ii) states obtained from such complete extensions by some operations valid in this theory. We exemplify this approach by arriving at the structure of the PR box, assuming (i) in the case that A is a local system from NS with two binary inputs and two binary outputs. Together with (ii) the local transformations in NS (including e.g. relabelings of inputs and outputs), we can reach all the non-local vertices of the non-signalling polytope of behaviours with two binary inputs and two binary outputs. Given the fact, that the complete extension of a local deterministic behavior of system A is a product of two deterministic behaviors on systems $A \otimes B$, via operation of mixing two or more behaviors we show that the state space of non-signalling behaviours includes the whole non-signalling polytope of behaviours with two binary inputs and two binary outputs.

Interestingly, a post-quantum theory that does not satisfy the purification postulate, but rather the CEP, can have a property that there is an infinite sequence of complete extensions of complete extensions as none of them is pure. This will always hold unless the number of pure states in the theory is of cardinality continuum at least. This is indeed the case in both classical theory and the theory of non-signalling behaviours.

Verifying if CEP holds in the case of other beyond-quantum theories and studying its consequences is an important direction to follow. We formulate the hypothesis that there exists a generalised probabilistic theory with complete extensions that may naturally hyper-decohere to quantum mechanics. This is supported by demonstrating that the proof of an existing no-go result, which applies to generalised probabilistic theories with purifications, no longer holds. The confirmation of this hypothesis would open a new arena in which new physical laws and phenomena may be searched.

It is also interesting to study whether complementing CEP with a more dynamical axiom, e.g., the one linked to the Neumark extension or Stinespring dilations, can lead to a more powerful postulate. In the case of the purification postulate such a dynamical postulate can be derived, it is therefore interesting to investigate whether this is also the case for a dynamical version of the CEP, or whether this must be additionally postulated. Determining what form this dynamical postulate should take, and demonstrating that it indeed holds in the theory of non-signalling behaviours, is perhaps the most important direction for follow up research.

Acknowledgments

MW, TD and KH acknowledge grant Sonata Bis 5 (grant number: 2015/18/E/ST2/00327) from the National Science Center. TD acknowledge Omer Sakarya for his help in numerical simulation. KH, PH and LP acknowledges ERC grant QOLAPS and Polish Ministry of Science and Higher Education Grant no. IdP2011 000361. s MP acknowledges support from European Union’s Horizon 2020 Research and Innovation Programme under the Marie Skłodowska-Curie Action OPERACQC (Grant Agreement No. 661338), and from the Foundational Questions Institute under the Physics of the Observer Programme (Grant No. FQXi-RFP-1601). RR acknowledges support from the research project “Causality in quantum theory: foundations and applications” of the Foundation Wiener-Anspach and from the Inter-university Attraction Poles 5 program of the Belgian Science Policy Office under the grant IAP P7-35 photonics@be. KH was partially supported by the Foundation for Polish Science through IRAP project co-financed by EU within Smart Growth Operational Programme (contract no. 2018/MAB/5). JHS was supported by the Foundation for Polish Science through IRAP project co-financed by EU within Smart Growth Operational Programme (contract no. 2018/MAB/5). JHS thanks Ana Belén Sainz for helpful discussions.

Code availability

The code that supports the theoretical plots and tables within this paper is available from the corresponding author upon reasonable request.

References

- [1] Asher Peres. “Karl Popper and the Copenhagen interpretation”. *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics* **33**, 23–34 (2002).
- [2] Paul Adrian Dirac. “The principles of quantum mechanics (third ed.)”. *Clarendon Press Oxford*. (1948).
- [3] J. von Neumann. “Mathematische grundlagen der quantenmechanik”. *Springer, Berlin*. (1932).
- [4] Johann von Neumann. “Mathematische grundlagen der quantenmechanik”. *Springer*. (1932).
- [5] S. Popescu and D. Rohrlich. “Quantum nonlocality as an axiom”. *Found. Phys.* **24**, 379–385 (1994). url: doi.org/10.1007/BF02058098.
- [6] Wolfgang Bertram. “An Essay on the Completion of Quantum Theory. I: General Setting” (2017). [arXiv:1711.08643](https://arxiv.org/abs/1711.08643).
- [7] Wolfgang Bertram. “An Essay on the Completion of Quantum Theory. II: Unitary Time Evolution” (2018). [arXiv:1807.04650](https://arxiv.org/abs/1807.04650).
- [8] Lucien Hardy. “Quantum theory from five reasonable axioms” (2001). url: arxiv.org/abs/quant-ph/0101012.
- [9] Karol Życzkowski. “Quartic quantum theory: an extension of the standard quantum mechanics”. *J. Phys. A: Math. Theor.* **41**, 355302 (2008). url: doi.org/10.1088/1751-8113/41/35/355302.
- [10] Lee Smolin. “Could quantum mechanics be an approximation to another theory?” (2006). [arXiv:quant-ph/060910](https://arxiv.org/abs/quant-ph/060910).
- [11] Marshall H Stone. “On one-parameter unitary groups in hilbert space”. *Ann. Math.* **33**, 643–648 (1932). url: doi.org/10.2307/1968538.
- [12] Andrew M Gleason. “Measures on the closed subspaces of a hilbert space”. In *The Logico-Algebraic Approach to Quantum Mechanics*. Pages 123–133. Springer (1975).
- [13] Lluís Masanes, Thomas D Galley, and Markus P Müller. “The measurement postulates of quantum mechanics are operationally redundant”. *Nat. Comm.* **10**, 1–6 (2019).
- [14] Borivoje Dakic and Caslav Brukner. “Quantum theory and beyond: Is entanglement special?” (2009).
- [15] G. Chiribella, G. M. D’Ariano, and P. Perinotti. “Probabilistic theories with purification”. *Phys. Rev. A* **81**, 062348 (2010). [arXiv:0908.1583](https://arxiv.org/abs/0908.1583).
- [16] Lucien Hardy. “Reformulating and Reconstructing Quantum Theory” (2011). [arXiv:1104.2066](https://arxiv.org/abs/1104.2066).
- [17] Rob Clifton, Jeffrey Bub, and Hans Halvorson. “Characterizing quantum theory in terms of information-theoretic constraints”. *Foundations of Physics* **33**, 1561–1591 (2003).
- [18] Philip Goyal. “Information-geometric reconstruction of quantum theory”. *Phys. Rev. A* **78**, 052120 (2008).
- [19] Lluís Masanes and Markus P Müller. “A derivation of quantum theory from physical requirements”. *New J. Phys.* **13**, 063001 (2011).
- [20] Howard Barnum, Markus P Müller, and Cozmin Ududec. “Higher-order interference and single-system postulates characterizing quantum theory”. *New Journal of Physics* **16**, 123029 (2014).
- [21] Alexander Wilce. “A Royal Road to Quantum Theory (or Thereabouts)” (2016). [arXiv:1606.09306](https://arxiv.org/abs/1606.09306).

- [22] Philipp Höhn. “Quantum theory from rules on information acquisition”. *Entropy* **19**, 98 (2017).
- [23] Agung Budiyo and Daniel Rohrlich. “Quantum mechanics as classical statistical mechanics with an ontic extension and an epistemic restriction”. *Nature Communications* **8** (2017).
- [24] John H. Selby, Carlo Maria Scandolo, and Bob Coecke. “Reconstructing quantum theory from diagrammatic postulates”. *Quantum* **5**, 445 (2021).
- [25] Sean Tull. “A categorical reconstruction of quantum theory”. *Logical Methods in Computer Science ; Volume 16* Pages Issue 1 ; 1860–5974 (2020).
- [26] John van de Wetering. “An effect-theoretic reconstruction of quantum theory”. *Compositionality* **1**, 1 (2019).
- [27] Kenji Nakahira. “Derivation of quantum theory with superselection rules”. *Physical Review A* **101** (2020).
- [28] G. Chiribella, G. M. D’Ariano, and P. Perinotti. “Informational derivation of quantum theory”. *Phys. Rev. A* **84**, 012311 (2011). [arXiv:1011.6451](https://arxiv.org/abs/1011.6451).
- [29] Ciarán M Lee and John H Selby. “Generalised phase kick-back: the structure of computational algorithms from physical principles”. *New J. Phys.* **18**, 033023 (2016). url: doi.org/10.1088/1367-2630/18/3/033023.
- [30] Ciarán M Lee and John H Selby. “Deriving grover’s lower bound from simple physical principles”. *New J. Phys.* **18**, 093047 (2016).
- [31] Howard Barnum, Ciarán M Lee, and John H Selby. “Oracles and query lower bounds in generalised probabilistic theories”. *Found. Phys.* **48**, 954–981 (2018).
- [32] Giulio Chiribella, Giacomo Mauro D’Ariano, and Paolo Perinotti. “Probabilistic theories with purification”. *Physical Review A* **81**, 062348 (2010).
- [33] Jamie Sikora and John Selby. “Simple proof of the impossibility of bit commitment in generalized probabilistic theories using cone programming”. *Phys. Rev. A* **97**, 042302 (2018).
- [34] Giulio Chiribella and Carlo Maria Scandolo. “Entanglement and thermodynamics in general probabilistic theories”. *New J. Phys.* **17**, 103027 (2015).
- [35] Giulio Chiribella and Carlo Maria Scandolo. “Microcanonical thermodynamics in general physical theories”. *New J. Phys.* **19**, 123043 (2017).
- [36] Howard Barnum, Ciarán M Lee, Carlo Maria Scandolo, and John H Selby. “Ruling out higher-order interference from purity principles”. *Entropy* **19**, 253 (2017).
- [37] Ciarán M Lee and John H Selby. “A no-go theorem for theories that decohere to quantum mechanics”. *Proc. R. Soc. A: Math. Phys. Eng. Sci.* **474**, 20170732 (2018).
- [38] Roman V. Buny, Stephen D.H. Hsu, and A. Zee. “Is Hilbert space discrete?”. *Physics Letters B* **630**, 68–72 (2005).
- [39] Markus Mueller. “Does probability become fuzzy in small regions of spacetime?”. *Physics Letters B* **673**, 166–167 (2009).
- [40] T. N. Palmer. “Discretisation of the Bloch Sphere, Fractal Invariant Sets and Bell’s Theorem” (2020). [arXiv:1804.01734](https://arxiv.org/abs/1804.01734).
- [41] Bas Westerbaan and John van de Wetering. “A computer scientist’s reconstruction of quantum theory”. *J. Phys. A: Math. Theor.* **55**, 384002 (2022).
- [42] L. Hardy. “Probability theories with dynamic causal structure: a new framework for quantum gravity” (2005). [arXiv:gr-qc/0509120](https://arxiv.org/abs/gr-qc/0509120).
- [43] L. Hardy. “Towards quantum gravity: a framework for probabilistic theories with non-fixed causal structure”. *J. Phys. A* **40**, 3081–3099 (2007). [arXiv:gr-qc/0608043](https://arxiv.org/abs/gr-qc/0608043).
- [44] Ognjan Oreshkov, Fabio Costa, and Časlav Brukner. “Quantum correlations with no causal order”. *Nat. Comm.* **3**, 1–8 (2012).

- [45] Giulio Chiribella, Giacomo Mauro D’Ariano, Paolo Perinotti, and Benoit Valiron. “Quantum computations without definite causal structure”. *Phys. Rev. A* **88**, 022318 (2013).
- [46] Mateus Araújo, Adrien Feix, Miguel Navascués, and Časlav Brukner. “A purification postulate for quantum mechanics with indefinite causal order”. *Quantum* **1**, 10 (2017).
- [47] M. A. Nielsen and I. L. Chuang. “Quantum computation and quantum information”. *Cambridge University Press, Cambridge*. (2000).
- [48] J. Barrett. “Information processing in generalized probabilistic theories”. *Phys. Rev. A* **75**, 032304 (2007).
- [49] N. Brunner, D. Cavalcanti, S. Pironio, V. Scarani, and S. Wehner. “Bell nonlocality”. *Rev. Mod. Phys.* **86**, 839 (2014). [arXiv:quant-ph/1303.2849](https://arxiv.org/abs/1303.2849).
- [50] Howard Barnum, Oscar CO Dahlsten, Matthew Leifer, and Ben Toner. “Nonclassicality without entanglement enables bit commitment”. In 2008 IEEE Information Theory Workshop. *Pages 386–390*. IEEE (2008).
- [51] Marek Winczewski, Tamoghna Das, and Karol Horodecki. “Limitations on device independent secure key via squashed non-locality” (2019). [arXiv:1903.12154](https://arxiv.org/abs/1903.12154).
- [52] Martin Plávala. “General probabilistic theories: An introduction” (2021). [arXiv:2103.07469](https://arxiv.org/abs/2103.07469).
- [53] Markus Müller. “Probabilistic theories and reconstructions of quantum theory”. *SciPost Phys. Lect. Notes* **Page 28** (2021).
- [54] Ludovico Lami. “Non-classical correlations in quantum mechanics and beyond” (2018). [arXiv:1803.02902](https://arxiv.org/abs/1803.02902).
- [55] Bob Coecke. “Terminality implies non-signalling” (2014). url: arxiv.org/abs/1405.3681v3.
- [56] Aleks Kissinger, Matty Hoban, and Bob Coecke. “Equivalence of relativistic causal structure and process terminality” (2017). url: doi.org/10.48550/arXiv.1708.04118.
- [57] Stefano Gogioso and Carlo Maria Scandolo. “Categorical probabilistic theories” (2017). url: doi.org/10.4204/EPTCS.266.23.
- [58] C. Pfister and S. Wehner. “If no information gain implies no disturbance, then any discrete physical theory is classical”. *Nat. Comm.* **4**, 1851 (2013). url: doi.org/10.1038/ncomms2821.
- [59] Ł. Czekaj, M. Horodecki, P. Horodecki, and R. Horodecki. “Information content of systems as a physical principle”. *Phys. Rev. A* **95**, 022119 (2017).
- [60] P. Janotta, C. Gogolin, J. Barrett, and N. Brunner. “Limits on non-local correlations from the structure of the local state space”. *New J. Phys.* **13**, 063024 (2011).
- [61] Howard Barnum and Alexander Wilce. “Ordered linear spaces and categories as frameworks for information-processing characterizations of quantum and classical theory” (2009). [arXiv:0908.2354](https://arxiv.org/abs/0908.2354).
- [62] Peter Janotta and Raymond Lal. “Generalized probabilistic theories without the no-restriction hypothesis”. *Phys. Rev. A* **87**, 052131 (2013). url: doi.org/10.1103/PhysRevA.87.052131.
- [63] K. Kuratowski. “Introduction to set theory & topology”. *Volume 101 of International series of monographs in pure and applied mathematics*. PWN. Warsaw (1961).
- [64] Kenta Cho and Bart Jacobs. “Disintegration and bayesian inversion, both abstractly and concretely”. *Math. Struct. Comput. Sci.* **29**, 938–971 (2017). url: doi.org/10.1017/S0960129518000488.
- [65] Manuel Blum. “Coin flipping by telephone”. In *Advances in Cryptology: A Report on CRYPTO 81*, IEEE Workshop on Communications Security. *Pages 11–15*. (1981).
- [66] Shafi Goldwasser, Silvio Micali, and Charles Rackoff. “The knowledge complexity of interactive proof systems”. *SIAM J. Comput.* **18**, 186–208 (1989).
- [67] Dominic Mayers. “Unconditionally secure quantum bit commitment is impossible”. *Phys. Rev. Lett.* **78**, 3414–3417 (1997).

- [68] Hoi-Kwong Lo and Hoi Fung Chau. “Why quantum bit commitment and ideal quantum coin tossing are impossible”. *Physica D: Nonlinear Phenomena* **120**, 177–187 (1998).
- [69] Stephen Boyd and Lieven Vandenberghe. “Convex optimization”. *Cambridge University Press*. (2004).
- [70] Sevag Gharibian, Jamie Sikora, and Sarvagya Upadhyay. “QMA variants with polynomially many provers”. *Quantum Information & Computation* **13**, 0135–0157 (2013). [arXiv:1108.0617](https://arxiv.org/abs/1108.0617).
- [71] Somshubhro Bandyopadhyay, Alessandro Cosentino, Nathaniel Johnston, Vincent Russo, John Watrous, and Nengkun Yu. “Limitations on separable measurements by convex optimization”. *IEEE Transactions on Information Theory* **61**, 3593–3604 (2015). url: doi.org/10.1109/TIT.2015.2417755.
- [72] Monique Laurent and Teresa Piovesan. “Conic approach to quantum graph parameters using linear optimization over the completely positive semidefinite cone”. *Siam J. Optim.* **25**, 2461–2493 (2015). url: doi.org/10.1137/14097865X.
- [73] Ashwin Nayak, Jamie Sikora, and Levent Tunçel. “A search for quantum coin-flipping protocols using optimization techniques”. *Math. Program.* **156**, 581–613 (2016). url: doi.org/10.1007/s10107-015-0909-y.
- [74] Jamie Sikora and Antonios Varvitsiotis. “Linear conic formulations for two-party correlations and values of nonlocal games”. *Math. Program.* **162**, 431–463 (2017). url: doi.org/10.1007/s10107-016-1049-8.
- [75] Samuel Fiorini, Serge Massar, Manas K Patra, and Hans Raj Tiwary. “Generalized probabilistic theories and conic extensions of polytopes”. *J. Phys. A: Math. Theor.* **48**, 025302 (2014). url: doi.org/10.1088/1751-8113/48/2/025302.
- [76] Anna Jenčová and Martin Plávala. “Conditions on the existence of maximally incompatible two-outcome measurements in general probabilistic theory”. *Phys. Rev. A* **96**, 022113 (2017). url: doi.org/10.1103/PhysRevA.96.022113.
- [77] Joonwoo Bae, Dai-Gyoung Kim, and Leong-Chuan Kwek. “Structure of optimal state discrimination in generalized probabilistic theories”. *Entropy* **18**, 39 (2016). url: doi.org/10.3390/e18020039.
- [78] L. Lami, C. Palazuelos, and A. Winter. “Ultimate data hiding in quantum mechanics and beyond”. *Commun. Math. Phys.* **361**, 661–708 (2018).
- [79] Jamie Sikora and John H. Selby. “Impossibility of coin flipping in generalized probabilistic theories via discretizations of semi-infinite programs”. *Phys. Rev. Research* **2**, 043128 (2020).
- [80] John H Selby and Jamie Sikora. “How to make unforgeable money in generalised probabilistic theories”. *Quantum* **2**, 103 (2018). url: doi.org/10.22331/q-2018-11-02-103.
- [81] Bob Coecke, John Selby, and Sean Tull. “Two roads to classicality” (2017). url: doi.org/10.4204/EPTCS.266.7.
- [82] John Selby and Bob Coecke. “Leaks: quantum, classical, intermediate and more”. *Entropy* **19**, 174 (2017). url: doi.org/10.3390/e19040174.
- [83] M. Pawłowski, T. Paterek, D. Kaszlikowski, V. Scarani, A. Winter, and M. Żukowski. “Information causality as a physical principle”. *Nature* **461**, 1101–1104 (2009). url: doi.org/10.1038/nature08400.
- [84] J. Barrett. “Nonsequential positive-operator-valued measurements on entangled mixed states do not always violate a Bell inequality”. *Phys. Rev. A* **65**, 042302 (2002). url: doi.org/10.1103/PhysRevA.65.042302.
- [85] A. J. Short, S. Popescu, and N. Gisin. “Entanglement swapping for generalized nonlocal correlations”. *Phys. Rev. A* **73**, 012101 (2006). url: doi.org/10.1103/PhysRevA.73.012101.

- [86] C. H. Bennett, G. Brassard, C. Crépeau, R. Jozsa, A. Peres, and W. K. Wootters. “Teleporting an unknown quantum state via dual classical and einstein–podolsky–rosen channels”. *Phys. Rev. Lett.* **70**, 1895 (1993).
- [87] M. Żukowski, A. Zeilinger, M. A. Horne, and A. K. Ekert. “Event-ready detectors bell experiment via entanglement swapping”. *Phys. Rev. Lett.* **71**, 4287 (1993).
- [88] A. Acin, N. Brunner, N. Gisin, S. Massar, S. Pironio, and V. Scarani. “Device-independent security of quantum cryptography against collective attacks”. *Phys. Rev. Lett.* **98**, 230501 (2007). url: doi.org/10.1103/PhysRevLett.98.230501.
- [89] E. Hänggi, R. Renner, and S. Wolf. “Efficient quantum key distribution based solely on bell’s theorem”. *EUROCRYPT*Pages 216–234 (2010). [arXiv:org:0911.4171](https://arxiv.org/abs/0911.4171).
- [90] J. Barrett, L. Hardy, and A. Kent. “No signaling and quantum key distribution”. *Phys. Rev. Lett* **95**, 010503 (2005).
- [91] A. Acin, N. Gisin, and L. Masanes. “From bell’s theorem to secure quantum key distribution”. *Phys. Rev. Lett* **97**, 120405 (2006).
- [92] E. Hänggi. “Device-independent quantum key distribution”. PhD thesis. PhD Thesis, 2010. (2010).
- [93] R. Colbeck and R. Renner. “Free randomness can be amplified”. *Nat. Phys.* **8**, 450–454 (2012). url: doi.org/10.1038/nphys2300.
- [94] R. Gallego, L. Masanes, G. DeLaTorre, C. Dhara, L. Aolita, and A. Acin. “Full randomness from arbitrarily deterministic events”. *Nat. Comm.* **4**, 2654 (2013). url: doi.org/10.1038/ncomms3654.
- [95] P. Mironowicz, R. Gallego, and M. Pawłowski. “Amplification of arbitrarily weak randomness”. *Phys. Rev. A* **91**, 032317 (2015). url: doi.org/10.1103/PhysRevA.91.032317.
- [96] F. G. S. L. Brandão, R. Ramanathan, A. Grudka, K. Horodecki, P. Horodecki M. Horodecki, T. Szarek, and H. Wojewódka. “Robust device-independent randomness amplification with few devices”. *Nat. Comm.* **7**, 11345 (2016). url: doi.org/10.1038/ncomms11345.
- [97] R. Ramanathan, F. G. S. L. Brandão, K. Horodecki, M. Horodecki, P. Horodecki, and H. Wojewódka. “Randomness amplification against no-signaling adversaries using two devices”. *Phys. Rev. Lett.* **117**, 230501 (2016). url: doi.org/10.1103/PhysRevLett.117.230501.
- [98] H. Wojewódka, F. G. S. L. Brandão, A. Grudka, M. Horodecki, K. Horodecki, P. Horodecki, M. Pawłowski, and R. Ramanathan. “Randomness amplification against no-signaling adversaries using two devices”. *IEEE Trans. Inf. Theory* **63**, 7592 (2017). url: doi.org/10.1109/TIT.2017.2738010.
- [99] J. F. Clauser, M. A. Horne, A. Shimony, and R. A. Holt. “Proposed experiment to test local hidden-variable theories”. *Phys. Rev. Lett.* **23**, 880–884 (1969).
- [100] Marek Winczewski, Tamoghna Das, and Karol Horodecki. “Limitations on device independent secure key via squashed non-locality” (2020). [arXiv:1903.12154](https://arxiv.org/abs/1903.12154).
- [101] P. Horodecki and R. Ramanathan. “The relativistic causality versus no-signaling paradigm for multi-party correlations”. *Nat Commun* **10**, 1701 (2019).
- [102] J. Barrett, N. Linden, S. Massar, S. Pironio, S. Popescu, and D. Roberts. “Non-local correlations as an information theoretic resource”. *Phys. Rev. A* **71**, 022101 (2005).
- [103] R. Horodecki, P. Horodecki, M. Horodecki, and K. Horodecki. “Quantum entanglement”. *Rev. Mod. Phys.* **81**, 865 (2009). url: doi.org/10.1103/RevModPhys.81.865.
- [104] S. Pironio. “Lifting bell inequalities”. *Journal of Mathematical Physics* **46**, 062112 (2005). [arXiv:1210.0194](https://arxiv.org/abs/1210.0194).
- [105] A. Schrijver. “Combinatorial optimization polyhedra and efficiency”. Springer. Berlin (2003). url: link.springer.com/book/9783540443896.

- [106] C. Carathéodory. “Über den variabilitätsbereich der fourier’schen konstanten von positiven harmonischen funktionen”. Aus: *Rendiconti del Circolo Matematico di Palermo*. Direzione e Redazione. (1911). url: books.google.co.in/books?id=n4SkjwEACAAJ.
- [107] Günter M. Ziegler. “Lectures on polytopes”. *Springer New York*. (1995).
- [108] D. Collins, N. Gisin, N. Linden, S. Massar, and S. Popescu. “Bell inequalities for arbitrarily high-dimensional systems”. *Phys. Rev. Lett.* **88**, 040404 (2002). url: doi.org/10.1103/PhysRevLett.88.040404.
- [109] P. McMullen. “The maximum numbers of faces of a convex polytope”. *Mathematika* **17**, 179–184 (1970). arXiv:<https://londmathsoc.onlinelibrary.wiley.com/doi/pdf/10.1112/S0025579300002850>.
- [110] Khaled Elbassioni, Zvi Lotker, and Raimund Seidel. “Upper bound on the number of vertices of polyhedra with 0,1-constraint matrices”. *Information Processing Letters* **100**, 69 – 71 (2006).
- [111] Samson Abramsky and Adam Brandenburger. “The sheaf-theoretic structure of non-locality and contextuality”. *New J. Phys.* **13**, 113036 (2011). url: doi.org/10.1088/1367-2630/13/11/113036.
- [112] M. Araújo, M. Túlio Quintino, C. Budroni, M. Terra Cunha, and A. Cabello. “All noncontextuality inequalities for the n-cycle scenario”. *Phys. Rev. A* **88**, 022118 (2013). url: doi.org/10.1103/PhysRevA.88.022118.
- [113] Ernst Specker. “Die logik nicht gleichzeitig entscheidbarer aussagen”. In *Ernst Specker Selecta*. Pages 175–182. Springer (1990).
- [114] Yeong-Cherng Liang, Robert W Spekkens, and Howard M Wiseman. “Specker’s parable of the overprotective seer: A road to contextuality, nonlocality and complementarity”. *Phys. Rep.* **506**, 1–39 (2011). url: doi.org/10.1016/j.physrep.2011.05.001.
- [115] Ravi Kunjwal, Chris Heunen, and Tobias Fritz. “Quantum realization of arbitrary joint measurability structures”. *Phys. Rev. A* **89**, 052126 (2014). url: doi.org/10.1103/PhysRevA.88.022118.
- [116] B. Tsirelson. “Quantum generalizations of Bell’s inequality”. *Lett. Math. Phys.* **4**, 93–100 (1980). url: doi.org/10.1007/BF00417500.
- [117] A. Grudka, K. Horodecki, M. Horodecki, P. Horodecki, R. Horodecki, P. Joshi, W. Kłobus, and A. Wójcik. “Quantifying Contextuality”. *Phys. Rev. Lett.* **112**, 120401 (2014). url: doi.org/10.1103/PhysRevLett.112.120401.

A Proofs

A.1 Proof of Proposition 13

For simplicity in this proof we adopt the diagrammatic notation of [15] using the conventions of [24].

Consider some state s and some purification $\sigma^p \in \mathbf{Ext}_P[s]$ and an arbitrary extension $\Sigma \in \mathbf{Ext}[s]$. We want to demonstrate that we can achieve generation, which we will do by showing that there exists a transformation mapping σ^p to Σ , as Σ is an arbitrary extension we will therefore have demonstrated that σ^p is generating.

To start, let us consider another purification, $\Sigma^p \in \mathbf{Ext}_P[\sigma]$ this time of the state Σ . It is clear that, as this is pure and an extension of s then it is also a purification of s . We therefore have two purifications of s and so one may be tempted to connect them via the essential uniqueness property. However, this will not always be the case, to see this let us examine the systems more carefully. σ^p is a state on some purifying system B :

$$\begin{array}{c} \begin{array}{|c|} \hline A \\ \hline \end{array} \quad \begin{array}{|c|} \hline B \\ \hline \end{array} \\ \hline \sigma^p \end{array} \quad (27)$$

whilst, σ could have been an extension on some other system X and the purification of this, Σ^p could involve a third system Y :

$$\begin{array}{c} \begin{array}{|c|} \hline A \\ \hline \end{array} \quad \begin{array}{|c|} \hline X \\ \hline \end{array} \quad \begin{array}{|c|} \hline Y \\ \hline \end{array} \\ \hline \Sigma^p \end{array} \quad (28)$$

Unless we are in a highly contrived scenario whereby $B = X \otimes Y$ then we are not able to directly employ the essential uniqueness property. We therefore have to further extend our systems by composing with some extra pure states, ϕ and χ , as follows:

$$\begin{array}{c} \begin{array}{|c|} \hline A \\ \hline \end{array} \quad \begin{array}{|c|} \hline B \\ \hline \end{array} \\ \hline \sigma^p \end{array} \quad \begin{array}{c} \begin{array}{|c|} \hline X \\ \hline \end{array} \quad \begin{array}{|c|} \hline Y \\ \hline \end{array} \\ \hline \chi \end{array} \quad \text{and} \quad \begin{array}{c} \begin{array}{|c|} \hline A \\ \hline \end{array} \quad \begin{array}{|c|} \hline B \\ \hline \end{array} \quad \begin{array}{|c|} \hline X \\ \hline \end{array} \quad \begin{array}{|c|} \hline Y \\ \hline \end{array} \\ \hline \Sigma^p \end{array} \quad \begin{array}{c} \begin{array}{|c|} \hline \phi \\ \hline \end{array} \\ \hline \phi \end{array} \quad (29)$$

Then, thanks to our additional assumption that the parallel composite of pure states is pure, then these composite states define further purifications of the original states s . Now however, they have the same systems and hence we can use essential uniqueness to conclude that there is a reversible transformation T mapping between them, that is:

$$\begin{array}{c} \begin{array}{|c|} \hline A \\ \hline \end{array} \quad \begin{array}{|c|} \hline B \\ \hline \end{array} \quad \begin{array}{|c|} \hline X \\ \hline \end{array} \quad \begin{array}{|c|} \hline Y \\ \hline \end{array} \\ \hline T \\ \hline \begin{array}{|c|} \hline B \\ \hline \end{array} \quad \begin{array}{|c|} \hline X \\ \hline \end{array} \quad \begin{array}{|c|} \hline Y \\ \hline \end{array} \\ \hline \begin{array}{|c|} \hline \sigma^p \\ \hline \end{array} \quad \begin{array}{|c|} \hline \chi \\ \hline \end{array} \end{array} = \begin{array}{c} \begin{array}{|c|} \hline A \\ \hline \end{array} \quad \begin{array}{|c|} \hline B \\ \hline \end{array} \quad \begin{array}{|c|} \hline X \\ \hline \end{array} \quad \begin{array}{|c|} \hline Y \\ \hline \end{array} \\ \hline \Sigma^p \end{array} \quad \begin{array}{c} \begin{array}{|c|} \hline \phi \\ \hline \end{array} \\ \hline \phi \end{array} \quad (30)$$

Now, we can simply discard the B and Y systems on both sides of this equation to see that we can achieve generation:

$$\begin{array}{c} \begin{array}{|c|} \hline A \\ \hline \end{array} \quad \begin{array}{|c|} \hline X \\ \hline \end{array} \\ \hline \tilde{T} \\ \hline \begin{array}{|c|} \hline B \\ \hline \end{array} \\ \hline \sigma^p \end{array} := \begin{array}{c} \begin{array}{|c|} \hline A \\ \hline \end{array} \quad \begin{array}{|c|} \hline X \\ \hline \end{array} \\ \hline T \\ \hline \begin{array}{|c|} \hline B \\ \hline \end{array} \quad \begin{array}{|c|} \hline \chi \\ \hline \end{array} \\ \hline \sigma^p \quad \chi \end{array} = \begin{array}{c} \begin{array}{|c|} \hline A \\ \hline \end{array} \quad \begin{array}{|c|} \hline X \\ \hline \end{array} \\ \hline T \\ \hline \begin{array}{|c|} \hline B \\ \hline \end{array} \quad \begin{array}{|c|} \hline X \\ \hline \end{array} \quad \begin{array}{|c|} \hline Y \\ \hline \end{array} \\ \hline \sigma^p \quad \chi \end{array} \quad (31)$$

$$= \begin{array}{c} \begin{array}{|c|} \hline A \\ \hline \end{array} \quad \begin{array}{|c|} \hline X \\ \hline \end{array} \\ \hline \Sigma^p \end{array} \quad \begin{array}{c} \begin{array}{|c|} \hline \phi \\ \hline \end{array} \\ \hline \phi \end{array} = \begin{array}{c} \begin{array}{|c|} \hline A \\ \hline \end{array} \quad \begin{array}{|c|} \hline X \\ \hline \end{array} \\ \hline \Sigma^p \end{array} = \begin{array}{c} \begin{array}{|c|} \hline A \\ \hline \end{array} \quad \begin{array}{|c|} \hline X \\ \hline \end{array} \\ \hline \Sigma \end{array} \quad (32)$$

That is, there is a transformation \tilde{T} which maps σ^p to Σ which completes the proof.

A.2 Proof of Theorem 16

To see this, first note that the set of all PME's, S_P is a convex set as any two PME's of P , $\mathcal{E}_{pure}^1(P)$, $\mathcal{E}_{pure}^2(P) \in S_P$, their convex combination⁵ $\lambda \mathcal{E}_{pure}^1(P) + (1-\lambda) \mathcal{E}_{pure}^2(P) \in S_P$, $\forall \lambda \in (0, 1)$. Such a mixture of PME's is defined as follows: Suppose $\{(p_i, P_E^i)\}_i = \mathcal{E}_{pure}^1(P)$ and $\{(q_i, P_E^i)\}_i = \mathcal{E}_{pure}^2(P)$ are two PME's of the behaviour P , where $V(\mathcal{E}_{pure}^1)$ and $V(\mathcal{E}_{pure}^2)$ are the set of pure behaviours corresponds to the ensembles. Now define $V_{int} = V(\mathcal{E}_{pure}^1) \cap V(\mathcal{E}_{pure}^2)$, $V_1 = V(\mathcal{E}_{pure}^1) \setminus V(\mathcal{E}_{pure}^2)$ and $V_2 = V(\mathcal{E}_{pure}^2) \setminus V(\mathcal{E}_{pure}^1)$. Then the convex combination of $\mathcal{E}_{pure}^1(P)$ and $\mathcal{E}_{pure}^2(P)$ is defined as

$$\lambda \mathcal{E}_{pure}^1(P) + (1-\lambda) \mathcal{E}_{pure}^2(P) := \{(r_i, P_E^i)\}_{P_E^i \in V(\mathcal{E}_{pure}^1) \cup V(\mathcal{E}_{pure}^2)}, \quad (33)$$

where $r_i = \lambda p_i + (1-\lambda)q_i$, $\forall P_E^i \in V_{int}$, $r_i = \lambda p_i$, $\forall P_E^i \in V_1$ and $r_i = (1-\lambda)q_i$, $\forall P_E^i \in V_2$. Clearly

$$\sum_{i: P_E^i \in V(\mathcal{E}_{pure}^1) \cup V(\mathcal{E}_{pure}^2)} r_i P_E^i = \sum_{i: P_E^i \in V_{int}} (\lambda p_i + (1-\lambda)q_i) P_E^i + \sum_{i: P_E^i \in V_1} \lambda p_i P_E^i + \sum_{i: P_E^i \in V_2} (1-\lambda)q_i P_E^i \quad (34)$$

$$= \sum_{i: P_E^i \in V_{int} \cup V_1} \lambda p_i P_E^i + \sum_{i: P_E^i \in V_{int} \cup V_2} (1-\lambda)q_i P_E^i \quad (35)$$

$$= \lambda P + (1-\lambda)P = P. \quad (36)$$

Here we use the fact that $V_{int} \cup V_1 = V(\mathcal{E}_{pure}^1)$, $V_{int} \cup V_2 = V(\mathcal{E}_{pure}^2)$ and $\sum_{i: P_E^i \in V(\mathcal{E}_{pure}^1)} p_i P_E^i = P$ and $\sum_{i: P_E^i \in V(\mathcal{E}_{pure}^2)} q_i P_E^i = P$. The above equation proves that the convex combination of two PME's of a behaviour P is also an ensemble of P , and its members are all pure. Hence, S_P forms a convex set (as we have shown) with only finite number of vertices (as we show below), and all of them are minimal ensembles, as the following lemma proves.

Lemma 28 (Extremal \implies Minimal). *In the set of all pure members ensembles of the behaviour P , denoted by S_P , the ensembles that are extremal of S_P , are minimal.*

Proof. Suppose by contradiction, this is not true, i.e. there exists a PME, $\mathcal{E}_{pure}(P) = \{(p_i, P_E^i)\}_{i=1}^n$, with $p_i > 0$, $\forall i$, which is extremal in S_P , but is not minimal. Then, there must exist a proper subset $\mathcal{I} \subset V(\mathcal{E}_{pure})$ such that for all $P_E^j \in \mathcal{I}$, there is some other choices of probabilities $\{q_j\}$, which forms an ensemble $\mathcal{M}(P) = \{(q_j, P_E^j)\}_{P_E^j \in \mathcal{I}}$, and it is minimal. Let us now embed the distribution $\{q_j\}_{j \in \mathcal{I}}$ which has less than n elements, to obtain new but equivalent distribution with n elements, by letting $p'_i := q_i$, $\forall P_E^i \in \mathcal{I}$ and $p'_i := 0$, $\forall P_E^i \in V(\mathcal{E}_{pure}) \setminus \mathcal{I}$. Let us note that the minimal ensemble $\mathcal{M}(P)$, is now equivalent to the PME, $\{(p'_i, P_E^i)\}_{i=1}^n$.

Consider now an ensemble defined as:

$$\mathcal{N} = \left\{ \frac{p_i - pp'_i}{(1-p)}, P_E^i \right\}_{i=1}^n \equiv \{(r_i, P_E^i)\}_{i=1}^n, \quad (37)$$

where we define $p = \frac{p_{min}}{p_{max}}$ with $p_{min} = \min_i \{p_i : p_i > 0\}$ and $p_{max} = \max_i \{p'_i\}$. Let us first note, that \mathcal{N} is an ensemble of P . indeed, note that

$$\begin{aligned} P &= \sum_i p_i P_E^i = \sum_i (p_i - pp'_i) P_E^i + p \sum_i p'_i P_E^i, \\ &= (1-p) \sum_i r_i P_E^i + p \sum_i p'_i P_E^i, \end{aligned} \quad (38)$$

now by assumption $\sum_i p'_i P_E^i = P$, as $\{(p'_i, P_E^i)\}_{i=1}^n = \mathcal{M}(P)$. Thus:

$$P = (1-p) \sum_i r_i P_E^i + pP \quad (39)$$

which implies that $\sum_i r_i P_E^i = P$, i.e. \mathcal{N} is an ensemble of P , if $0 < p < 1$.

⁵Abstractly we can view PME's as probability distributions over the set of pure behaviours and hence is a convex set (in particular a simplex) the set S_P is then a subset of this simplex which we will show is closed under convex combinations.

We will argue now that the latter fact holds. Indeed, by definition, p is nonzero, resulting $p_{min} > 0$. To see $p < 1$ we will prove that $p_{min} < p'_{max}$. If $p_{min} \geq p'_{max}$, then $1 = \sum_i p_i \geq np_{min} \geq np'_{max} \geq n/|\mathcal{I}| \Rightarrow |\mathcal{I}| \geq n$, which is a contradiction as \mathcal{I} is a proper subset of $V(\mathcal{E}_{pure})$. Here we use the fact that $p'_{max} \geq 1/|\mathcal{I}|$. Hence, $p = \frac{p_{min}}{p'_{max}} < 1$ and \mathcal{N} is an ensemble of P , and can be denoted as $\mathcal{N}(P)$.

Now, we observe that by construction the PME

$$\mathcal{E}(P) = p\mathcal{M}(P) + (1-p)\mathcal{N}(P), \quad (40)$$

i.e., $\mathcal{E}(P)$ is a mixture of two ensembles, that are not equal to each other. This is a contradiction with assumed extremality of the ensemble $\mathcal{E}(P)$, since the mixture, as shown above, is non-trivial, and the assertion follows. \square

The above theorem proves that all the extremal points in S_P , are minimal ensembles, i.e., there is no extremal points in S_P other than the $\mathcal{M}(P)$. Now we will prove the converse, that no interior point from S_P is a minimal ensemble, i.e., all minimal ensembles are also extremal. To prove it, we will need the following lemma, interesting in its own right, as it characterizes minimal ensembles as those with a unique distribution:

Lemma 29. *The pure members ensemble $\mathcal{E}_{pure}(P) = \{(p_i, P_E^i)\}_i$ of a behaviour P is minimal, \mathcal{M} , iff the decomposition of this behaviour into the elements $\{p_i : p_i > 0\}$ is unique, given by corresponding probabilities p_i .*

Proof. The “if” direction is trivial: if the elements $\{p_i : p_i > 0\}$, of the decomposition of $\mathcal{E}_{pure}(P)$ are unique, then it is not possible to set any probability to zero. Hence, there is no proper subset $\mathcal{I} \subset V(\mathcal{E}_{pure})$, which forms an ensemble of P , with another choice of probabilities.

For the “only if” part, suppose $\mathcal{M}(P) = \{(p_i, P_E^i)\}_{i=1}^m$, is a minimal ensemble of P , we have to prove that the decomposition $\{p_i, p_i > 0\}$, is unique. Assume that the $\{p_i\}$ is not unique, but being a minimal ensemble it should follow $\sum_{i=1}^m p_i P_E^i = P$, or in other words the set of following linear equations

$$\sum_{i=1}^m a_{ki} y_i = c_k, \quad (41)$$

for some $k = 1, \dots, l'$, has solution in form $y_i = p_i$. Here c_k are the entries of the behaviour P , for the pair (a, x) , or in other words the probability of getting a , when the input is x , $P(a|x) = c_k$. Similarly the coefficients $\{a_{ki}\}$ are the same entries for the pair (a, x) of the pure behaviours $\{P_E^i(a|x) = a_{ki}\}$. As the behaviour P should follow some equality constraint, and due to some internal symmetry of it, not all l' equations in Eq. (41), are linearly independent. Here, by linear independence we mean $\sum_{k \neq k'} \lambda_k (\sum_{i=1}^m a_{ki} y_i - c_k) \neq \sum_{i=1}^m a_{k'i} y_i - c'_{k'}$, for some k, k' and λ_k . Suppose, there are only l linearly independent equations. (There is also a constraint on the $\{y_i\}$, that $\sum_{i=1}^m y_i = 1$ as they represent the probability of the ensemble, but we don't need to consider it separately, as the behaviour P is normalized, so Eq. (41), will take care of it.) Now the number of linearly independent equations and the number of variables can be in one of the three orders which we consider separately: 1) $l > m$, 2) $l < m$ 3) $l = m$. Notice first, that it can not be $l > m$, i.e., for l number of linear equation pertaining m number of variables. Otherwise there would be no solution of the set of equation:

$$\left\{ \sum_{i=1}^m a_{ki} y_i = c_k \right\}_{k=1}^l, \quad (42)$$

with variables y_i but we already have a solution, the initial one: $y_i = p_i > 0, \forall i$. On the other hand, if $l < m$, then one can always write down any set of l , $\{y_i\}_{i=1}^l$, as a linear functions of the remaining $(m-l)$ $\{y_j\}_{j=l+1}^m$. And in that case one can always set any one (or more) $y_i = 0$ for some i , which violates the condition of minimal ensembles. Hence we are left with $l = m$.

In this case, we have the same number of linearly independent equation as the number of variables, and in that case the matrix $A = [a_{ki}]$ is non-singular and invertible, which gives a unique solution of $y_i = p_i > 0$ for all i . \square

We can pass now to prove the extremality of minimal ensembles:

Lemma 30 (Minimal \implies Extremal). *For a behaviour P , all of its minimal ensembles $\mathcal{M}(P)$ are extremal in the set S_P of all ensembles of a behaviour P .*

Proof. Suppose by contradiction, that $\mathcal{M}(P)$ is not extremal. Then, there exist pure members ensembles $\mathcal{E}_1(P)$ and $\mathcal{E}_2(P)$ such that:

$$\mathcal{M}(P) = \lambda \mathcal{E}_1(P) + (1 - \lambda) \mathcal{E}_2(P) \quad (43)$$

for some $0 < \lambda < 1$. By the above equality, $V(\mathcal{E}_1) \subseteq V(\mathcal{M})$ and $V(\mathcal{E}_2) \subseteq V(\mathcal{M})$. But by minimality of \mathcal{M} , $V(\mathcal{E}_1)$ can not be proper subset of $V(\mathcal{M})$, as there are no weights that together with any proper subset of $V(\mathcal{M})$ form an ensemble of P . Thus $V(\mathcal{E}_1) = V(\mathcal{M})$ and for similar reason $V(\mathcal{E}_2) = V(\mathcal{M})$. It would mean, that there is an ensemble (let us focus on \mathcal{E}_1) which has different distribution, but the same set of members. It would mean that the distribution of \mathcal{M} is not unique: there is another one which together with the same set of members yields an ensemble of P . This however is not possible, since by Lemma 29, any minimal ensemble has unique distribution. This proves desired contradiction, hence the assertion follows. \square

From these two Lemmas 28 and 30, we obtain, that the set S_P of all pure member ensembles (PMEs) of P is a convex hull of the set of minimal ensembles $\mathcal{M}(P)$. And for any behaviour P , the set of minimal ensembles is finite, as there are finite number of pure behaviours⁶ and corresponding to Lemma 29 the decomposition of the p_i in minimal ensembles are unique, implies S_P forms a convex polytope.

A.3 Proof of Corollary 20

Proof. Note that according to the Definition 18, if $\mathcal{E}(P)_{AE}(ae|xz)$ is the NSEA of the given behaviour $P_A(a|x)$, then the only ensembles realized for different choices of input z of the extending party E are the minimal ensembles $\mathcal{M}(P)$. If there are $\{\mathcal{M}_i(P)\}_{i=1}^N$, N numbers of such minimal ensembles of P_A , then there should be $|\mathcal{Z}| = N$, number of distinct inputs in part of the extending system. Due to Lemma 28 all the extremal points in S_P are minimal ensembles. From the latter, one can generate any pure members ensemble $\mathcal{E}(P)$ by properly mixing the minimal ensembles by using appropriate distribution $\{\tilde{p}(k)\}_{k=1}^N$, with $\sum_{k=1}^N \tilde{p}(k) = 1$. For an arbitrary $\mathcal{E}(P) \in S_P$, the Lemma 28, certifies the existence of at least one such $\{\tilde{p}(k)\}_{k=1}^N$, which will generate it, hence, $\mathcal{E}(P) = \sum_{k=1}^N \tilde{p}(k) \mathcal{M}_k(P)$. Each $\mathcal{M}_k(P) = \{(p(e = i|z = k), P_{\mathbf{E}}^{ik}(a|x))\}_i$, has been obtained from $\mathcal{E}(P)_{AE}$, by setting the input $z = k$ of the extending party. If the inputs are now chosen probabilistically according to the distribution $\{\tilde{p}(k)\}_{k=1}^N$, and registering the output $e = i$, then

$$\mathcal{E}(P) = \sum_{k=1}^N \tilde{p}(k) \left\{ \left(p(e = i|z = k), P_{\mathbf{E}}^{ik}(a|x) \right) \right\} \quad (44)$$

$$= \left\{ \left(q(i), P_{\mathbf{E}}^i(a|x) \right) \right\}_{P_{\mathbf{E}}^i \in \bigcup_{k=1}^N V(\mathcal{M}_k)}, \quad (45)$$

where $q(i)$ is the probability of getting the pure behaviour $P_{\mathbf{E}}^i$, as given in Lemma 28. \square

A.4 Proof of Theorem 23

Lemma 31. (ACCESS \implies GENERATION) *Access to all ensembles implies access to arbitrary extensions of the extended system.*

⁶The cardinality of the set of minimal ensembles is bounded by the cardinality of the set of all subsets of pure behaviours which is finite if the set of pure behaviours is finite. See section 4.2 for an explicit tighter upper bound.

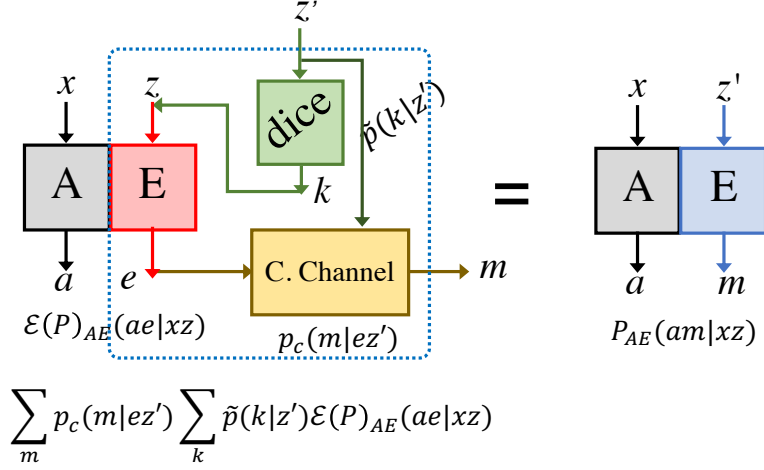


Figure 7: Pictorial depiction of Theorem 24. E holds the additional interfaces of the NSEA of behaviour P_A . She use an input randomizer (a dice) $\tilde{p}(k|z')$, which can be tuned by the parameter z' , and a classical post-processing channel $p_c(m|ez')$, which is also z' dependent. The dice, the extending system and the post-processing channel together form the new system in part of E . The input of the dice z' and the output of the channel m , can be considered as the input and output of the new system. E only choose some value of input z' , generates required randomness and also appropriate conditional probability distribution, resulting different set of mixed ensembles.

Proof. We first follow [102] (See Eq. (35) there) and observe that for any extension $P \equiv P(am|xz')$ of $P(a|x)$ there is

$$P(am|xz') = P(a|xz'm)P(m|xz') \quad (46)$$

$$= P(a|xz'm)P(m|z') \quad (47)$$

$$\equiv \{P^{ij}(a|x)P(m=i|z'=j)\}_{i,j} \quad (48)$$

In the first of the above equalities we use Bayes rule, and in the second the non-signalling from A to E . Such obtained equality implies that every bipartite behaviour can be viewed from perspective of the system E as having access to $|z'|$ ensembles of the form $\mathcal{E}_j(P) = \{(P(m=i|z'=j), P^{ij}(a|x))\}_i$. We argue now, that from NSEA one can generate any of these ensembles. Thanks to Theorem 22, for $|z'|$ inputs $z'=j$ there exists a dice D_j with probability distribution $\tilde{p}(k|z'=j)$ and a classical post-processing channel C_j , described with a conditional distribution $p_c(m|e, z'=j)$, such that when applied on system E of the NSEA $\mathcal{E}(P)(ae|xz)$ they generate the ensemble \mathcal{E}_j . Hence from a collection of $\{D_j\}_{j=1}^{|z'|}$ and $\{C_j\}_{j=1}^{|z'|}$ one can build via appropriate “wiring” a behaviour P' which upon input $z'=j$ performs according to a dice D_j and further post-process the output e (and that of D_j) through C_j to give the final output m (see Fig. 7). In this way one assures that a new behaviour P' has access to all the ensembles \mathcal{E}_j (and no other), forming (thanks to Eq. (48) above) an extension equivalent to P up to relabelling of inputs in the extending system E . \square

From the above theorem it is clear that the NSEA $\mathcal{E}(P)_{AE}$, of P_A together with access of arbitrary randomness (input randomizer followed by a classical post-processing channel) can generate any collection of ensembles (Theorem 22) even mixed. If we consider the whole setup as a single behaviour, as depicted in Fig. 7 (with the blue rectangular part is in possession of E), then it becomes a proper arbitrary extension generating a particular collection of ensembles. Note that it can be easily verified that it fulfills all properties of non-signalling behaviour.

Lemma 32. *GENERATION* \implies *ACCESS*

Proof. The proof for any GPT is given in Proposition 10. For the sake of completeness, we show it now using solely arguments from the non-signalling theory. Namely NSEA is a particular extension, hence access to arbitrary extension implies access to NSEA. This further via Theorem 22 implies access to any (possibly mixed) ensemble. \square

B Explicit examples in the theory of no-signaling behaviours

B.1 Complete extensions of binary input-output behaviours

In section 2.1, we proved that the theory of non-signaling behaviours (NS) does not possess the property of purification as there is not a pure extension of every behaviour. However, there do exist some behaviours for which the NSEA is an extremal, i.e., pure, behaviour. In other words, whilst the purification postulate fails as not every behaviour has a purification, there nonetheless do exist purifications of certain behaviours.

Here we are going to show an example of this. Namely, that if one considers a maximally mixed behaviour with a single binary input and a single binary output, then it can be extended to an extremal (pure) behaviour in a higher dimensional state space⁷ This pure behaviour is the maximally non-local behaviour equivalent to the Popescu-Rohrlich box [5] (up to proper labeling on the extending system) defined as

$$P^{\text{PR}}(a, e|x, z) = \begin{cases} \frac{1}{2} & \text{for } a \oplus e = x \cdot z, \\ 0 & \text{otherwise.} \end{cases} \quad (49)$$

where $a, e, x, z \in \{0, 1\}$. This is a close analogue to the fact that in quantum theory: a purification of the maximally mixed state $\frac{\mathbb{1}_d}{d}$ is the maximally entangled Bell state (up to local isometry) [103].

We will now prove this result. Note that the maximally mixed behaviour with a single binary input and single binary output is given by

$$P_A^m(a|x) = \begin{array}{c|cc} \swarrow \alpha^x & 0 & 1 \\ \hline 0 & 1/2 & 1/2 \\ \hline 1 & 1/2 & 1/2 \end{array} \quad (50)$$

here x being the input and a being the output of the behaviour on system A . This maximally mixed behaviour lies in the ‘‘center’’ of the polytope (of the set of behaviours with a single binary input and a single binary output), the extremal points (vertices) of the polytope are

$$P_E^0 = \begin{array}{c|cc} \swarrow \alpha^x & 0 & 1 \\ \hline 0 & 1 & 1 \\ \hline 1 & 0 & 0 \end{array}, \quad P_E^1 = \begin{array}{c|cc} \swarrow \alpha^x & 0 & 1 \\ \hline 0 & 0 & 0 \\ \hline 1 & 1 & 1 \end{array}, \quad P_E^2 = \begin{array}{c|cc} \swarrow \alpha^x & 0 & 1 \\ \hline 0 & 1 & 0 \\ \hline 1 & 0 & 1 \end{array}, \quad P_E^3 = \begin{array}{c|cc} \swarrow \alpha^x & 0 & 1 \\ \hline 0 & 0 & 1 \\ \hline 1 & 1 & 0 \end{array}. \quad (51)$$

These pure behaviours are deterministic behaviours as $P_E^i(a|x) = \delta_{a, g^i(x)}$, for some function $g^i : \{0, 1\} \rightarrow \{0, 1\}$. The polytope has been depicted in Fig. 8, with a blue square, the corners of the square are the extremal (pure) behaviours $\{P_E^i\}$, represented by the black solid circles. The center point of the polytope is the maximally mixed behaviour, P_A^m , which is depicted by a white solid circle. Any point inside the polytope can be expanded as a convex combination of the vertices. It is easy to see that, as P_A^m lies at the intersection of the two diagonals of the square (behaviour polytope), it can be expanded in terms of vertex pairs $\{P_E^0, P_E^1\}$ and $\{P_E^2, P_E^3\}$ with equal probabilities. In particular these form the two minimal ensembles of P_A^m namely

$$\mathcal{M}_0(P_A^m) = \{(1/2, P_E^0); (1/2, P_E^1)\} = \{p(i|0), P_E^{i0}\}, \quad (52)$$

$$\mathcal{M}_1(P_A^m) = \{(1/2, P_E^2); (1/2, P_E^3)\} = \{p(i|1), P_E^{i1}\}, \quad (53)$$

and there are no other minimal ensembles. Now from the Definition 18, of NSEA to the system E , the above two minimal ensembles are obtained in part of system A , for two different measurement choices on E . We choose that the first ensemble is obtained by setting the input $z = 0$, and the second one

⁷Throughout the paper we use the following notation for behaviours,

- P – *italic* represent any generic behaviour.
- P – normal font represent a particular example of a behaviour.

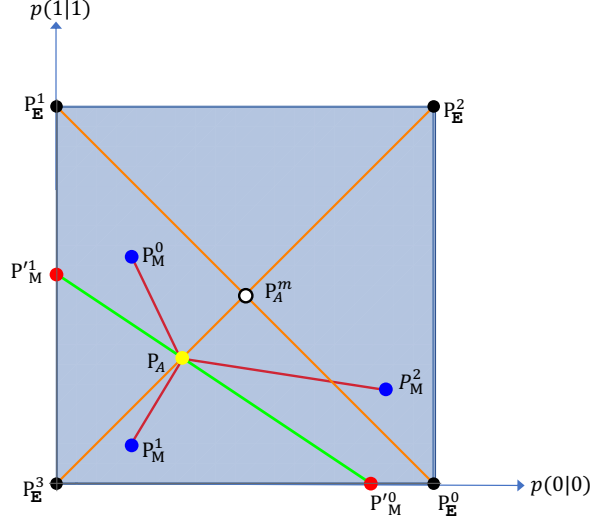


Figure 8: Polytope of the set of single party binary input output behaviours. Here the behaviour of consideration given in Eq. (56) is the yellow bullet. The white bullet is the maximally mixed behaviour. The black bullets are the deterministic, pure or extremal behaviours as given in Eq. (51). The required behaviour can be decomposed as linear combination of the deterministic behaviours as given in Eqs. (57) and (58). It can also be expanded as convex combination of mixed behaviours which are the red and blue bullets, these form the mixed ensemble of P_A .

is obtained by $z = 1$. Labeling the member behaviours of each ensemble with the output e , i.e.,

z	e	behaviour on A	Probability
0	0	P_E^0	1/2
	1	P_E^1	1/2
1	0	P_E^2	1/2
	1	P_E^3	1/2

(54)

we finally obtain the non-signaling complete extension (NSCE) of the maximally mixed behaviour that presents the same behaviour as the PR box in equation (49):

$$P_{AE}^{\text{PR}}(ae|xz) = \begin{array}{c|cc|cc} & x & 0 & 1 & & \\ & e & a & & & \\ \hline 0 & 0 & 1/2 & 0 & 1/2 & 0 \\ & 1 & 0 & 1/2 & 0 & 1/2 \\ \hline 1 & 0 & 1/2 & 0 & 0 & 1/2 \\ & 1 & 0 & 1/2 & 1/2 & 0 \end{array} \quad (55)$$

It is easy to check that if $x \cdot z = 0$, the resulting probability distribution on a and e is perfectly correlated, while for $x \cdot z = 1$ it is perfectly anti-correlated (i.e., $e = 0$ implies $a = 1$ and $e = 1$ implies $a = 0$). Thus in a sense we have derived a PR box solely from the principle of the complete extension (CEP). It is easy to see, that by negating z and e , one can instead obtain another maximally non-local bipartite behaviour. Similarly, all of the other maximally non-local behaviours (i.e., all non-local vertices of the polytope of two party binary input and binary output behaviours) can be obtained by proper relabeling of the z and e .

Note that the PR box is a vertex in the polytope of two party binary input and binary output behaviours [102]. Hence, we have the following conclusion:

Corollary 33. *The PR box is a purification of a maximally mixed behaviour with a single binary input and a single binary output.*

We have therefore constructed the PR box without any reference to CHSH inequality [99] (in contrast to how it was originally in [5]), as the non-signaling complete extension (NSCE) of a maximally

mixed behaviour (50). One could argue that the PR box is present in the theory of non-signaling behaviours from the very beginning. Whilst this is true, our derivation is based solely on the Definition 18, rather than relying on the structure of the polytope of non-signaling behaviours.

Remark 34. *It is tempting to say that the non-signaling complete extension (NSCE) of two party maximally mixed binary input and binary output behaviours is a tensor product of the PR boxes. It is however not the case. This is due to the fact, that one of the valid ensembles of a maximally mixed state $\frac{\mathbb{1}_{AB}}{4} = \frac{1}{2}PR_{AB} + \frac{1}{2}\overline{PR}_{AB}$, where $\overline{PR}_{AB}(ab|xy) = \frac{1}{2}\delta_{a\oplus b, xy\oplus 1}$ is a non-local behaviour supported on the orthogonal subspace to that of the support of PR. Since this ensemble is clearly minimal, having 2 members, in definition of NSCE there should be the input z which allows the owner of extending system to collapse the system AB into one of these maximally non-local behaviours (each with probability half). Suppose now, by contradiction that the NSCE is of the form $PR_{AX_A} \otimes PR_{BX_B}$. It is then clear to see, that in such a behaviour none of direct measurements (choosing the inputs) has outcome behaviour on AB of the form expected by measurement of demanded input z . However one should consider some other possible ways of measuring system $X_A X_B$ e.g. via wiring. Yet there is no such action on systems $X_A X_B$, simulating joint outcomes of z , since that would lead to the so called non-locality swapping, which is proven to be impossible in Refs. [84, 85].*

By virtue of Theorem 7, getting a pure behaviour in the higher dimensional state space through the construction of non-signaling complete extension (NSCE) is not always possible for any generic behaviour. Indeed, if we choose any other behaviour in the polytope of single binary input and single binary output, except for the maximally mixed and the four vertices behaviours, then its NSCE is not a vertex. For example, let us consider the following behaviour

$$P_A = \begin{array}{c|cc} & x & \\ \hline & 0 & 1 \\ \hline a & 0 & 1/3 & 2/3 \\ \hline & 1 & 2/3 & 1/3 \end{array} \quad (56)$$

which lies in the polytope of single party binary input and binary output behaviours, and it is represented by the yellow point in Fig. 8. Each behaviour can be expanded in terms of the pure behaviours of the polytope, hence

$$P_A = x_0 P_E^0 + x_1 P_E^1 + x_2 P_E^2 + x_3 P_E^3, \quad (57)$$

where $x_i \geq 0, \forall i$, and $\sum_{i=0}^3 x_i = 1$. The general solutions of Eq. (57) is:

$$\begin{cases} x_0 = \frac{1}{3}(2 - 3x_3), \\ x_1 = \frac{1}{3}(2 - 3x_3), \\ x_2 = \frac{1}{3}(3x_3 - 1), \\ x_3 = x_3 \end{cases}, \quad \frac{1}{3} \leq x_3 \leq \frac{2}{3} \quad (58)$$

To construct the minimal ensembles of behaviour P_A , we have to find out the set of decomposition over the pure points $\{P_E^i\}$, such that any proper subset of each choice can not be the ensemble of P_A with another set of probabilities. This implies that we have to find those solutions of Eq. (58), where the minimal number of x_i 's are nonzero. There are two of such choices, given by

$$\left\{ \begin{array}{l} x_0 = \frac{1}{3}, \\ x_1 = \frac{1}{3}, \\ x_2 = 0, \\ x_3 = \frac{1}{3} \end{array} \right\}, \quad \left\{ \begin{array}{l} x_0 = 0, \\ x_1 = 0, \\ x_2 = \frac{1}{3}, \\ x_3 = \frac{2}{3} \end{array} \right\}, \quad (59)$$

which, hence, form the two minimal ensembles of P_A :

$$\mathcal{M}_0(P_A) = \{(1/3, P_E^0); (1/3, P_E^1); (1/3, P_E^3)\} = \{p(i|0), P_E^{i0}\}, \quad (60)$$

$$\mathcal{M}_1(P_A) = \{(1/3, P_E^2); (2/3, P_E^3)\} = \{p(i|1), P_E^{i1}\}. \quad (61)$$

Here we label the ensembles by $\{0, 1\}$, according to the inputs z and the members as $P^{00} = P_E^0$, $P^{10} = P_E^1$, $P^{20} = P_E^3$ and $P^{01} = P_E^2$, $P^{11} = P_E^3$. Finally then, the NSCE of P_A to system E is given by

$$P_{AE}(ae|xz) = \begin{array}{c|cc|cc} & x & 0 & 1 & & \\ z & e^a & 0 & 1 & 0 & 1 \\ \hline 0 & 0 & 1/3 & 0 & 1/3 & 0 \\ & 1 & 0 & 1/3 & 0 & 1/3 \\ & 2 & 0 & 1/3 & 1/3 & 0 \\ \hline 1 & 0 & 1/3 & 0 & 0 & 1/3 \\ & 1 & 0 & 2/3 & 2/3 & 0 \end{array} \quad (62)$$

which lies in the polytope of two party behaviours, in which one party has a binary input and output, whereas the other party has a binary input but a ternary output. Moreover, as we will prove in the following section (Sec. B.1.1), this extended behaviour is not pure in the respective behaviour polytope.

An example of Theorem 20 has been given in Sec. B.1.2, where we have shown that any pure members ensemble of P_A , given in Eq. (56), can be constructed by taking the convex combination of its two minimal ensembles. This convex combination has been simulated by using an additional randomness (a dice) at the input of the extending party. An arbitrary ensemble of P_A , can be constructed by passing the output of the extending party through a classical post-processing channel (Theorem 22) has also been exemplified in Sec. B.1.3. The members of the mixed ensembles have been depicted by the blue dots (for mixed ensemble 1) and by red dots (for mixed ensemble 2). By enumerating the mixed ensembles with the input and the member behaviours with the output of an additional system, an extension of P_A has been generated. As the mixed ensembles of P_A are arbitrarily chosen, the extension made out of it is also an arbitrary extension of P_A – given in Eq. (79) of Sec. B.1.4. Moreover, for a generic behaviour, the non-signaling complete extension (NSCE) introduces correlations (possibly non-local) between the extending and the extended system. It is quite clear that, for a pure behaviour (a vertex of a polytope), there will be no correlation due to NSCE, whereas it can inject a maximal amount of correlation when the given behaviour is maximally mixed. For P_A , we calculate the non-local correlation in Sec. B.1.5.

Another important aspect of non-signaling complete extensions (NSCE) is that, unlike quantum purifications, the NSCE of an arbitrary behaviour and its conjugate behaviour are not the same. If $\mathcal{E}(P)_{AE}(ae|xz)$ is the NSCE of the behaviour $P_A(a|x)$, then the behaviour $P_E(e|z) = \sum_a P_{AE}(ae|xz)$, is the conjugate behaviour of P_A . The NSCE of $P_E(e|z)$, according to the Definition 18, $\mathcal{E}(P)_{A'E}(a'e|x'z)$ is not equal to the $\mathcal{E}(P)_{AE}(ae|xz)$. In section B.1.6, we provide an explicit example in favour of this argument, for the behaviour P_A , given in Eq. (56).

B.1.1 Example: the non-signaling complete extension that is not a vertex

In this section we will prove that the NSCE of P_A (given in equation (56)), is not a vertex in the higher dimensional behaviour polytope. Consider an arbitrary behaviour $P_{A_1 A_2 \dots A_n}(a_1 a_2 \dots a_n | x_1 x_2 \dots x_n)$, with n parties $A_1 A_2 \dots A_n$ where the x_i and a_i are respectively the input and output of the i th party A_i . Suppose there are m_i possible measurement choices of A_i , i.e., $x_i \in \{1, 2, \dots, m_i\}$, and corresponding to each measurement $x_i = j$ the number of possible outcomes is d_{ij} , such that $a_i \in \{1, 2, \dots, d_{ij}\}$. The total number of parameters involved in defining the behaviour is given by Ref. [104] as

$$t = \prod_{i=1}^n \left(\sum_{j=1}^{m_i} d_{ij} \right). \quad (63)$$

Among these t parameters not all of them are independent as the behaviour should obey normalization and non-signaling conditions. These conditions together imply that the behaviour must lie within a particular polytope of \mathbb{R}^t . If we construct a vector $v \in \mathbb{R}^t$, whose entries are those t probabilities, then the polytope can be defined as

$$\mathcal{P} = \{v \mid \mathcal{A}v \leq w\} \quad (64)$$

for some $w \in \mathbb{R}^s$, and \mathcal{A} an $t \times s$ matrix. Here the condition $\mathcal{A}u \leq w$ captures all of the constraints that the probabilities need to satisfy.

Let us take an arbitrary element $u \in \mathcal{P}$, and suppose $\mathcal{A}_u u \leq w_u$ are those inequality constraints among all possible constraints that are satisfied by u with equality [105], i.e., $\mathcal{A}_u u = w_u$. Here, \mathcal{A}_u and w_u are sub-matrices of \mathcal{A} and w respectively. Then from Theorem 5.7 of Ref. [105], u will be a pure point of \mathcal{P} if and only if $\text{rank}(\mathcal{A}_u) = t$.

The behaviour P_{AE} , given in Eq. (62), belongs to a polytope which lies in a space of $t = 20$. We now compute the rank of \mathcal{A}_u to demonstrate that this is not a vertex of the polytope. First we count how many independent equality constraints are acting on P_{AE} . The $t = 20$ probabilities are $P_{AE}(00|00), P_{AE}(01|00), \dots, P_{AE}(12|11)$, i.e., $P_{AE}(ae|xz)$ for $a, x \in \{0, 1\}$ and $e \in \{0, 1, 2\}$ for $z = 0$ and $e \in \{0, 1\}$ for $z = 1$. The equality constraints they need to satisfy are the non-signaling constraint and normalization conditions, namely

$$\sum_a P_{AE}(ae|xz) = \sum_a P_{AE}(ae|x'z), \quad \forall e, x, x', z, \quad (65)$$

$$\sum_e P_{AE}(ae|xz) = \sum_e P_{AE}(ae|xz'), \quad \forall a, x, z, z', \quad (66)$$

and

$$\sum_{ae} P_{AE}(ae|xz) = 1, \quad \forall x, z. \quad (67)$$

There are 5 from Eq. (65)⁸, 4 from Eq. (66)⁹, and another 4 from Eq. (67), totalling 13 equality constraints. Moreover, among the 40 inequality constraints of the form $0 \leq P_{AE}(ae|xz) \leq 1$, $\forall a, e, x, z$, (2 inequalities for each of the 20 probabilities $P_{AE}(ae|xz)$), only 10 of them find equality with zeroes. Note that there is no probability which finds equality with 1, as there is no entry of 1, in Eq. (62). Hence, the total number of equality constraint are 23, and the matrix \mathcal{A}_u for which $\mathcal{A}_u u = w_u$, is of the dimension 23×20 . To obtain $\text{rank}(\mathcal{A}_u)$, we need to find out the number of linearly independent equality constraint out of the 23. The 10 equality constraint with zero, on the probabilities are all linearly independent, but among those 13 (non-signaling constraint and normalization conditions) only 9 of them are linearly independent – by choosing any 1 normalization condition, and any 8 non-signaling condition one can generate the remaining 4 conditions. Hence, the total number of the linearly independent constraints, on the behaviour P_{AE} is $9 + 10 = 19$, and $\text{rank}(\mathcal{A}_u) = 19 < 20$. Which allows us to state that behaviour P_{AE} is not a vertex (extreme point) of the given polytope.

On the other hand the NSCE of the maximally mixed behaviour, the PR box given in Eq. (55) is a pure behaviour in the polytope of two binary input output behaviours. This can be shown in the following way: the space where the PR box lives is of $t = 16$ [102]. Amongst the 16 parameters there are 8 equality constraints with zeroes and $4 + 4 = 8$ linearly independent equality comes from the 8 non-signaling conditions and 4 normalization condition, hence $\text{rank}(\mathcal{A}_u) = 16$, which exactly matches with the dimension of the space.

In the following lines we study particular properties of NSEA in a low dimensional case.

Observation 35. *All non-deterministic behaviours in single party, binary input, binary output scenario have two minimal ensembles. All minimal ensembles of those have either two or three members.*

Proof. The polytope of single party binary input binary output behaviour belongs to set of reals of dimension $d = 2$ and it has been given in Fig. 8. Suppose, P is any arbitrary behaviour in the polytope, then from the theorem of Carathéodory [106, 107], the maximal number of pure behaviours in each minimal ensembles of P is 3. It is also clear from the figure that any 3 pure behaviours of the polytope form a triangle, and any arbitrary point inside the polytope can be inside at most two overlapping triangles. Hence, it has at most two minimal ensembles. \square

⁸ $\sum_a P_{AE}(ae|xz) = P_E(e|z)$, $e \in \{0, 1, 2\}$ for $z = 0$ and $e \in \{0, 1\}$ for $z = 1$.

⁹ $\sum_e P_{AE}(ae|xz) = P_A(a|x)$, $a, x \in \{0, 1\}$

For any arbitrary behaviour inside that polytope, the minimal ensemble consists of any combination of the following number of members.

Corollary 36. *Among single party, binary input, binary output behaviours, only five of them have NSCE that is a purification (NSCE which is a vertex).*

Proof. Due to the Observation 35 we know that the non-signaling complete extensions of behaviours in considered scenario are bipartite states in one of the following polytopes. Polytope of

- (i) one binary input and one unary input with binary outputs: behaviours lying on the edges of the polytope has single minimal ensembles with only two members.
- (ii) two binary inputs, two binary outputs behaviours: when the initial behaviour has two minimal ensembles and each of the ensembles has two members.
- (iii) two binary inputs, one binary output, one binary/ternary output (depending on the corresponding input setting) behaviours: behaviours lying on any one diagonal of the polytope.
- (iv) two binary inputs, and two ternary outputs behaviours: for any arbitrary behaviour, not belongs to the above sets.

For each local vertex (deterministic box) of the listed polytopes, if we trace out the second party by summation over all of its outcomes, the result is one of the deterministic behaviours of the initial polytope. Due to Theorem 1 of [102], we know the form of all non-local vertices in (i,ii,iii) polytopes. In each case after tracing out the extending system (summation over outcomes), the result is the maximally mixed behaviour of the initial system.

As we have investigated all the vertices which were suspected of being a purifications of behaviours from single party, one binary input, one binary output scenario and in each case we obtained one of the five states we conclude there are no other behaviours (in the initial polytope) that have purification. \square

B.1.2 Example: NSCE of P_A gives ACCESS to any PME of P_A

In Theorem 20, we state that, the extended system of the NSCE, can access any PMEs of the behaviour P_A , if it is equipped with arbitrary randomness. Any pure ensemble of P_A , $\mathcal{E}(P_A) = \{x_i, P_i\}_i$, where the $\{x_i\}$ satisfy Eq. (58), can be written down as convex combination of the minimal ensembles, which is given below

$$\mathcal{E}(P_A) = \lambda \mathcal{M}_0(P_A) + (1 - \lambda) \mathcal{M}_1(P_A), \quad (68)$$

with $\lambda = 2 - 3x_3 \in [0, 1]$, as $x_3 \in [\frac{1}{3}, \frac{2}{3}]$. If the extending party X chooses to toss a coin p_t , (binary output) and feed it to the input z of her part of the completely extended behaviour with $p_t(0) = \lambda$, and $p_t(1) = 1 - \lambda$, then the extending system has ACCESS to any pure ensemble of P_A .

B.1.3 Example: NSCE of P_A gives ACCESS to any mixed ensemble of P_A

In this section, we will explicitly exemplify that the extending system E can access all possible mixed ensemble $\mathcal{E}_{mix}(P_A) = \{p_m, P_M^m\}_m$, of an arbitrary behaviour P_A , (given in Theorem 22). Here the behaviours P_M^m are any arbitrary behaviours.

Example 1: Suppose X wants to access the following ensemble of mixed behaviours

$$\mathcal{E}_{mix}(P_A) = \left\{ \left(\frac{33}{81}, P_M^0 \right); \left(\frac{32}{81}, P_M^1 \right); \left(\frac{16}{81}, P_M^2 \right) \right\}, \quad (69)$$

in part of system A^{10} , where the mixed behaviours (the blue points in Fig. 8), are given by

$$P_M^0 = \begin{array}{c|cc} \diagdown & 0 & 1 \\ \hline 0 & 1/5 & 2/5 \\ \hline 1 & 4/5 & 3/5 \end{array}, \quad P_M^1 = \begin{array}{c|cc} \diagdown & 0 & 1 \\ \hline 0 & 1/5 & 9/10 \\ \hline 1 & 4/5 & 1/10 \end{array}, \quad P_M^2 = \begin{array}{c|cc} \diagdown & 0 & 1 \\ \hline 0 & 7/8 & 3/4 \\ \hline 1 & 1/8 & 1/4 \end{array}. \quad (70)$$

¹⁰Here we use *tilde*, on the symbol of ensemble to denote a particular ensemble among the set of ensembles.

Each of these mixed behaviours has some decompositions over the pure behaviours, which are certainly not unique, consider the following minimal one, which are

$$\mathcal{M}(P_M^0) = \left\{ \left(\frac{2}{5}, P_E^1 \right); \left(\frac{1}{5}, P_E^2 \right); \left(\frac{2}{5}, P_E^3 \right) \right\}, \quad (71)$$

$$\mathcal{M}(P_M^1) = \left\{ \left(\frac{1}{10}, P_E^0 \right); \left(\frac{1}{10}, P_E^2 \right); \left(\frac{4}{5}, P_E^3 \right) \right\}, \quad (72)$$

and

$$\mathcal{M}(P_M^2) = \left\{ \left(\frac{3}{4}, P_E^0 \right); \left(\frac{1}{8}, P_E^1 \right); \left(\frac{1}{8}, P_E^2 \right) \right\}, \quad (73)$$

Put them into Eq. (69), the mixed ensemble then turn out to be the pure one, given by

$$\mathcal{E}_{pure}(P_A) = \left\{ \left(\frac{76}{405}, P_E^0 \right); \left(\frac{76}{405}, P_E^1 \right); \left(\frac{59}{405}, P_E^2 \right); \left(\frac{194}{405}, P_E^3 \right) \right\} = \{r(e), P_E^e\} \quad (74)$$

One can check that $\mathcal{E}_{pure}(P_A) = \frac{76}{135} \mathcal{M}_0(P_A) + \frac{59}{135} \mathcal{M}_1(P_A)$, and E can access $\mathcal{E}_{pure}(P_A)$ by choosing the input according to the probability distribution, $\{p_t(0) = \frac{76}{135}, p_t(1) = \frac{59}{135}\}$. This can be done by feeding the output of a flipped coin to the input z of her part of the NSCE, $\mathcal{E}(P)_{AE}(ae|xz)$.

Once the PME in part of A has been prepared, the prefixed mixed ensembles has been constructed by passing the output e through a classical channel (post-processing channel) $P_c(m|e)$, which is

$m \backslash e$	0	1	2	3
0	0	33/38	33/59	33/97
1	4/19	0	16/59	64/97
2	15/19	5/38	10/59	0

$$P_c(m|e) = \quad (75)$$

the index m is the flag in part of E , different m give the access to different mixed behaviour P_M^m with probability p_m . One can check that $p_m = \sum_e P_c(m|e)r(e)$, and $P_M^m = \frac{1}{p_m} \sum_e P_c(m|e)r(e)P_E^e$. Thus we can see that the extending system can be able to access any ensemble of behaviour P_A , by NSCE with arbitrary randomness which will mix the minimal ensembles by mixing the input z , and then gluing the output e by a conditional classical channel.

Example 2: P_A can also be expanded as another mixed ensemble $\mathcal{E}'_{mix}(P_A) = \left\{ \left(\frac{2}{5}, P_M^0 \right); \left(\frac{3}{5}, P_M^1 \right) \right\}$, (the red points in Fig. 8) where

$$P_M^0 = \frac{5/6}{1/6} \Big| \frac{1}{0} = \left\{ \left(\frac{5}{6}, P_E^0 \right); \left(\frac{1}{6}, P_E^3 \right) \right\}, \quad (76)$$

$$P_M^1 = \frac{0}{1} \Big| \frac{4/9}{5/9} = \left\{ \left(\frac{5}{9}, P_E^1 \right); \left(\frac{4}{9}, P_E^3 \right) \right\}. \quad (77)$$

Now the pure ensemble turn out to $\mathcal{E}'(P_A) = \mathcal{M}_0(P_A) = \{(1/3, P_E^0); (1/3, P_E^1); (1/3, P_E^3)\} = \{r'(e), P_E^e\}$. For this, E will chose a completely biased coin, $p_t(0) = 1, p_t(1) = 0$, and feed its output to the input of the NSCE, and the post-processing channel P_c is

$m \backslash e$	0	1	2	3
0	1	0	-	1/5
1	0	1	-	4/5

$$P_c(m|e) = \quad (78)$$

which will be used to post-process the output of the extending part. Here we keep the column for $e = 2$ ‘‘blank’’ as there is no such incidence that the pure behaviour P_E^2 occur. Clearly $p_0 = \sum_e r'(e)P_c(m|e) = \frac{1}{3} + \frac{1}{3} \times \frac{1}{5} = \frac{2}{5}$ and $P_M^0 = \frac{1}{p_0} \sum_e r'(e)P_c(m=0|e)P_E^e = \frac{5}{6}P_E^0 + \frac{1}{6}P_E^3$.

B.1.4 NSCE can generate any extension

(Example of Theorem 24). Numbering these two examples of mixed ensembles with $z' = 0$ and $z' = 1$, we obtain an arbitrary extension of P_A to the behaviour $P_{AE}(am|xz')$. Such that $\{p(m|z' = 0), P^{m0}(a|x)\} = \mathcal{E}_{mix}(P_A)$ and $\{p(m|z' = 1), P^{m1}(a|x)\} = \mathcal{E}'_{mix}(P_A)$. And the arbitrary extended behaviour is

$$P_{AE}^{\text{mix}}(am|xz') = \begin{array}{c|cc|cc} & x & & & & & \\ & a & & & & & \\ z' \backslash m & & 0 & 1 & | & 0 & 1 \\ \hline 0 & 0 & \frac{11}{135} & \frac{44}{135} & | & \frac{22}{135} & \frac{11}{45} \\ & 1 & \frac{32}{405} & \frac{128}{405} & | & \frac{16}{45} & \frac{16}{405} \\ & 2 & \frac{14}{81} & \frac{2}{81} & | & \frac{4}{27} & \frac{4}{81} \\ \hline 1 & 0 & \frac{1}{3} & \frac{1}{15} & | & \frac{2}{5} & 0 \\ & 1 & 0 & \frac{3}{5} & | & \frac{4}{15} & \frac{1}{3} \end{array} \quad (79)$$

We can consider all possible extension of $P_A \rightarrow P_{AE}$, which will take care of all possible ensembles of P_A .

B.1.5 Quantifying non-locality introduced in NSCE

Here we quantify the amount of non-locality introduced among the extending and the extended system in the process of the construction of the NSCE, following Definition 18.

We have observed the fact that the completely extended behaviour of the maximally mixed single input output box, has turned out to be the Popescu-Rohrlich box [5], which (under suitable pre and post processing) can violate any kind of bipartite Bell expression maximally. In that case, the maximal amount of non-locality was introduced in the process of the construction of the NSCE. On the other hand, to quantify the non-locality of the NSCE of the non-maximally mixed single input output behaviour given in Eq. (56), the NSCE is shown in Eq. (62), to have different cardinalities of outputs. To get rid of this asymmetry in the extending system of X , we can do two possible surgeries.

Case 1: One can add one more outputs in the purified system and calculate the Bell like inequality defined by Collins et. al. (CGLMP) [108] (the acronym is after it finds D. Collins, N. Gisin, N. Linden, S. Massar and S. Popescu). The behaviour which maximizes the CGLMP bound has the following form after a local relabeling of the inputs and outputs

$$P_{AE}(ae|xz) = \begin{array}{c|ccc|ccc} & x & & & & & \\ & a & & & & & \\ z \backslash e & & 0 & 1 & 2 & | & 0 & 1 & 2 \\ \hline 0 & 0 & 1/3 & 0 & 0 & | & 0 & 1/3 & 0 \\ & 1 & 0 & 2/3 & 0 & | & 2/3 & 0 & 0 \\ & 2 & 0 & 0 & 0 & | & 0 & 0 & 0 \\ \hline 1 & 0 & 1/3 & 0 & 0 & | & 1/3 & 0 & 0 \\ & 1 & 0 & 1/3 & 0 & | & 0 & 1/3 & 0 \\ & 2 & 0 & 1/3 & 0 & | & 1/3 & 0 & 0 \end{array} \quad (80)$$

For this bipartite two inputs and three-output box, the CGLMP bound turns out to be 3, which is beyond the quantum limit quoted to be 2.87 in Ref. [108].

Case 2: Another way to calculate the non-locality of this asymmetric behaviour by following the prescription giving in Ref. [104]. It proposes to merge the extra outcomes in the following way

$$P'_{AE}(ae = 1|xz) = P_{AE}(ae = 1|xz) + P_{AE}(ae = 2|xz) \quad (81)$$

Hence the behaviour in Eq. (62) can be transformed to a bipartite binary input output box,

$$P'_{AE}(ae|xz) = \begin{array}{c|cc|cc} & x & 0 & 1 & & \\ z & \swarrow e & a & & & \\ \hline 0 & 0 & 1/3 & 0 & 1/3 & 0 \\ & 1 & 0 & 2/3 & 1/3 & 1/3 \\ \hline 1 & 0 & 1/3 & 0 & 0 & 1/3 \\ & 1 & 0 & 2/3 & 2/3 & 0 \end{array} \quad (82)$$

For this behaviour we have the well known CHSH inequality to quantify the non-locality, and it is 3.33, which is also beyond the quantum limit. However, the amount of non-locality for this NSCE is substantially less than the amount of non-locality present in a PR box. It therefore seems that it may be possible to quantify the non-locality between subsystem and its extending system as a measure of how close a behaviour is to being a vertex for NSCE in the theory of non-signaling behaviours.

Until now, we have given examples in favor of the various properties of NSCE we have discovered so far. Now we want to shed some light on another aspects of NSCE which shows a sharp disparity with the purification principle of the QT. If $|\psi_{AE}\rangle$, is the purification of a quantum state ρ_A , to system E , then the same pure state is also the purification of quantum state $\rho_X = \text{tr}_A|\psi_{AE}\rangle\langle\psi_{AE}|$. In the latter section we give an example to show that this is not the case for the NSCE. If we have a behaviour P_A , and P_{AE} is its NSCE, then we say the behaviour $P_E = \text{tr}_A P_{AE}$ is the conjugate box. We are going now to construct the NSCE of the conjugate box.

B.1.6 Complete extension of the conjugate box

In this section, we will find the NSCE of the conjugate behaviour of the behaviour given in Eq. (56). The conjugate box, P_E can be obtained from Eq. (62), by $P_E(e|z) = \sum_a P_{AE}(ae|xz)$, and it is given by

$$P_E(e|z) = \begin{array}{c|cc|c} & z & 0 & 1 \\ \hline e & \swarrow & & \\ \hline 0 & 0 & 1/3 & 1/3 \\ & 1 & 1/3 & 2/3 \\ \hline 2 & 2 & 1/3 & 0 \end{array} \quad (83)$$

This behaviour lies in a 4 dimensional behaviour polytope whose vertices are given by

$$\begin{array}{ccc} P_E^0 = \frac{\begin{array}{c|cc|c} & z & 0 & 1 \\ \hline e & \swarrow & & \\ \hline 0 & 0 & 1 & 1 \\ & 1 & 0 & 0 \\ \hline 2 & 2 & 0 & 0 \end{array}}{2}, & P_E^1 = \frac{\begin{array}{c|cc|c} & z & 0 & 1 \\ \hline e & \swarrow & & \\ \hline 0 & 0 & 0 & 0 \\ & 1 & 1 & 1 \\ \hline 2 & 2 & 0 & 0 \end{array}}{2}, & P_E^2 = \frac{\begin{array}{c|cc|c} & z & 0 & 1 \\ \hline e & \swarrow & & \\ \hline 0 & 0 & 0 & 0 \\ & 1 & 0 & 0 \\ \hline 2 & 2 & 1 & 1 \end{array}}{2}, \\ P_E^3 = \frac{\begin{array}{c|cc|c} & z & 0 & 1 \\ \hline e & \swarrow & & \\ \hline 0 & 0 & 1 & 0 \\ & 1 & 0 & 1 \\ \hline 2 & 2 & 0 & 0 \end{array}}{2}, & P_E^4 = \frac{\begin{array}{c|cc|c} & z & 0 & 1 \\ \hline e & \swarrow & & \\ \hline 0 & 0 & 1 & 0 \\ & 1 & 0 & 0 \\ \hline 2 & 2 & 0 & 1 \end{array}}{2}, & P_E^5 = \frac{\begin{array}{c|cc|c} & z & 0 & 1 \\ \hline e & \swarrow & & \\ \hline 0 & 0 & 0 & 1 \\ & 1 & 1 & 0 \\ \hline 2 & 2 & 0 & 0 \end{array}}{2}, \\ P_E^6 = \frac{\begin{array}{c|cc|c} & z & 0 & 1 \\ \hline e & \swarrow & & \\ \hline 0 & 0 & 0 & 1 \\ & 1 & 0 & 0 \\ \hline 2 & 2 & 1 & 0 \end{array}}{2}, & P_E^7 = \frac{\begin{array}{c|cc|c} & z & 0 & 1 \\ \hline e & \swarrow & & \\ \hline 0 & 0 & 0 & 0 \\ & 1 & 0 & 1 \\ \hline 2 & 2 & 1 & 0 \end{array}}{2}, & P_E^8 = \frac{\begin{array}{c|cc|c} & z & 0 & 1 \\ \hline e & \swarrow & & \\ \hline 0 & 0 & 0 & 0 \\ & 1 & 1 & 0 \\ \hline 2 & 2 & 0 & 1 \end{array}}{2}. \end{array} \quad (84)$$

To obtain the NSCE of this box, we need to find the minimal ensembles of P_E , which are

$$\mathcal{M}_0(P_E) = \{(1/3, P_E^0); (1/3, P_E^1); (1/3, P_E^7)\}, \quad (85)$$

$$\mathcal{M}_1(P_E) = \{(1/3, P_E^1); (1/3, P_E^3); (1/3, P_E^6)\}, \quad (86)$$

$$\mathcal{M}_2(P_E) = \{(1/3, P_E^3); (1/3, P_E^5); (1/3, P_E^7)\}. \quad (87)$$

Consider the NSCE of P_E to a system A' , as $P_{A'E}(a'e|x'z)$, where $\{p(a' = i|x' = k), P_E^{ik}(e|z)\} = \mathcal{M}_k$, the NSCE of P_E is

$$P_{A'E}(a'e|x'z) = \begin{array}{c} z \\ \begin{array}{c|ccc|ccc|ccc} & \begin{array}{c} x' \\ e \end{array} & 0 & 1 & 2 & 0 & 1 & 2 & 0 & 1 & 2 \\ \hline 0 & 0 & \frac{1}{3} & 0 & 0 & 0 & \frac{1}{3} & 0 & 0 & \frac{1}{3} & 0 & 0 \\ 1 & 1 & 0 & \frac{1}{3} & 0 & \frac{1}{3} & 0 & 0 & 0 & 0 & \frac{1}{3} & 0 \\ 2 & 2 & 0 & 0 & \frac{1}{3} & 0 & 0 & \frac{1}{3} & 0 & 0 & 0 & \frac{1}{3} \\ \hline 1 & 0 & \frac{1}{3} & 0 & 0 & 0 & 0 & \frac{1}{3} & 0 & 0 & \frac{1}{3} & 0 \\ 1 & 1 & 0 & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} & 0 & \frac{1}{3} & 0 & \frac{1}{3} \end{array} \end{array} \quad (88)$$

One can clearly see that $P_{AE} \neq P_{A'E}$, for example, by noticing the mismatch of the cardinality of the inputs of the extended party of the conjugate system.

B.2 Dimensionality of the No-signaling complete extension

In this subsection we provide a proof for Theorem 25 of the main text.

Proof. Let us recall, that in quantum theory (due to the Schmidt decomposition) purifications can always be found with an extending system with the same dimension as the extended system. We are interested in similarly quantifying the size of the minimal NSCE. However, there are many possible quantifiers that we could use here, for example, we can measure the size of the NSCE using the total number of its outputs, inputs and outputs, or, the dimension of its state space (vector space). The theory of non-signalling behaviours is a discrete convex theory, hence we call its state space a convex polytope, a multidimensional generalization of a convex polyhedron. In this section, we strive to upper bound the dimension of the polytope that contains NSEA of some fixed but arbitrary behaviour P . We identify the dimension of behaviour with the dimension of a polytope it belongs. Although we chose the dimension of NSEA, as a quantifier of its size, other aspects of it, like the number of its inputs and outputs, will also be discussed in the following.

From Theorem 1 of [104] we know that the dimension of a certain behaviour polytope \mathcal{B} is:

$$\dim \mathcal{B} = \prod_{i=1}^n \left(\sum_{j=1}^{m_i} (v_{ij} - 1) + 1 \right) - 1, \quad (89)$$

where n is the total number of non-signalling parties, m_i is the total number of the inputs of the i^{th} party, with v_{ij} being the number of outputs for the j^{th} input. This polytope of $\dim \mathcal{B}$ is contained in the vector space of \mathbb{R}^t , where $t = \prod_{i=1}^n \sum_{j=1}^{m_i} v_{ij}$ is the total number of outputs.

Suppose now that $P \in \mathcal{B}$, then there exists a polytope $\tilde{\mathcal{B}}$, such that the NSEA $\mathcal{E}(P) \in \tilde{\mathcal{B}}$. The dimension of $\tilde{\mathcal{B}}$ can be determined by:

$$\dim \tilde{\mathcal{B}} = \prod_{i=1}^{n+1} \left(\sum_{j=1}^{m_i} (v_{ij} - 1) + 1 \right) - 1 \quad (90)$$

$$= \left(\sum_{j=1}^{m_{n+1}} (v_{n+1,j} - 1) + 1 \right) (\dim \mathcal{B} + 1) - 1, \quad (91)$$

where $n + 1$ is the index of the extending party, and the last equality is due to expansion of product with respect to $n + 1$ -th term.

To obtain the upper bound on $\dim \tilde{\mathcal{B}}$ it is enough now to find upper bounds on the number of inputs and outputs of the extending system. The following considerations have a qualitative character, we are much more interested in the fact that an upper bound on $\dim \tilde{\mathcal{B}}$ exists than in its tightness. The upper bound on the number of outputs can be found via Carathéodory theorem [107], i.e., $v_{n+1,j} \leq \dim \mathcal{B} + 1$, we obtain:

$$\dim \tilde{\mathcal{B}} \leq (m_{n+1} \dim \mathcal{B} + 1) (\dim \mathcal{B} + 1) - 1. \quad (92)$$

It is instructive to notice that the above inequality can be saturated for some generic behaviours. The number of inputs of the extending system is equal to the number of minimal ensembles of the behaviour \mathcal{P} . Suppose now that V is the number of vertices of the polytope \mathcal{B} . Then, using Carathéodory theorem again, each choice of $\dim \mathcal{B} + 1$ vertices (out of V) leads to a minimal ensemble via elimination of the vertices, therefore $m_{n+1} \leq \binom{V}{\dim \mathcal{B} + 1}$, and hence:

$$\dim \tilde{\mathcal{B}} < \left(\binom{V}{\dim \mathcal{B} + 1} \dim \mathcal{B} + 1 \right) (\dim \mathcal{B} + 1), \quad (93)$$

where the last term (-1) was neglected.

The characterisation of the vertices of polytopes of behaviours with two binary inputs and two binary outputs in bipartite and tripartite scenarios was provided in [102]. In general, however, finding vertices of a polytope is a hard problem known as the ‘‘vertex enumeration problem’’. Using McMullen’s Upper Bound Theorem [107, 109, 110], we can, however, still obtain an upper bound for the number of vertices in terms of $\dim \mathcal{B}$ and t . Here we use a fact that a polytope can be defined as an intersection of a set of halfspaces. Then following [110] (up to notation) we have the subsequent statement.

Let A be an $m \times t$ matrix of reals and let $b \in \mathbb{R}^m$ be a real vector. Consider a polytope

$$\mathcal{P} = \{x \in \mathbb{R}^t : Ax \leq b\}, \quad (94)$$

and $V_{\mathcal{P}}$ be the number of vertices of \mathcal{P} then:

$$V_{\mathcal{P}} \leq \binom{m-t-s}{s} + \binom{m-s-1}{t-s-1}, \quad (95)$$

where $s = \lfloor t/2 \rfloor$.

To achieve our goal it is enough to determine the number of rows m of matrix A . We divide now all the constraints on the non-signalling polytope into three classes, written in terms of marginal probabilities p_{ijk} :

1. Probabilistic constraints: $0 \leq p_{ijk} \leq 1$.
2. Normalization constraints, i.e., $\sum_k^{v_{ij}} p_{ijk} = 1$.
3. non-signalling constraints (see Section 4.1).

It is easy to see that constraints of type 1 contribute to $2t$ rows of matrix A . In the next step we observe that constraints of type 2 and 3 are linear and together reduce the dimension of the space of correlations from t to $\dim \mathcal{B}$, and hence can be encoded in $t - \dim \mathcal{B}$ linearly independent rows of matrix A (irrespective of the actual number of constraints). The total number of rows is therefore upper bound with $m \leq 3t - \dim \mathcal{B}$ (there still is some redundancy because of the relation between constraints of types 1 and 2), hence:

$$V \leq \binom{2t - \lfloor t/2 \rfloor - \dim \mathcal{B}}{\lfloor t/2 \rfloor} + \binom{3t - \lfloor t/2 \rfloor - (\dim \mathcal{B} + 1)}{t - \lfloor t/2 \rfloor - 1}. \quad (96)$$

Finally, we obtain

$$\begin{aligned} \dim \tilde{\mathcal{B}} &< (\dim \mathcal{B} + 1) \\ &\times \left(\left(\binom{2t - \lfloor t/2 \rfloor - \dim \mathcal{B}}{\lfloor t/2 \rfloor} + \binom{3t - \lfloor t/2 \rfloor - (\dim \mathcal{B} + 1)}{t - \lfloor t/2 \rfloor - 1} \right) \dim \mathcal{B} + 1 \right), \end{aligned} \quad (97)$$

where:

$$\dim \mathcal{B} = \prod_{i=1}^n \left(\sum_{j=1}^{m_i} (v_{ij} - 1) + 1 \right) - 1, \quad t = \prod_{i=1}^n \sum_{j=1}^{m_i} v_{ij}. \quad (98)$$

□

In Theorem 25 we were interested only in showing that the dimension of the complete extension is bounded from above, and hence our result is very loose. In particular the number of candidates for a minimal ensemble is very inefficient upper bound on the number of minimal ensembles, although obtained with simple combinatorics. Therefore, the first place for improvement is to find tighter upper bound on the number of minimal ensembles. A more rigorous treatment can show that many candidates lead to the same minimal ensemble and moreover that some of them are not valid. Furthermore, the upper bound on the number of vertices that we used is very general and does not incorporate, for example, symmetries of the non-signalling polytope.

B.3 Complete extensions of three-cycle contextual behaviour

In the previous sections, we have considered the NSCE of an arbitrary behavior which is inherently of the form $P_A(a|x)$, i.e., a conditional probability distribution consist of only single set of input x and output a . The polytope of those behaviors are completely determined by the constraints probability distributions satisfy, like non-signaling and normalization condition. Here, within the theory of non-signalling behaviours (NS) we will study contextuality scenarios, and, hence, investigate NSCE of contextual behaviours. The polytope of contextual behaviors are not simply the non-signalling polytopes but rather they are termed as no-disturbance polytopes. And while considering the minimal ensembles of a contextual behavior we will focus on the pure members of the no-disturbance polytopes.

In particular we are going to focus on the three-cycle contextuality scenario [111, 112] (aka Specker's triangle [113, 114]).

First we will introduce this contextuality scenario in standard notation before showing how to recast it within NS. The three-cycle scenario consists of a triple of binary observables $\{X_0, X_1, X_2\}$ with outcomes $\{a, b, c\}$ valued in $\{-1, 1\}$, together with the constraint that these are pairwise compatible – that is, any pair $\{X_i, X_j\}$ can be jointly measured without disturbance, but are globally incompatible – that is, there is no way to measure all three without disturbance. Such a setup can be realised in quantum theory [115] provided that the observables are not taken to correspond to projective measurements. This compatibility structure can be captured by the set of maximal contexts [111, 112]

$$\mathcal{C}_3 = \{\{X_0, X_1\}, \{X_1, X_2\}, \{X_2, X_0\}\}. \quad (99)$$

The behaviour of a particular realisation of this scenario is captured by the joint distributions over the outcomes of the compatible observables, i.e., $p(a, b)$; $p(b, c)$; and $p(c, a)$, which can be represented in the following table:

	$a \backslash b$	+1	-1	$b \backslash c$	+1	-1	$c \backslash a$	+1	-1	

Any such behaviour must satisfy certain no-disturbance constraints¹¹ [111, 112], which together define

¹¹As it should follow the no-disturbance condition like the non-signaling one in non-locality, which is

$$p(a) = \sum_{b=-1}^1 p(a, b) = \sum_{c=-1}^1 p(c, a) \quad (101)$$

$$p(b) = \sum_{a=-1}^1 p(a, b) = \sum_{c=-1}^1 p(b, c) \quad (102)$$

$$p(c) = \sum_{b=-1}^1 p(b, c) = \sum_{a=-1}^1 p(c, a) \quad (103)$$

the no-disturbance polytope. Formally defining contextuality is beyond the scope of this paper, but, for our purposes it suffices to note that a behaviour is non-contextual if it is a convex combination of the eight deterministic vertices of the no-disturbance polytope

$$\begin{aligned} \mathcal{N}_0^{\mathcal{C}} &= \frac{1}{0} \left| \frac{0}{0} \right| \left| \frac{1}{0} \right| \left| \frac{0}{0} \right| \left| \frac{1}{0} \right| \left| \frac{0}{0} \right| & \mathcal{N}_1^{\mathcal{C}} &= \frac{1}{0} \left| \frac{0}{0} \right| \left| \frac{0}{0} \right| \left| \frac{1}{0} \right| \left| \frac{0}{1} \right| \left| \frac{0}{0} \right| \\ \mathcal{N}_2^{\mathcal{C}} &= \frac{0}{0} \left| \frac{1}{0} \right| \left| \frac{0}{1} \right| \left| \frac{0}{0} \right| \left| \frac{1}{0} \right| \left| \frac{0}{0} \right| & \mathcal{N}_3^{\mathcal{C}} &= \frac{0}{1} \left| \frac{0}{0} \right| \left| \frac{1}{0} \right| \left| \frac{0}{0} \right| \left| \frac{0}{0} \right| \left| \frac{1}{0} \right| & \mathcal{N}_4^{\mathcal{C}} &= \frac{0}{0} \left| \frac{1}{0} \right| \left| \frac{0}{0} \right| \left| \frac{0}{1} \right| \left| \frac{0}{1} \right| \left| \frac{0}{0} \right| \\ \mathcal{N}_5^{\mathcal{C}} &= \frac{0}{1} \left| \frac{0}{0} \right| \left| \frac{0}{0} \right| \left| \frac{1}{0} \right| \left| \frac{0}{0} \right| \left| \frac{0}{1} \right| & \mathcal{N}_6^{\mathcal{C}} &= \frac{0}{0} \left| \frac{0}{1} \right| \left| \frac{0}{1} \right| \left| \frac{0}{0} \right| \left| \frac{0}{0} \right| \left| \frac{1}{0} \right| & \mathcal{N}_7^{\mathcal{C}} &= \frac{0}{0} \left| \frac{0}{1} \right| \left| \frac{0}{0} \right| \left| \frac{0}{1} \right| \left| \frac{0}{0} \right| \left| \frac{0}{1} \right| \end{aligned}$$

otherwise, it is a contextual behaviour. Any other vertex of the no-disturbance polytope is therefore contextual, indeed, there are four of these which are

$$\begin{aligned} \mathcal{C}_0 &= \frac{1/2}{0} \left| \frac{0}{1/2} \right| \left| \frac{1/2}{0} \right| \left| \frac{0}{1/2} \right| \left| \frac{0}{1/2} \right| \left| \frac{1/2}{0} \right| & \mathcal{C}_1 &= \frac{1/2}{0} \left| \frac{0}{1/2} \right| \left| \frac{0}{1/2} \right| \left| \frac{1/2}{0} \right| \left| \frac{1/2}{0} \right| \left| \frac{0}{1/2} \right| \\ \mathcal{C}_2 &= \frac{0}{1/2} \left| \frac{1/2}{0} \right| \left| \frac{1/2}{0} \right| \left| \frac{0}{1/2} \right| \left| \frac{1/2}{0} \right| \left| \frac{0}{1/2} \right| & \mathcal{C}_3 &= \frac{0}{1/2} \left| \frac{1/2}{0} \right| \left| \frac{0}{1/2} \right| \left| \frac{1/2}{0} \right| \left| \frac{0}{1/2} \right| \left| \frac{1/2}{0} \right|. \end{aligned}$$

Note that the dimension of the no-disturbance polytope is 6, hence, according to the Carathéodory theorem [107], the set of minimal ensembles of an arbitrary behaviour inside the polytope has at most 7 elements.

We will now focus on a particular class of behaviours within this polytope, namely those lying on an isotropic line – a mixture of a contextual vertex, say \mathcal{C}_0 and the maximally mixed behaviour¹² which is a non-contextual behaviour. We denote an element on this isotropic line (parameterised by $\lambda \in (0, 1)$) as P_λ which can be explicitly written as

$$\begin{aligned} P_\lambda &= (1 - \lambda)\mathcal{C}_0 + \lambda M \\ &= \frac{2-\lambda}{4} \left| \frac{\lambda}{4} \right| \left| \frac{2-\lambda}{4} \right| \left| \frac{\lambda}{4} \right| \left| \frac{\lambda}{4} \right| \left| \frac{2-\lambda}{4} \right| \left| \frac{\lambda}{4} \right| \end{aligned} \quad (105)$$

The minimal ensembles of P_λ , have been found by applying numerical techniques to obtain the solution of a set of linear equations, and they are as follows.

For all $\lambda \in (0, 1)$ we have the following

$$v = \left[\frac{\lambda}{4}, \frac{\lambda}{4}, 1 - \frac{\lambda}{2} \right] \quad (106)$$

$$\mathcal{M}_1(P) = \left[\mathcal{N}_2^{\mathcal{C}}, \mathcal{N}_5^{\mathcal{C}}, \mathcal{C}_0 \right] \quad (107)$$

$$v = \left[\frac{\lambda}{4}, \frac{\lambda}{4}, 1 - \lambda, \frac{\lambda}{2} \right] \quad (108)$$

$$\mathcal{M}_2 = \left[\mathcal{N}_3^{\mathcal{C}}, \mathcal{N}_4^{\mathcal{C}}, \mathcal{C}_0, \mathcal{C}_1 \right] \quad (109)$$

$$\mathcal{M}_3 = \left[\mathcal{N}_1^{\mathcal{C}}, \mathcal{N}_6^{\mathcal{C}}, \mathcal{C}_0, \mathcal{C}_2 \right] \quad (110)$$

$$\mathcal{M}_4 = \left[\mathcal{N}_0^{\mathcal{C}}, \mathcal{N}_7^{\mathcal{C}}, \mathcal{C}_0, \mathcal{C}_3 \right] \quad (111)$$

¹²The maximally mixed behaviour M is given by

$$M = \frac{1}{4} \left| \frac{1}{4} \right| \left| \frac{1}{4} \right| \left| \frac{1}{4} \right| \left| \frac{1}{4} \right| \left| \frac{1}{4} \right| \left| \frac{1}{4} \right| \quad (104)$$

$$v = \left[1 - \frac{3\lambda}{4}, \frac{\lambda}{4}, \frac{\lambda}{4}, \frac{\lambda}{4} \right] \quad (112)$$

$$\mathcal{M}_5 = \left[\mathcal{C}_0, \mathcal{C}_1, \mathcal{C}_2, \mathcal{C}_3 \right] \quad (113)$$

$$v = \left[\frac{\lambda}{4}, \frac{\lambda}{4}, \frac{\lambda}{4}, \frac{\lambda}{4}, 1 - \lambda \right] \quad (114)$$

$$\mathcal{M}_6 = \left[\mathcal{N}^{\mathcal{C}_1}, \mathcal{N}^{\mathcal{C}_2}, \mathcal{N}^{\mathcal{C}_3}, \mathcal{N}^{\mathcal{C}_7}, \mathcal{C}_0 \right] \quad (115)$$

$$\mathcal{M}_7 = \left[\mathcal{N}^{\mathcal{C}_0}, \mathcal{N}^{\mathcal{C}_4}, \mathcal{N}^{\mathcal{C}_5}, \mathcal{N}^{\mathcal{C}_6}, \mathcal{C}_0 \right] \quad (116)$$

In $\lambda \in (0, \frac{2}{3})$, we additionally have

$$v = \left[\frac{\lambda}{4}, \frac{\lambda}{4}, \frac{\lambda}{4}, \frac{\lambda}{4}, \frac{\lambda}{4}, \frac{\lambda}{4}, 1 - \frac{3\lambda}{2} \right] \quad (117)$$

$$\mathcal{M}_8 = \left[\mathcal{N}^{\mathcal{C}_0}, \mathcal{N}^{\mathcal{C}_1}, \mathcal{N}^{\mathcal{C}_3}, \mathcal{N}^{\mathcal{C}_4}, \mathcal{N}^{\mathcal{C}_6}, \mathcal{N}^{\mathcal{C}_7}, \mathcal{C}_0 \right] \quad (118)$$

Whilst in $\lambda \in (\frac{2}{3}, 1)$, the behaviour has another 5 minimal ensembles. However, in this range of λ , P_λ is non-contextual (see Theorem 1 in [112]) and we are interested in exploring the case in which P_λ displays contextuality so we will not go into detail about these.

To study the NSEAs for contextuality scenarios we must first recast the scenario into the theory of non-signalling behaviours. In particular, rather than describing the scenario by a triple of bipartite distributions (i.e., $p(a,b)$, $p(b,c)$, and $p(c,a)$) we will instead describe the scenario by a conditional probability distribution. We do this by associating the input $x \in \{0, 1, 2\}$ to the maximal contexts, i.e., $0 \sim \{X_0, X_1\}$, $1 \sim \{X_1, X_2\}$, and $2 \sim \{X_2, X_0\}$; and associating an outcome $a' \in \{0, 1, 2, 3\}$ to the outcome of the bipartite distributions $\{+1, -1\} \times \{+1, -1\}$. This defines an embedding of the no-disturbance polytope into the polytope of single input output behaviours (with the relevant dimensions).

In our particular example of interest, namely P_λ using this embedding we can associate the behaviour of the contextuality scenario P_λ to the single input output behaviour

$$P'_{A\lambda} = \begin{array}{c|ccc} & a' & 0 & 1 & 2 \\ \hline & 0 & \frac{2-\lambda}{4} & \frac{2-\lambda}{4} & \frac{\lambda}{4} \\ \hline & 1 & \frac{\lambda}{4} & \frac{\lambda}{4} & \frac{2-\lambda}{4} \\ \hline & 2 & \frac{\lambda}{4} & \frac{\lambda}{4} & \frac{2-\lambda}{4} \\ \hline & 3 & \frac{2-\lambda}{4} & \frac{2-\lambda}{4} & \frac{\lambda}{4} \end{array} \quad (119)$$

Now we are in the position that we can construct the NSCE of $P'_{A\lambda}$, in particular, focusing on the contextual case in which $\lambda \in (0, \frac{2}{3})$. Note however, that we want to construct the NSCE not for the minimal ensembles of the behaviour within the single input output polytope, but for the ensembles within the embedded no-disturbance polytope. The NSCE of this will however still live inside the polytope of no signalling behaviours with a pair of inputs and outputs (of suitable dimensions).

In particular, one can show that the NSEA of $P'_{A\lambda}$ in $\lambda \in (0, \frac{2}{3})$ is given by

z		$x=0$				$x=1$				$x=2$			
		$a=0$	1	2	3	0	1	2	3	0	1	2	3
0	0	0	$\frac{\lambda}{4}$	0	0	0	0	$\frac{\lambda}{4}$	0	$\frac{\lambda}{4}$	0	0	0
	1	0	0	$\frac{\lambda}{4}$	0	0	$\frac{\lambda}{4}$	0	0	0	0	0	$\frac{\lambda}{4}$
	2	$\frac{2-\lambda}{4}$	0	0	$\frac{2-\lambda}{4}$	$\frac{2-\lambda}{4}$	0	0	$\frac{2-\lambda}{4}$	0	$\frac{2-\lambda}{4}$	$\frac{2-\lambda}{4}$	0
1	0	0	0	$\frac{\lambda}{4}$	0	$\frac{\lambda}{4}$	0	0	0	0	$\frac{\lambda}{4}$	0	0
	1	0	$\frac{\lambda}{4}$	0	0	0	0	0	$\frac{\lambda}{4}$	0	0	$\frac{\lambda}{4}$	0
	2	$\frac{1-\lambda}{2}$	0	0	$\frac{1-\lambda}{2}$	$\frac{1-\lambda}{2}$	0	0	$\frac{1-\lambda}{2}$	0	$\frac{1-\lambda}{2}$	$\frac{1-\lambda}{2}$	0
	3	$\frac{\lambda}{4}$	0	0	$\frac{\lambda}{4}$	0	$\frac{\lambda}{4}$	$\frac{\lambda}{4}$	0	$\frac{\lambda}{4}$	0	0	$\frac{\lambda}{4}$
2	0	$\frac{\lambda}{4}$	0	0	0	0	$\frac{\lambda}{4}$	0	0	0	0	$\frac{\lambda}{4}$	0
	1	0	0	0	$\frac{\lambda}{4}$	0	0	$\frac{\lambda}{4}$	0	0	$\frac{\lambda}{4}$	0	0
	2	$\frac{1-\lambda}{2}$	0	0	$\frac{1-\lambda}{2}$	$\frac{1-\lambda}{2}$	0	0	$\frac{1-\lambda}{2}$	0	$\frac{1-\lambda}{2}$	$\frac{1-\lambda}{2}$	0
	3	0	$\frac{\lambda}{4}$	$\frac{\lambda}{4}$	0	$\frac{\lambda}{4}$	0	0	$\frac{\lambda}{4}$	$\frac{\lambda}{4}$	0	0	$\frac{\lambda}{4}$
3	0	$\frac{\lambda}{4}$	0	0	0	$\frac{\lambda}{4}$	0	0	0	$\frac{\lambda}{4}$	0	0	0
	1	0	0	0	$\frac{\lambda}{4}$	0	0	0	$\frac{\lambda}{4}$	0	0	0	$\frac{\lambda}{4}$
	2	$\frac{1-\lambda}{2}$	0	0	$\frac{1-\lambda}{2}$	$\frac{1-\lambda}{2}$	0	0	$\frac{1-\lambda}{2}$	0	$\frac{1-\lambda}{2}$	$\frac{1-\lambda}{2}$	0
	3	0	$\frac{\lambda}{4}$	$\frac{\lambda}{4}$	0	0	$\frac{\lambda}{4}$	$\frac{\lambda}{4}$	0	0	$\frac{\lambda}{4}$	$\frac{\lambda}{4}$	0
4	0	$\frac{4-3\lambda}{8}$	0	0	$\frac{4-3\lambda}{8}$	$\frac{4-3\lambda}{8}$	0	0	$\frac{4-3\lambda}{8}$	0	$\frac{4-3\lambda}{8}$	$\frac{4-3\lambda}{8}$	0
	1	$\frac{\lambda}{8}$	0	0	$\frac{\lambda}{8}$	0	$\frac{\lambda}{8}$	$\frac{\lambda}{8}$	0	$\frac{\lambda}{8}$	0	0	$\frac{\lambda}{8}$
	2	0	$\frac{\lambda}{8}$	$\frac{\lambda}{8}$	0	$\frac{\lambda}{8}$	0	0	$\frac{\lambda}{8}$	$\frac{\lambda}{8}$	0	0	$\frac{\lambda}{8}$
	3	0	$\frac{\lambda}{8}$	$\frac{\lambda}{8}$	0	0	$\frac{\lambda}{8}$	$\frac{\lambda}{8}$	0	0	$\frac{\lambda}{8}$	$\frac{\lambda}{8}$	0
5	0	$\frac{\lambda}{4}$	0	0	0	0	$\frac{\lambda}{4}$	0	0	0	0	$\frac{\lambda}{4}$	0
	1	0	$\frac{\lambda}{4}$	0	0	0	0	$\frac{\lambda}{4}$	0	$\frac{\lambda}{4}$	0	0	0
	2	0	0	$\frac{\lambda}{4}$	0	$\frac{\lambda}{4}$	0	0	0	0	$\frac{\lambda}{4}$	0	0
	3	0	0	0	$\frac{\lambda}{4}$	0	0	0	$\frac{\lambda}{4}$	0	0	0	$\frac{\lambda}{4}$
	4	$\frac{1-\lambda}{2}$	0	0	$\frac{1-\lambda}{2}$	$\frac{1-\lambda}{2}$	0	0	$\frac{1-\lambda}{2}$	0	$\frac{1-\lambda}{2}$	$\frac{1-\lambda}{2}$	0
6	0	$\frac{\lambda}{4}$	0	0	0	$\frac{\lambda}{4}$	0	0	0	$\frac{\lambda}{4}$	0	0	0
	1	0	$\frac{\lambda}{4}$	0	0	0	0	0	$\frac{\lambda}{4}$	0	0	$\frac{\lambda}{4}$	0
	2	0	0	$\frac{\lambda}{4}$	0	0	$\frac{\lambda}{4}$	0	0	0	0	0	$\frac{\lambda}{4}$
	3	0	0	0	$\frac{\lambda}{4}$	0	0	$\frac{\lambda}{4}$	0	0	$\frac{\lambda}{4}$	0	0
	4	$\frac{1-\lambda}{2}$	0	0	$\frac{1-\lambda}{2}$	$\frac{1-\lambda}{2}$	0	0	$\frac{1-\lambda}{2}$	0	$\frac{1-\lambda}{2}$	$\frac{1-\lambda}{2}$	0
7	0	$\frac{\lambda}{4}$	0	0	0	$\frac{\lambda}{4}$	0	0	0	$\frac{\lambda}{4}$	0	0	0
	1	$\frac{\lambda}{4}$	0	0	0	0	$\frac{\lambda}{4}$	0	0	0	0	$\frac{\lambda}{4}$	0
	2	0	0	$\frac{\lambda}{4}$	0	$\frac{\lambda}{4}$	0	0	0	0	$\frac{\lambda}{4}$	0	0
	3	0	$\frac{\lambda}{4}$	0	0	0	0	0	$\frac{\lambda}{4}$	0	0	$\frac{\lambda}{4}$	0
	4	0	0	0	$\frac{\lambda}{4}$	0	0	$\frac{\lambda}{4}$	0	0	$\frac{\lambda}{4}$	0	0
	5	0	0	0	$\frac{\lambda}{4}$	0	0	0	$\frac{\lambda}{4}$	0	0	0	$\frac{\lambda}{4}$
	6	$\frac{2-3\lambda}{4}$	0	0	$\frac{2-3\lambda}{4}$	$\frac{2-3\lambda}{4}$	0	0	$\frac{2-3\lambda}{4}$	0	$\frac{2-3\lambda}{4}$	$\frac{2-3\lambda}{4}$	0

(120)

One can then observe that non-locality is a special case of contextuality in which the no-disturbance polytope coincides with a non-signalling polytope – in such a case the NSCE that we construct for the contextuality scenario will be precisely that which we would construct in the non-signalling scenario.

Note that as the no-disturbance polytope is a discrete theory, that theorem 7 applies and so in general the NSEAs that we construct will not be purifications. However, it would be interesting to study the cases in which they are.

B.4 Complete extension of the behaviours lying on the isotropic line

The aim of this section is to construct explicitly the NSCE of the isotropic behaviour, that is, a mixture of the PR and anti- PR box. We therefore focus on the polytope of behaviours, $P_{AB}(ab|xy)$, with two binary inputs $x, y \in \{0, 1\}$, two binary outputs $a, b \in \{0, 1\}$, and satisfying the non-signaling condition. There are 24 extremal (pure) behaviours (vertices) of this polytope, among which 16 are local behaviours given by

$$L_{\alpha\beta\gamma\delta}(ab|xy) = \begin{cases} 1 & \text{if } a = \alpha x \oplus b, \\ & b = \gamma y \oplus \delta \\ 0 & \text{otherwise.} \end{cases} \quad (121)$$

with α, β, γ and $\delta \in \{0, 1\}$. And another 8 non-local behaviours, which are

$$B_{rst}(ab|xy) = \begin{cases} 1/2 & \text{if } a \oplus b = xy \oplus rx \oplus sy \oplus t \\ 0 & \text{otherwise.} \end{cases} \quad (122)$$

with r, s, t taking values either 0 or 1 [102]. The triple (r, s, t) enumerates the non-local behaviours in the polytope of behaviours with two binary inputs and two binary outputs. Hence, the isotropic behaviour can be formulated as, $B(\eta)_{AB}(ab|xy) = \eta B_{000}(ab|xy) + (1 - \eta)B_{111}(ab|xy)$, where $\eta \in (0, 1)$, and where B_{000} and B_{111} are the PR and anti- PR boxes respectively. One can easily check that $B(\frac{1}{2}) = P_{AB}^m(ab|xy)$, is a maximally mixed behaviour in the polytope of two input two output behaviours (the one taking value $P(ab|xy) = \frac{1}{4}$ for all $a, b, x, y \in \{0, 1\}$). All the behaviours $B(\eta)$, for $\eta \in (1, \frac{1}{2})$ can be transformed into to the behaviours $B(\eta)$, for $\eta \in (\frac{1}{2}, 0)$ by local relabeling of inputs and outputs. Thus, in our investigation we will consider only the behaviours lying on the isotropic line from the PR box to maximally mixed behaviour, i.e., for $B(\eta)$, where $\eta \in [\frac{1}{2}, 1)$.

We know from the definition of NSCE, Definition 18, that, corresponding to each input of the extending party, there is a minimal ensemble $\mathcal{M}(B(\eta))$, where the members of the ensembles are enumerated by the outputs of the extending system. Hence, finding all possible minimal ensembles of a behaviour is sufficient for the construction of NSCE.

B.4.1 Minimal ensembles for isotropic behaviours

Our aim is to find the minimal ensembles for the behaviours lying on the isotropic line, but, in this section we first focus on the minimal ensembles for a particular behaviour, the Bell Tsirelson behaviour [116]. The Bell Tsirelson behaviour, lying on the isotropic line for $\eta = \frac{2+\sqrt{2}}{4}$, reaches the quantum limit in violating the CHSH inequality [99]. The challenge in finding the minimal ensembles is that the Bell Tsirelson behaviour is specified by an irrational number, however, despite this, we are able to find all of its minimal ensembles analytically, due to the following observations and a theorem.

Observation 37. *All the pure members ensembles of the behaviours $B(\eta)$, in the isotropic line, contains the PR box, B_{000} as one of its member element, for all $\eta \in (\frac{3}{4}, 1]$.*

Proof. By direct inspection of the 24 vertices of the polytope of behaviours with two binary inputs and two binary outputs we observe that PR -box is the only behaviour, which has value 4 of the CHSH functional which reads

$$\beta(P) = \sum_{x,y} \sum_{(a,b) \in \text{supp}(B_{000})} P(ab|xy) \quad (123)$$

Moreover, for all other vertices the functional achieves values ≤ 3 . It is also clear that the functional β is linear (in fact it is equal to $2\langle P|B_{000} \rangle$ where $\langle \cdot | \cdot \rangle$ denotes Euclidean scalar product of vectors (see Lemma 4 of [117]). Let us then suppose that there exists an ensemble $\{p_i, V_i\}_i$ of a behaviour P from the isotropic line $(3/4, 1]$, which does not have PR in its set of members (here V_i are the vertices of the polytope). We have then $\beta(P) = \beta(\sum_i p_i V_i) = \sum_i p_i \beta(V_i) \leq \sum_i p_i 3 \leq 3$. This contradicts the fact that all behaviours from the considered set satisfy $\beta(P) > 3$ i.e. they violate CHSH inequality. \square

Let $S_{\mathcal{M}}(\eta)$ denote the set of all minimal ensembles of the isotropic behaviour $B(\eta)$.

Lemma 38. *Let $1 > \eta' > \eta > \frac{3}{4}$, then for any minimal ensemble from $S_{\mathcal{M}}(\eta)$, there exists a unique minimal ensemble of $S_{\mathcal{M}}(\eta')$ with the same set of members.*

Proof. Since $\eta, \eta' \in (\frac{3}{4}, 1)$ by observation 37, all the minimal ensembles of $B(\eta)$ must contain B_{000} as its members elements. Now, from lemma 29, for the set $V(\mathcal{M}_1)$ there exists unique weights $\{p_i > 0\}_{i=0}^n$ such that

$$\sum_{V_i \in V(\mathcal{M}_1)} p_i V_i = B(\eta), \quad (124)$$

where $V_0 = B_{000}$ and V_i for $i \in \{1, \dots, n\}$ are vertices of the polytope of behaviours with two binary inputs and two binary outputs. Now, from $\eta' > \eta$ there exists unique weight $1 > q > 0$ such, that $B(\eta') = qB(\eta) + (1 - q)B_{000}$. Hence, for the set $\{V_i\}_{i=0}^n$ there exist unique weights $\{r_i\}_{i=0}^n$ given by equations $r_0 = qp_0 + (1 - q)$ and $r_i = p_i(1 - q)$ for $i \in \{1, \dots, n\}$, such that $\sum_i r_i V_i = B(\eta')$. Hence, by lemma 29, the ensemble $\{r_i, V_i\}_{i=0}^n$ is minimal ensemble of $B(\eta')$, and the assertion follows. \square

We will now adopt the notation, that \mathcal{V} is the set of the sets of members of minimal ensembles:

$$\mathcal{V}(P) := \bigcup_{\mathcal{M} \in S_{\mathcal{M}}(P)} \{V(\mathcal{M}(P))\}. \quad (125)$$

From lemma 38 we have the immediate corollary:

Corollary 39. *Let $1 > \eta' > \eta > \frac{3}{4}$, and $B(\eta) = \eta B_{000} + (1 - \eta)B_{111}$ then, there is $\mathcal{V}(P(\eta)) \subseteq \mathcal{V}(P(\eta'))$*

Theorem 40. *Let $\frac{3}{4} < \eta' < \eta < \eta'' < 1$. If $\mathcal{V}(B(\eta')) = \mathcal{V}(B(\eta'')) \equiv \mathcal{V}$, then we also have $\mathcal{V}(\eta) = \mathcal{V}$.*

Proof. Since $\eta', \eta'' \in (\frac{3}{4}, 1)$, from the Corollary 39, we have

$$\mathcal{V}(B(\eta')) \subseteq \mathcal{V}(B(\eta)) \subseteq \mathcal{V}(B(\eta'')) \quad (126)$$

However by assumption there is $\mathcal{V}(B(\eta'')) \subseteq \mathcal{V}(B(\eta'))$, hence $\mathcal{V}(B(\eta)) \subseteq \mathcal{V}(B(\eta'))$, and the assertion follows. \square

The above theorem holds for the range of the isotropic parameter $(\frac{3}{4}, 1)$. Moreover, due to this theorem we are able to find out the minimal ensembles of all non-local $B(\eta)$ for $\eta \in (\frac{1}{2}, 1)$. However, numerical investigate indicates that it holds true also for the $(\frac{1}{2}, \frac{3}{4}]$, hence we conjecture as follows:

Conjecture 41. *Let $\frac{1}{2} < \eta' < \eta < \eta'' < 1$. If $\mathcal{V}(B(\eta')) = \mathcal{V}(B(\eta'')) \equiv \mathcal{V}$, then we also have $\mathcal{V}(\eta) = \mathcal{V}$.*

The above observation and theorem help us to find out the NSEA of the Bell-Tsirelson's box. There are a total of 354 minimal ensembles of $B(\eta)$, for $\eta = \frac{2+\sqrt{2}}{4}$. Surprisingly, the NSEA of a noisy PR box leads to an extending system E embedded in a vector space of dimension 2,837 (taking into account the normalization constraints it effectively lives in a 2,483 dimensional space) and the entire NSEA, is lying in a polytope which is embedded in a vector space of dimension 45,392 taking into account the normalization and non-signaling constraints effectively lives in space of dimension 22,355. The list of these ensembles are given in Sec. B.4.3.

B.4.2 Dimension of the extending party for the behaviours lying on the isotropic line

In the previous section, we discussed how one can construct the NSCE of the Bell-Tsirelson box. In this section, due to the large number of minimal ensembles for a generic $B(\eta)$, we will discuss some of the statistics of the NSCE of more general isotropic behaviours.

For any value of $\eta \in [\frac{1}{2}, 1)$, we have provided an estimate of the NSCE for the bipartite behaviour $B(\eta)$ for $\eta \in (\frac{1}{2}, 1)$. This is indeed an NSCE of this behaviour provided that conjecture 41 holds true. The number of elements in each of the minimal ensembles is bounded between 2 and $\dim \mathcal{B} + 1$, which is 9, for two input and two output binary behaviours. In the entire range of $\eta \in (\frac{1}{2}, 1)$, we have numerically calculated the total number of inputs (number of minimal ensembles $|\mathcal{M}|$) of the

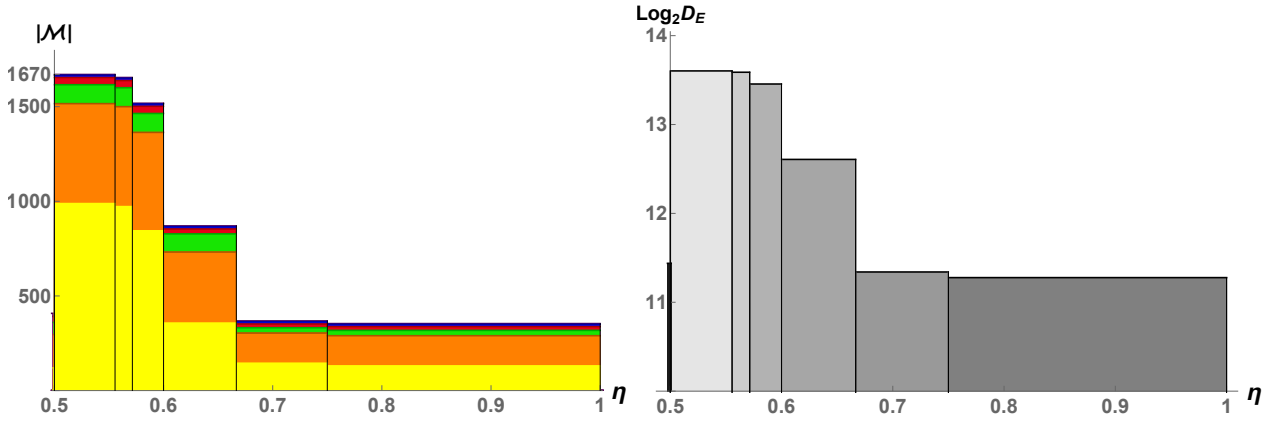


Figure 9: Panel a) Histogram plot of the total number of inputs $|\mathcal{Z}| = |\mathcal{M}|$ of the extending system of the NSEA of the behaviours along the isotropic line. Different color in each column of the histogram shows the the number of inputs having different numbers of outputs. The height of the yellow color in each column represents the number of inputs with 9 outputs, orange represents the same for 8 outputs, green is for 7 outputs, red is for 6 outputs and so on. The inputs of the extending system for various values of outputs are given in Table 1. Panel b) Histogram plot of the memory required in the extending system to store the information about the minimal ensembles of the behaviours along the isotropic line. Here D_E represents the dimension of the extending system of NSEA. All the plots have been given from $\eta = 0$ to $\eta = \frac{1}{2}$, i.e., from the PR box to the maximally mixed box.

extending system, which have an equal number of outputs $v_j \in \{2, 3, \dots, 9\}$. This is given in table 1. The Bell-Tsirelson box, $\eta = \frac{2+\sqrt{2}}{4}$, lying in the range $\eta \in (\frac{3}{4}, 1)$, is given in the last column of Table 1, having total number of inputs as 354. Among these 354 inputs there are 130 inputs having 9 outputs, 160 inputs having 8 outputs and so on. The minimal ensembles associated with each input of the Bell-Tsirelson box have been given in Sec. B.4.3 of the supplemental material. Here, \mathcal{M} represent the number of minimal ensembles of a behaviour and the length of the column represent the number of pure behaviours in each of the minimal ensembles.

$v_j \backslash \eta$	$\frac{1}{2}$	$(\frac{1}{2}, \frac{5}{9}]$	$(\frac{5}{9}, \frac{4}{7}]$	$(\frac{4}{7}, \frac{3}{5}]$	$(\frac{3}{5}, \frac{2}{3}]$	$(\frac{2}{3}, \frac{3}{4}]$	$(\frac{3}{4}, 1)$
2	4	1	1	1	1	1	1
3	0	3	3	3	3	3	3
4	12	0	0	0	0	0	0
5	0	12	12	12	12	12	12
6	32	38	38	38	26	20	20
7	64	100	100	100	96	28	28
8	176	528	528	520	376	160	160
9	120	988	972	844	356	144	130
total	408	1670	1654	1518	870	368	354

Table 1: Table shows the number of inputs (number of minimal ensembles $|\mathcal{M}|$), of the extending system which have same number of outputs v_j , for various values of $v_j \in \{2, 3, \dots, 9\}$, in the entire range of the parameter η .

The number of inputs of the extending system for various values of η , has been given in figure 9(a). The color represent the number of inputs among the total number of inputs having equal number of outputs. The yellow color stands for the set of inputs having 9 outputs, orange is for 8, green is for 7 and red is for 6 and so on. We have also plotted the memory required by the extending system, to store the information about the all possible minimal ensembles of the behaviours lying on the isotropic

line, in figure 9(b). Here D_E represent the dimension of the extending system.

$\dim \eta$	$\frac{1}{2}$	$(\frac{1}{2}, \frac{5}{9}]$	$(\frac{5}{9}, \frac{4}{7}]$	$(\frac{4}{7}, \frac{3}{5}]$	$(\frac{3}{5}, \frac{2}{3}]$	$(\frac{2}{3}, \frac{3}{4}]$	$(\frac{3}{4}, 1)$
E	2776	12445	12317	11237	6241	2595	2483
$NSCE$	24992	112013	110861	101141	56177	23363	22355

Table 2: Table shows the dimension of the extending system as well as the total dimension of the NSCE, of the behaviours lying on the isotropic line.

The dimension of the extending system as well as the total dimension of the NSCE, as given in Eq. (89), for various values of $\eta \in (\frac{1}{2}, 1)$ has been enlisted in Table 2. We have observed that the dimension of the NSCE is a maximum when the behaviour is in the vicinity of the maximally mixed behaviour. The system size is considerably lower in the range of our interest i.e., when $\eta \in (\frac{3}{4}, 1)$, that is, when the behaviour $B(\eta)$ starts showing non-classical features.

B.4.3 Minimal ensembles for non-local isotropic behaviours

Here we present the minimal ensembles of the Bell-Tsirelson box, which lies on the isotropic line

$$B(\eta) = \eta B_{000} + (1 - \eta) B_{001}, \quad (127)$$

for $\eta = \frac{2+\sqrt{2}}{4}$. This behaviour lies in a polytope of dimension $d = 8$ [102], and, according to the theorem of Carathéodory [106], the minimal ensembles of $B(\eta)$, consists of at most $d+1$ pure behaviours. Hence, we present only those minimal ensembles having at most 9 elements.

We group the ensembles by the intervals of the parameter η for which they are valid minimal ensembles. They are minimal, as otherwise we would find their subsets which would be minimal and also valid for higher α . The interval of the parameter α is the due to the requirement that the B_{000} coefficient must be greater or equal 0.

We also group the ensembles by the vector of coefficients corresponding to the behaviours of the decomposition which summed up give the $B(\alpha)$.

1. Having $\eta \in (0, 1)$.

$$v = [\eta, 1 - \eta]$$

$$\mathcal{M}_1 = [B_{000}, B_{001}]$$

2. Having $\eta \in (\frac{1}{4}, 1)$.

$$v = \left[\frac{4\eta}{3} - \frac{1}{3}, \frac{1}{6} - \frac{\eta}{6}, \frac{1}{6} - \frac{\eta}{6}, \frac{1}{6} - \frac{\eta}{6}, \frac{1}{6} - \frac{\eta}{6}, \frac{1}{6} - \frac{\eta}{6}, \frac{1}{6} - \frac{\eta}{6}, \frac{1}{6} - \frac{\eta}{6}, \frac{1}{6} - \frac{\eta}{6} \right]$$

$$\mathcal{M}_2 = [B_{000}, L_{0001}, L_{0011}, L_{0100}, L_{0110}, L_{1001}, L_{1010}, L_{1100}, L_{1111}]$$

3. Having $\eta \in (\frac{1}{3}, 1)$.

$$v = \left[\frac{3\eta}{2} - \frac{1}{2}, \frac{1}{4} - \frac{\eta}{4}, \frac{1}{4} - \frac{\eta}{4}, \frac{1}{4} - \frac{\eta}{4}, \frac{1}{4} - \frac{\eta}{4}, \frac{1}{4} - \frac{\eta}{4}, \frac{1}{4} - \frac{\eta}{4}, \frac{1}{4} - \frac{\eta}{4} \right]$$

$$\begin{aligned}
\mathcal{M}_3 &= [\mathbf{B}_{000}, L_{0000}, L_{0001}, L_{0100}, L_{0110}, L_{1001}, L_{1111}] \\
\mathcal{M}_4 &= [\mathbf{B}_{000}, L_{0001}, L_{0011}, L_{0100}, L_{0101}, L_{1010}, L_{1100}] \\
\mathcal{M}_5 &= [\mathbf{B}_{000}, L_{0001}, L_{0011}, L_{0110}, L_{0111}, L_{1010}, L_{1100}] \\
\mathcal{M}_6 &= [\mathbf{B}_{000}, L_{0001}, L_{0110}, L_{1000}, L_{1001}, L_{1100}, L_{1111}] \\
\mathcal{M}_7 &= [\mathbf{B}_{000}, L_{0001}, L_{0110}, L_{1010}, L_{1011}, L_{1100}, L_{1111}] \\
\mathcal{M}_8 &= [\mathbf{B}_{000}, L_{0010}, L_{0011}, L_{0100}, L_{0110}, L_{1001}, L_{1111}] \\
\mathcal{M}_9 &= [\mathbf{B}_{000}, L_{0011}, L_{0100}, L_{1001}, L_{1010}, L_{1100}, L_{1101}] \\
\mathcal{M}_{10} &= [\mathbf{B}_{000}, L_{0011}, L_{0100}, L_{1001}, L_{1010}, L_{1110}, L_{1111}]
\end{aligned}$$

$$v = \left[\frac{3\eta}{2} - \frac{1}{2}, \frac{1}{2} - \frac{\eta}{2}, \frac{1}{4} - \frac{\eta}{4}, \frac{1}{4} - \frac{\eta}{4}, \frac{1}{4} - \frac{\eta}{4}, \frac{1}{4} - \frac{\eta}{4} \right]$$

$$\begin{aligned}
\mathcal{M}_{11} &= [\mathbf{B}_{000}, \mathbf{B}_{010}, L_{0001}, L_{0011}, L_{0100}, L_{0110}] \\
\mathcal{M}_{12} &= [\mathbf{B}_{000}, \mathbf{B}_{011}, L_{1001}, L_{1010}, L_{1100}, L_{1111}] \\
\mathcal{M}_{13} &= [\mathbf{B}_{000}, \mathbf{B}_{100}, L_{0001}, L_{0100}, L_{1001}, L_{1100}] \\
\mathcal{M}_{14} &= [\mathbf{B}_{000}, \mathbf{B}_{101}, L_{0011}, L_{0110}, L_{1010}, L_{1111}] \\
\mathcal{M}_{15} &= [\mathbf{B}_{000}, \mathbf{B}_{110}, L_{0011}, L_{0110}, L_{1001}, L_{1100}] \\
\mathcal{M}_{16} &= [\mathbf{B}_{000}, \mathbf{B}_{111}, L_{0001}, L_{0100}, L_{1010}, L_{1111}]
\end{aligned}$$

4. Having $\eta \in \left(\frac{2}{5}, 1\right)$.

$$v = \left[\frac{5\eta}{3} - \frac{2}{3}, \frac{1}{3} - \frac{\eta}{3}, \frac{1}{6} - \frac{\eta}{6}, \frac{1}{6} - \frac{\eta}{6}, \frac{1}{6} - \frac{\eta}{6}, \frac{1}{6} - \frac{\eta}{6}, \frac{1}{6} - \frac{\eta}{6}, \frac{1}{6} - \frac{\eta}{6}, \frac{1}{3} - \frac{\eta}{3} \right]$$

$$\begin{aligned}
\mathcal{M}_{17} &= [\mathbf{B}_{000}, \mathbf{B}_{110}, L_{0001}, L_{0011}, L_{0101}, L_{0110}, L_{1010}, L_{1011}, L_{1100}] \\
\mathcal{M}_{18} &= [\mathbf{B}_{000}, \mathbf{B}_{111}, L_{0001}, L_{0010}, L_{0100}, L_{0110}, L_{1000}, L_{1001}, L_{1111}]
\end{aligned}$$

$$v = \left[\frac{5\eta}{3} - \frac{2}{3}, \frac{1}{3} - \frac{\eta}{3}, \frac{1}{6} - \frac{\eta}{6}, \frac{1}{6} - \frac{\eta}{6}, \frac{1}{6} - \frac{\eta}{6}, \frac{1}{6} - \frac{\eta}{6}, \frac{1}{3} - \frac{\eta}{3}, \frac{1}{6} - \frac{\eta}{6}, \frac{1}{6} - \frac{\eta}{6} \right]$$

$$\begin{aligned}
\mathcal{M}_{19} &= [\mathbf{B}_{000}, \mathbf{B}_{110}, L_{0000}, L_{0011}, L_{0100}, L_{0110}, L_{1001}, L_{1110}, L_{1111}] \\
\mathcal{M}_{20} &= [\mathbf{B}_{000}, \mathbf{B}_{111}, L_{0001}, L_{0011}, L_{0100}, L_{0111}, L_{1010}, L_{1100}, L_{1101}]
\end{aligned}$$

$$v = \left[\frac{5\eta}{3} - \frac{2}{3}, \frac{1}{3} - \frac{\eta}{3}, \frac{1}{6} - \frac{\eta}{6}, \frac{1}{6} - \frac{\eta}{6}, \frac{1}{6} - \frac{\eta}{6}, \frac{1}{6} - \frac{\eta}{6}, \frac{1}{3} - \frac{\eta}{3}, \frac{1}{3} - \frac{\eta}{3} \right]$$

$$\begin{aligned}
\mathcal{M}_{21} &= [\mathbf{B}_{000}, \mathbf{B}_{011}, L_{0000}, L_{0010}, L_{0100}, L_{0110}, L_{1001}, L_{1111}] \\
\mathcal{M}_{22} &= [\mathbf{B}_{000}, \mathbf{B}_{011}, L_{0001}, L_{0011}, L_{0101}, L_{0111}, L_{1010}, L_{1100}]
\end{aligned}$$

$$v = \left[\frac{5\eta}{3} - \frac{2}{3}, \frac{1}{3} - \frac{\eta}{3}, \frac{1}{6} - \frac{\eta}{6}, \frac{1}{6} - \frac{\eta}{6}, \frac{1}{6} - \frac{\eta}{6}, \frac{1}{3} - \frac{\eta}{3}, \frac{1}{6} - \frac{\eta}{6}, \frac{1}{3} - \frac{\eta}{3} \right]$$

$$\mathcal{M}_{23} = [\mathbf{B}_{000}, \mathbf{B}_{011}, L_{0010}, L_{0110}, L_{1000}, L_{1001}, L_{1100}, L_{1111}]$$

$$v = \left[\frac{5\eta}{3} - \frac{2}{3}, \frac{1}{3} - \frac{\eta}{3}, \frac{1}{6} - \frac{\eta}{6}, \frac{1}{6} - \frac{\eta}{6}, \frac{1}{6} - \frac{\eta}{6}, \frac{1}{3} - \frac{\eta}{3}, \frac{1}{3} - \frac{\eta}{3}, \frac{1}{6} - \frac{\eta}{6} \right]$$

$$\mathcal{M}_{24} = [\mathbf{B}_{000}, \mathbf{B}_{011}, L_{0011}, L_{0111}, L_{1001}, L_{1010}, L_{1100}, L_{1101}]$$

$$\mathcal{M}_{105} = [\mathbf{B}_{000}, \mathbf{B}_{010}, \mathbf{B}_{100}, \mathbf{L}_{0001}, \mathbf{L}_{0110}, \mathbf{L}_{0111}, \mathbf{L}_{1000}, \mathbf{L}_{1011}, \mathbf{L}_{1100}]$$

$$\mathcal{M}_{106} = [\mathbf{B}_{000}, \mathbf{B}_{010}, \mathbf{B}_{110}, \mathbf{L}_{0011}, \mathbf{L}_{0100}, \mathbf{L}_{0101}, \mathbf{L}_{1001}, \mathbf{L}_{1010}, \mathbf{L}_{1110}]$$

$$v = \left[\frac{7\eta}{4} - \frac{3}{4}, \frac{1}{4} - \frac{\eta}{4}, \frac{1}{4} - \frac{\eta}{4}, \frac{3}{8} - \frac{3\eta}{8}, \frac{1}{4} - \frac{\eta}{4}, \frac{1}{8} - \frac{\eta}{8}, \frac{1}{4} - \frac{\eta}{4}, \frac{1}{8} - \frac{\eta}{8}, \frac{1}{8} - \frac{\eta}{8} \right]$$

$$\mathcal{M}_{107} = [\mathbf{B}_{000}, \mathbf{B}_{010}, \mathbf{B}_{101}, \mathbf{L}_{0011}, \mathbf{L}_{0100}, \mathbf{L}_{0101}, \mathbf{L}_{1010}, \mathbf{L}_{1101}, \mathbf{L}_{1110}]$$

$$\mathcal{M}_{108} = [\mathbf{B}_{000}, \mathbf{B}_{010}, \mathbf{B}_{111}, \mathbf{L}_{0001}, \mathbf{L}_{0110}, \mathbf{L}_{0111}, \mathbf{L}_{1000}, \mathbf{L}_{1100}, \mathbf{L}_{1111}]$$

6. Having $\eta \in \left(\frac{4}{9}, 1\right)$.

$$v = \left[\frac{9\eta}{5} - \frac{4}{5}, \frac{1}{5} - \frac{\eta}{5}, \frac{1}{5} - \frac{\eta}{5}, \frac{1}{5} - \frac{\eta}{5}, \frac{1}{5} - \frac{\eta}{5}, \frac{1}{5} - \frac{\eta}{5}, \frac{1}{5} - \frac{\eta}{5}, \frac{1}{5} - \frac{\eta}{5}, \frac{2}{5} - \frac{2\eta}{5} \right]$$

$$\mathcal{M}_{109} = [\mathbf{B}_{000}, \mathbf{B}_{011}, \mathbf{B}_{100}, \mathbf{B}_{110}, \mathbf{L}_{0001}, \mathbf{L}_{0101}, \mathbf{L}_{1010}, \mathbf{L}_{1011}, \mathbf{L}_{1100}]$$

$$\mathcal{M}_{110} = [\mathbf{B}_{000}, \mathbf{B}_{011}, \mathbf{B}_{101}, \mathbf{B}_{111}, \mathbf{L}_{0010}, \mathbf{L}_{0110}, \mathbf{L}_{1000}, \mathbf{L}_{1001}, \mathbf{L}_{1111}]$$

$$v = \left[\frac{9\eta}{5} - \frac{4}{5}, \frac{1}{5} - \frac{\eta}{5}, \frac{1}{5} - \frac{\eta}{5}, \frac{1}{5} - \frac{\eta}{5}, \frac{1}{5} - \frac{\eta}{5}, \frac{1}{5} - \frac{\eta}{5}, \frac{2}{5} - \frac{2\eta}{5}, \frac{1}{5} - \frac{\eta}{5}, \frac{1}{5} - \frac{\eta}{5} \right]$$

$$\mathcal{M}_{111} = [\mathbf{B}_{000}, \mathbf{B}_{010}, \mathbf{B}_{100}, \mathbf{B}_{111}, \mathbf{L}_{0010}, \mathbf{L}_{0011}, \mathbf{L}_{0100}, \mathbf{L}_{1001}, \mathbf{L}_{1101}]$$

$$\mathcal{M}_{112} = [\mathbf{B}_{000}, \mathbf{B}_{010}, \mathbf{B}_{101}, \mathbf{B}_{110}, \mathbf{L}_{0000}, \mathbf{L}_{0001}, \mathbf{L}_{0110}, \mathbf{L}_{1011}, \mathbf{L}_{1111}]$$

$$\mathcal{M}_{113} = [\mathbf{B}_{000}, \mathbf{B}_{011}, \mathbf{B}_{100}, \mathbf{B}_{110}, \mathbf{L}_{0000}, \mathbf{L}_{0100}, \mathbf{L}_{1001}, \mathbf{L}_{1110}, \mathbf{L}_{1111}]$$

$$\mathcal{M}_{114} = [\mathbf{B}_{000}, \mathbf{B}_{011}, \mathbf{B}_{101}, \mathbf{B}_{111}, \mathbf{L}_{0011}, \mathbf{L}_{0111}, \mathbf{L}_{1010}, \mathbf{L}_{1100}, \mathbf{L}_{1101}]$$

$$v = \left[\frac{9\eta}{5} - \frac{4}{5}, \frac{1}{5} - \frac{\eta}{5}, \frac{1}{5} - \frac{\eta}{5}, \frac{1}{5} - \frac{\eta}{5}, \frac{2}{5} - \frac{2\eta}{5}, \frac{1}{5} - \frac{\eta}{5}, \frac{1}{5} - \frac{\eta}{5}, \frac{1}{5} - \frac{\eta}{5}, \frac{1}{5} - \frac{\eta}{5} \right]$$

$$\mathcal{M}_{115} = [\mathbf{B}_{000}, \mathbf{B}_{010}, \mathbf{B}_{100}, \mathbf{B}_{111}, \mathbf{L}_{0001}, \mathbf{L}_{0110}, \mathbf{L}_{0111}, \mathbf{L}_{1000}, \mathbf{L}_{1100}]$$

$$\mathcal{M}_{116} = [\mathbf{B}_{000}, \mathbf{B}_{010}, \mathbf{B}_{101}, \mathbf{B}_{110}, \mathbf{L}_{0011}, \mathbf{L}_{0100}, \mathbf{L}_{0101}, \mathbf{L}_{1010}, \mathbf{L}_{1110}]$$

7. Having $\eta \in \left(\frac{1}{2}, 1\right)$.

$$v = \left[2\eta - 1, \frac{1}{3} - \frac{\eta}{3}, \frac{1}{3} - \frac{\eta}{3}, \frac{1}{6} - \frac{\eta}{6}, \frac{1}{6} - \frac{\eta}{6}, \frac{1}{6} - \frac{\eta}{6}, \frac{1}{6} - \frac{\eta}{6}, \frac{1}{3} - \frac{\eta}{3}, \frac{1}{3} - \frac{\eta}{3} \right]$$

$$\mathcal{M}_{117} = [\mathbf{B}_{000}, \mathbf{B}_{100}, \mathbf{B}_{110}, \mathbf{L}_{0000}, \mathbf{L}_{0011}, \mathbf{L}_{0100}, \mathbf{L}_{0111}, \mathbf{L}_{1001}, \mathbf{L}_{1110}]$$

$$\mathcal{M}_{118} = [\mathbf{B}_{000}, \mathbf{B}_{100}, \mathbf{B}_{110}, \mathbf{L}_{0001}, \mathbf{L}_{0010}, \mathbf{L}_{0101}, \mathbf{L}_{0110}, \mathbf{L}_{1011}, \mathbf{L}_{1100}]$$

$$\mathcal{M}_{119} = [\mathbf{B}_{000}, \mathbf{B}_{101}, \mathbf{B}_{111}, \mathbf{L}_{0000}, \mathbf{L}_{0011}, \mathbf{L}_{0100}, \mathbf{L}_{0111}, \mathbf{L}_{1010}, \mathbf{L}_{1101}]$$

$$\mathcal{M}_{120} = [\mathbf{B}_{000}, \mathbf{B}_{101}, \mathbf{B}_{111}, \mathbf{L}_{0001}, \mathbf{L}_{0010}, \mathbf{L}_{0101}, \mathbf{L}_{0110}, \mathbf{L}_{1000}, \mathbf{L}_{1111}]$$

$$v = \left[2\eta - 1, \frac{1}{3} - \frac{\eta}{3}, \frac{1}{3} - \frac{\eta}{3}, \frac{1}{6} - \frac{\eta}{6}, \frac{1}{6} - \frac{\eta}{6}, \frac{1}{6} - \frac{\eta}{6}, \frac{1}{3} - \frac{\eta}{3}, \frac{1}{6} - \frac{\eta}{6}, \frac{1}{3} - \frac{\eta}{3} \right]$$

8. Having $\eta \in \left(\frac{5}{9}, 1\right)$.

$$v = \left[\frac{9\eta}{4} - \frac{5}{4}, \frac{1}{4} - \frac{\eta}{4}, \frac{1}{4} - \frac{\eta}{4}, \frac{1}{4} - \frac{\eta}{4}, \frac{1}{4} - \frac{\eta}{4}, \frac{1}{4} - \frac{\eta}{4}, \frac{1}{4} - \frac{\eta}{4}, \frac{1}{4} - \frac{\eta}{4}, \frac{1}{2} - \frac{\eta}{2} \right]$$

$$\mathcal{M}_{240} = [\mathbf{B}_{000}, \mathbf{B}_{010}, \mathbf{B}_{100}, \mathbf{B}_{110}, \mathbf{L}_{0011}, \mathbf{L}_{0111}, \mathbf{L}_{1000}, \mathbf{L}_{1001}, \mathbf{L}_{1110}]$$

$$\mathcal{M}_{241} = [\mathbf{B}_{000}, \mathbf{B}_{010}, \mathbf{B}_{101}, \mathbf{B}_{111}, \mathbf{L}_{0000}, \mathbf{L}_{0100}, \mathbf{L}_{1010}, \mathbf{L}_{1011}, \mathbf{L}_{1101}]$$

$$v = \left[\frac{9\eta}{4} - \frac{5}{4}, \frac{1}{4} - \frac{\eta}{4}, \frac{1}{4} - \frac{\eta}{4}, \frac{1}{4} - \frac{\eta}{4}, \frac{1}{4} - \frac{\eta}{4}, \frac{1}{4} - \frac{\eta}{4}, \frac{1}{2} - \frac{\eta}{2}, \frac{1}{4} - \frac{\eta}{4}, \frac{1}{4} - \frac{\eta}{4} \right]$$

$$\mathcal{M}_{242} = [\mathbf{B}_{000}, \mathbf{B}_{010}, \mathbf{B}_{100}, \mathbf{B}_{110}, \mathbf{L}_{0010}, \mathbf{L}_{0110}, \mathbf{L}_{1011}, \mathbf{L}_{1100}, \mathbf{L}_{1101}]$$

$$\mathcal{M}_{243} = [\mathbf{B}_{000}, \mathbf{B}_{010}, \mathbf{B}_{101}, \mathbf{B}_{111}, \mathbf{L}_{0001}, \mathbf{L}_{0101}, \mathbf{L}_{1000}, \mathbf{L}_{1110}, \mathbf{L}_{1111}]$$

$$\mathcal{M}_{244} = [\mathbf{B}_{000}, \mathbf{B}_{011}, \mathbf{B}_{100}, \mathbf{B}_{111}, \mathbf{L}_{0000}, \mathbf{L}_{0001}, \mathbf{L}_{0111}, \mathbf{L}_{1010}, \mathbf{L}_{1110}]$$

$$\mathcal{M}_{245} = [\mathbf{B}_{000}, \mathbf{B}_{011}, \mathbf{B}_{101}, \mathbf{B}_{110}, \mathbf{L}_{0010}, \mathbf{L}_{0011}, \mathbf{L}_{0101}, \mathbf{L}_{1000}, \mathbf{L}_{1100}]$$

$$v = \left[\frac{9\eta}{4} - \frac{5}{4}, \frac{1}{4} - \frac{\eta}{4}, \frac{1}{4} - \frac{\eta}{4}, \frac{1}{4} - \frac{\eta}{4}, \frac{1}{2} - \frac{\eta}{2}, \frac{1}{4} - \frac{\eta}{4}, \frac{1}{4} - \frac{\eta}{4}, \frac{1}{4} - \frac{\eta}{4}, \frac{1}{4} - \frac{\eta}{4} \right]$$

$$\mathcal{M}_{246} = [\mathbf{B}_{000}, \mathbf{B}_{011}, \mathbf{B}_{100}, \mathbf{B}_{111}, \mathbf{L}_{0010}, \mathbf{L}_{0100}, \mathbf{L}_{0101}, \mathbf{L}_{1011}, \mathbf{L}_{1111}]$$

$$\mathcal{M}_{247} = [\mathbf{B}_{000}, \mathbf{B}_{011}, \mathbf{B}_{101}, \mathbf{B}_{110}, \mathbf{L}_{0000}, \mathbf{L}_{0110}, \mathbf{L}_{0111}, \mathbf{L}_{1001}, \mathbf{L}_{1101}]$$

9. Having $\eta \in \left(\frac{4}{7}, 1\right)$.

$$v = \left[\frac{7\eta}{3} - \frac{4}{3}, \frac{1}{3} - \frac{\eta}{3}, \frac{1}{3} - \frac{\eta}{3}, \frac{1}{6} - \frac{\eta}{6}, \frac{1}{6} - \frac{\eta}{6}, \frac{1}{3} - \frac{\eta}{3}, \frac{1}{6} - \frac{\eta}{6}, \frac{1}{3} - \frac{\eta}{3}, \frac{1}{2} - \frac{\eta}{2} \right]$$

$$\mathcal{M}_{248} = [\mathbf{B}_{000}, \mathbf{B}_{010}, \mathbf{B}_{111}, \mathbf{L}_{0000}, \mathbf{L}_{0010}, \mathbf{L}_{0100}, \mathbf{L}_{1010}, \mathbf{L}_{1011}, \mathbf{L}_{1101}]$$

$$\mathcal{M}_{249} = [\mathbf{B}_{000}, \mathbf{B}_{100}, \mathbf{B}_{110}, \mathbf{L}_{0000}, \mathbf{L}_{0011}, \mathbf{L}_{0111}, \mathbf{L}_{1000}, \mathbf{L}_{1001}, \mathbf{L}_{1110}]$$

$$v = \left[\frac{7\eta}{3} - \frac{4}{3}, \frac{1}{3} - \frac{\eta}{3}, \frac{1}{3} - \frac{\eta}{3}, \frac{1}{6} - \frac{\eta}{6}, \frac{1}{6} - \frac{\eta}{6}, \frac{1}{3} - \frac{\eta}{3}, \frac{1}{3} - \frac{\eta}{3}, \frac{1}{6} - \frac{\eta}{6}, \frac{1}{2} - \frac{\eta}{2} \right]$$

$$\mathcal{M}_{250} = [\mathbf{B}_{000}, \mathbf{B}_{010}, \mathbf{B}_{100}, \mathbf{L}_{0001}, \mathbf{L}_{0011}, \mathbf{L}_{0111}, \mathbf{L}_{1000}, \mathbf{L}_{1001}, \mathbf{L}_{1110}]$$

$$v = \left[\frac{7\eta}{3} - \frac{4}{3}, \frac{1}{3} - \frac{\eta}{3}, \frac{1}{3} - \frac{\eta}{3}, \frac{1}{6} - \frac{\eta}{6}, \frac{1}{6} - \frac{\eta}{6}, \frac{1}{3} - \frac{\eta}{3}, \frac{1}{2} - \frac{\eta}{2}, \frac{1}{6} - \frac{\eta}{6}, \frac{1}{3} - \frac{\eta}{3} \right]$$

$$\mathcal{M}_{251} = [\mathbf{B}_{000}, \mathbf{B}_{010}, \mathbf{B}_{110}, \mathbf{L}_{0000}, \mathbf{L}_{0010}, \mathbf{L}_{0110}, \mathbf{L}_{1011}, \mathbf{L}_{1100}, \mathbf{L}_{1101}]$$

$$\mathcal{M}_{252} = [\mathbf{B}_{000}, \mathbf{B}_{101}, \mathbf{B}_{111}, \mathbf{L}_{0001}, \mathbf{L}_{0010}, \mathbf{L}_{0101}, \mathbf{L}_{1000}, \mathbf{L}_{1110}, \mathbf{L}_{1111}]$$

$$\begin{aligned}
\mathcal{M}_{328} &= [\mathbf{B}_{000}, \mathbf{B}_{011}, \mathbf{B}_{100}, \mathbf{L}_{0010}, \mathbf{L}_{0101}, \mathbf{L}_{0111}, \mathbf{L}_{1011}, \mathbf{L}_{1100}] \\
\mathcal{M}_{329} &= [\mathbf{B}_{000}, \mathbf{B}_{011}, \mathbf{B}_{101}, \mathbf{L}_{0000}, \mathbf{L}_{0101}, \mathbf{L}_{0111}, \mathbf{L}_{1010}, \mathbf{L}_{1101}] \\
\mathcal{M}_{330} &= [\mathbf{B}_{000}, \mathbf{B}_{011}, \mathbf{B}_{110}, \mathbf{L}_{0000}, \mathbf{L}_{0101}, \mathbf{L}_{0111}, \mathbf{L}_{1001}, \mathbf{L}_{1110}] \\
\mathcal{M}_{331} &= [\mathbf{B}_{000}, \mathbf{B}_{011}, \mathbf{B}_{111}, \mathbf{L}_{0010}, \mathbf{L}_{0101}, \mathbf{L}_{0111}, \mathbf{L}_{1000}, \mathbf{L}_{1111}] \\
\mathcal{M}_{332} &= [\mathbf{B}_{000}, \mathbf{B}_{011}, \mathbf{L}_{0000}, \mathbf{L}_{0111}, \mathbf{L}_{1001}, \mathbf{L}_{1010}, \mathbf{L}_{1101}, \mathbf{L}_{1110}] \\
\mathcal{M}_{333} &= [\mathbf{B}_{000}, \mathbf{B}_{011}, \mathbf{L}_{0010}, \mathbf{L}_{0101}, \mathbf{L}_{1000}, \mathbf{L}_{1011}, \mathbf{L}_{1100}, \mathbf{L}_{1111}] \\
\mathcal{M}_{334} &= [\mathbf{B}_{000}, \mathbf{B}_{100}, \mathbf{B}_{111}, \mathbf{L}_{0010}, \mathbf{L}_{0100}, \mathbf{L}_{0111}, \mathbf{L}_{1011}, \mathbf{L}_{1101}] \\
\mathcal{M}_{335} &= [\mathbf{B}_{000}, \mathbf{B}_{101}, \mathbf{B}_{110}, \mathbf{L}_{0000}, \mathbf{L}_{0101}, \mathbf{L}_{0110}, \mathbf{L}_{1011}, \mathbf{L}_{1101}]
\end{aligned}$$

$$v = \left[\frac{5\eta}{2} - \frac{3}{2}, \frac{1}{2} - \frac{\eta}{2}, \frac{1}{2} - \frac{\eta}{2}, \frac{1}{2} - \frac{\eta}{2}, \frac{1}{2} - \frac{\eta}{2}, \frac{1}{2} - \frac{\eta}{2}, \frac{1}{2} - \frac{\eta}{2} \right]$$

$$\begin{aligned}
\mathcal{M}_{336} &= [\mathbf{B}_{000}, \mathbf{B}_{010}, \mathbf{B}_{100}, \mathbf{B}_{110}, \mathbf{L}_{1011}, \mathbf{L}_{1110}] \\
\mathcal{M}_{337} &= [\mathbf{B}_{000}, \mathbf{B}_{010}, \mathbf{B}_{101}, \mathbf{B}_{111}, \mathbf{L}_{1000}, \mathbf{L}_{1101}] \\
\mathcal{M}_{338} &= [\mathbf{B}_{000}, \mathbf{B}_{011}, \mathbf{B}_{100}, \mathbf{B}_{111}, \mathbf{L}_{0010}, \mathbf{L}_{0111}] \\
\mathcal{M}_{339} &= [\mathbf{B}_{000}, \mathbf{B}_{011}, \mathbf{B}_{101}, \mathbf{B}_{110}, \mathbf{L}_{0000}, \mathbf{L}_{0101}]
\end{aligned}$$

11. Having $\eta \in \left(\frac{2}{3}, 1\right)$.

$$v = \left[3\eta - 2, \frac{1}{2} - \frac{\eta}{2}, \frac{1}{2} - \frac{\eta}{2}, \frac{1}{2} - \frac{\eta}{2}, \frac{1}{2} - \frac{\eta}{2}, \frac{1}{2} - \frac{\eta}{2}, \frac{1}{2} - \frac{\eta}{2} \right]$$

$$\begin{aligned}
\mathcal{M}_{340} &= [\mathbf{B}_{000}, \mathbf{L}_{0000}, \mathbf{L}_{0001}, \mathbf{L}_{0101}, \mathbf{L}_{0111}, \mathbf{L}_{1000}, \mathbf{L}_{1110}] \\
\mathcal{M}_{341} &= [\mathbf{B}_{000}, \mathbf{L}_{0000}, \mathbf{L}_{0010}, \mathbf{L}_{0100}, \mathbf{L}_{0101}, \mathbf{L}_{1011}, \mathbf{L}_{1101}] \\
\mathcal{M}_{342} &= [\mathbf{B}_{000}, \mathbf{L}_{0000}, \mathbf{L}_{0010}, \mathbf{L}_{0110}, \mathbf{L}_{0111}, \mathbf{L}_{1011}, \mathbf{L}_{1101}] \\
\mathcal{M}_{343} &= [\mathbf{B}_{000}, \mathbf{L}_{0000}, \mathbf{L}_{0111}, \mathbf{L}_{1000}, \mathbf{L}_{1001}, \mathbf{L}_{1101}, \mathbf{L}_{1110}] \\
\mathcal{M}_{344} &= [\mathbf{B}_{000}, \mathbf{L}_{0000}, \mathbf{L}_{0111}, \mathbf{L}_{1010}, \mathbf{L}_{1011}, \mathbf{L}_{1101}, \mathbf{L}_{1110}] \\
\mathcal{M}_{345} &= [\mathbf{B}_{000}, \mathbf{L}_{0010}, \mathbf{L}_{0011}, \mathbf{L}_{0101}, \mathbf{L}_{0111}, \mathbf{L}_{1000}, \mathbf{L}_{1110}] \\
\mathcal{M}_{346} &= [\mathbf{B}_{000}, \mathbf{L}_{0010}, \mathbf{L}_{0101}, \mathbf{L}_{1000}, \mathbf{L}_{1011}, \mathbf{L}_{1100}, \mathbf{L}_{1101}] \\
\mathcal{M}_{347} &= [\mathbf{B}_{000}, \mathbf{L}_{0010}, \mathbf{L}_{0101}, \mathbf{L}_{1000}, \mathbf{L}_{1011}, \mathbf{L}_{1110}, \mathbf{L}_{1111}]
\end{aligned}$$

$$v = \left[3\eta - 2, 1 - \eta, \frac{1}{2} - \frac{\eta}{2}, \frac{1}{2} - \frac{\eta}{2}, \frac{1}{2} - \frac{\eta}{2}, \frac{1}{2} - \frac{\eta}{2} \right]$$

$$\begin{aligned}
\mathcal{M}_{348} &= [\mathbf{B}_{000}, \mathbf{B}_{010}, \mathbf{L}_{1000}, \mathbf{L}_{1011}, \mathbf{L}_{1101}, \mathbf{L}_{1110}] \\
\mathcal{M}_{349} &= [\mathbf{B}_{000}, \mathbf{B}_{011}, \mathbf{L}_{0000}, \mathbf{L}_{0010}, \mathbf{L}_{0101}, \mathbf{L}_{0111}] \\
\mathcal{M}_{350} &= [\mathbf{B}_{000}, \mathbf{B}_{100}, \mathbf{L}_{0010}, \mathbf{L}_{0111}, \mathbf{L}_{1011}, \mathbf{L}_{1110}] \\
\mathcal{M}_{351} &= [\mathbf{B}_{000}, \mathbf{B}_{101}, \mathbf{L}_{0000}, \mathbf{L}_{0101}, \mathbf{L}_{1000}, \mathbf{L}_{1101}] \\
\mathcal{M}_{352} &= [\mathbf{B}_{000}, \mathbf{B}_{110}, \mathbf{L}_{0000}, \mathbf{L}_{0101}, \mathbf{L}_{1011}, \mathbf{L}_{1110}] \\
\mathcal{M}_{353} &= [\mathbf{B}_{000}, \mathbf{B}_{111}, \mathbf{L}_{0010}, \mathbf{L}_{0111}, \mathbf{L}_{1000}, \mathbf{L}_{1101}]
\end{aligned}$$

12. Having $\eta \in \left(\frac{3}{4}, 1\right)$.

$$v = \left[4\eta - 3, \frac{1}{2} - \frac{\eta}{2}, \frac{1}{2} - \frac{\eta}{2}, \frac{1}{2} - \frac{\eta}{2}, \frac{1}{2} - \frac{\eta}{2}, \frac{1}{2} - \frac{\eta}{2}, \frac{1}{2} - \frac{\eta}{2}, \frac{1}{2} - \frac{\eta}{2}, \frac{1}{2} - \frac{\eta}{2} \right]$$

$$\mathcal{M}_{354} = [\mathbf{B}_{000}, \mathbf{L}_{0000}, \mathbf{L}_{0010}, \mathbf{L}_{0101}, \mathbf{L}_{0111}, \mathbf{L}_{1000}, \mathbf{L}_{1011}, \mathbf{L}_{1101}, \mathbf{L}_{1110}]$$

List of Symbols and abbreviations:

PP: Purification postulate.

CEP: Complete extension postulate.

A: Access property.

G: Generation property.

EU: Essential uniqueness.

ONSEA: Overcomplete non-signaling extension with access.

NSEA: Complete non-signaling extension with access.

NSCE: Non-signaling complete extension.

GPT: Generalised probabilistic theory.

\mathcal{G} : A generalized probabilistic theory.

$\text{Syst}[\mathcal{G}]$: Systems associated with \mathcal{G} .

\otimes : Composition rule in a GPT.

\times : Cartesian product.

V_A : A finite dimensional real vector space associated to system A .

V_A^* : Dual space of V_A .

\mathbb{R}^Λ : Vector space of real valued functions for some finite sample space Λ in classical theory.

Ω_A : Convex set within the vector space V_A , represents the state space of system A .

Ω_A^* : subset of dual vectors in V_A^* which evaluate to (a subset of) the unit interval on Ω_A .

\mathcal{E}_A : Convex set within the dual space V_A^* , represents the effect space of system A .

K_A : Convex cone associated with Ω_A .

e : An instance of an effect in the effect space \mathcal{E}_A .

$e(\Omega_A)$: Image of e when its domain is restricted to Ω_A .

u_A : Unit effect on Ω_A .

\mathcal{T}_A^B : Space of transformations from system A to system B .

M : Measurement, belongs to \mathcal{T}_A^B .

Δ_I : Classical systems contained in every GPT, where I denotes the set of classical (deterministic) states.

δ_i : a vertex of Δ_I .

ϵ_i : a vertex of Δ_I^* .

$\mathbb{1}_{\Delta_I}$: Identity transformation for system Δ_I .

ω_A : Arbitrary state of system A .

ϵ_{AB} : A non-signalling extension of ω_A to system B .

E_{ω_A} : The set of all states that purify ω_A .

Tr_B : Partial trace over system B , of a composite system.

$T(T')$: A system type in $\text{Syst}[\mathcal{G}]$.

s : An arbitrary state in a GPT.

$s(T)$: An arbitrary state of type T .

$\{(p_i, s_i)\}$: An ensemble of states.

$\mathbf{Ens}[s]$: The set of all possible ensembles for a state s .

$\mathbf{Ens}_P[s]$: The set of all pure ensembles of s .

$\mathbf{Face}[s]$: Face of a state s .

$\mathbf{Ext}[s]$: The set of extensions of s .

$\mathbf{Ext}_P[s]$: The set of all pure extensions of s .

$\mathbf{Ext}_{class}[s]$: A class of extension of s , where the extending system is classical.

$\text{Vert}[\Delta_I]$: The set of vertices of Δ_I .

σ^P : A purification of system s , belongs to $\mathbf{Ext}_P[s]$.

Σ : Arbitrary extension of system s , belongs to $\mathbf{Ext}[s]$.

$T_{1 \rightarrow 2}$: A reversible transformation in the extending system.

p_A^* : Alice's cheating probability in integer-commitment.

p_B^* : Bob's cheating probability in integer-commitment.

Tr_B : Partial trace on system B of a composite quantum state ρ_{AB} .

$\text{Tr}_{\neq A_i}$: Partial trace on all systems despite of system A_i of a composite state $\rho_{A_1 A_2 \dots A_N}$.

V : Number of the vertices within a theory.

\aleph_0 : The cardinality of the set of natural numbers.

\mathfrak{c} : The cardinality of the continuum.

\mathcal{P} : A partition of a composite system into mutually non-signaling subsystems.

\mathcal{S}_i^P : An i^{th} system of a partition \mathcal{P} .

ρ_A : Arbitrary quantum state of system A .

$|\psi_{AE}\rangle$: A pure quantum state of the composite system A and E.

Θ_E : Quantum channel in part of system E.

$\{(p_i, \rho_A^i)\}$: Ensemble of a quantum state.

P_A : A behaviour of system A.

P_{AE} : A composite behaviour of system A and E.

$\{p_i, P_A^i\}$: Ensemble of the behaviour of system A.

$p_{A|X}(a|x)$: A conditional probability distribution.

$p_{AE|XZ}(ae|xz)$: A bipartite conditional probability distribution.

\mathcal{X} : The set of all input $\{x\}$, of the conditional probability distribution $p_{A|X}(a|x)$.

\mathcal{Z} : The set of all input $\{z\}$, of the bipartite conditional probability distribution $p_{AE|XZ}(ae|xz)$.

\mathcal{A} : The set of all output $\{a\}$, of the conditional probability distribution $p_{A|X}(a|x)$.

\mathcal{E} : The set of all output $\{e\}$, of the bipartite conditional probability distribution $p_{AE|XZ}(ae|xz)$.

\mathcal{P}_E : Classical pre(post)-processing channel acting on the inputs(outputs) of system E.

P_E : A behaviour which is an extremal point.

$\mathcal{E}(P)$: An arbitrary ensemble of the behaviour P .

$V(\mathcal{E})$: The set of members of the ensemble \mathcal{E} .

$\mathcal{E}_{pure}(P)$: Pure members ensemble of the behaviour P .

$\mathcal{M}(P)$: Minimal ensemble of the behaviour P .

$\mathcal{E}(P)_{AE}$: Non-signaling Extension with Access of behaviour P_A .

D_j : A dice, characterized by the probability distribution $\tilde{p}(k|z' = j)$.

C_j : Classical post-processing channel, characterized by conditional distribution $p_c(m|e, z' = j)$.

\mathcal{B} : A behaviour polytope.

$\dim \mathcal{B}$: Dimension of a behaviour polytope.

P^{PR} : Popescu-Rohrlich box.

P^m : Maximally mixed behaviour.

$P_{A'}$: Conjugate behaviour of behaviour P_A .

$L_{\alpha\beta\gamma\delta}$: Local vertices of the polytope of binary input out behaviours, $\alpha, \beta, \gamma, \delta \in \{0, 1\}$.

B_{rst} : Non-local vertices of the polytope of binary input out behaviours, $r, s, t \in \{0, 1\}$.

$B(\eta)$: Arbitrary behaviour on the line joining PR box and anti-PR box.

$\beta(P)$: Linear map on the behaviour P , computing the CHSH functional.

$\mathcal{V}(P)$: Set of the sets of members of minimal ensembles.

\mathcal{C}_3 : The set of maximal contexts for the three-cycle contextuality scenario.

\mathcal{N}_j^C : j^{th} non-contextual vertex of the no-disturbance polytope.

\mathcal{C}_j : j^{th} contextual vertex of the no-disturbance polytope.