Department of Mathematical Sciences
University of Copenhagen

PhD Thesis

An Information-Theoretic Framework for Quantum Repeaters

Roberto Ferrara

Advisor:	Matthias Christandl University of Copenhagen
Assessment Committee:	Laura Mančinska University of Copenhagen
	Peter Harremoës Niels Brock, Copenhagen Business College
	Michał Horodecki University of Gdańsk

This thesis has been submitted to the PhD School of The Faculty of Science,
University of Copenhagen

August 1st, 2018

Roberto Ferrara
Department of Mathematical Sciences
University of Copenhagen
Universitetsparken 5
2100 Copenhagen Ø

roberto@math.ku.dk
roberto.ferrara87@outlook.com

Advisor:

Matthias Christandl
QMATH, Department of Mathematical Sciences
University of Copenhagen

Assessment Committee:

Laura Mančinska
QMATH, Department of Mathematical Sciences
University of Copenhagen

Peter Harremoës
Niels Brock, Copenhagen Business College

Michał Horodecki
Institute of Theoretical Physics and Astrophysics
National Quantum Information Centre
Faculty of Mathematics, Physics and Informatics
University of Gdańsk

ISBN: 978-87-7078-912-7

Submitted on the 1st of August, 2018
Corrected for printing on the 1st of October, 2018

Acknowledgments

To my supervisor, the committee, my collaborators and QMATH for making my PhD possible. I will also never forget the high quality coffee breaks and discussions ... all hail the Rat King. In particular to Chris, Johannes, mAlexander and Florian for helping me go through my month of speed writing.
To Dino, Dimitrios and Francesco for a proper mediterranean office.
To Dimitrios, Erica, Luca e Artù per le partite a briscola.
To the Frederiksberg Rugby Klub og til Flipper Undervandsrugby for helping to keep me sane, and the Copenhagen Touch Rugby for letting me be insane with barbecues, dancing-burgers and watermelons. E tutto ciò non sarebbe mai successo se i Torelli Sudati non mi avessero introdotto al rugby.
Ai Fisici Bestiali incluso Elisa OBiCi e quel tipo che era con noi anche alle superiori, perchè se mi ricordo del fatto che ti dimentico, allora un posticino te lo meriti. In particolare a Tommy con la promessa che un giorno scannerizzerò il plico di puttanate che ho messo da parte a casa . . . fortunatamente non ho conservato una certa bottiglia di plastica.
Ai Polli e al Pollaio per gli invasi, gli scacchi, i lucchetti, grigliate, le nespole e una camminata da non dimeticare.
A Zanna, Campa, Virgu, Verio, Zocca, Dalillo, Pettenó, ZANETTI, Ergest, Giuseppe aka icsileva, Ele, Fra, Giorgia, Maila, Luca, Marcolin, Menego, Tommy, Luca, Enrico, Bane, Malaponti, Rino, Alan e Fabio per la migliore classe che mi poteva capitare.
Alla Chiara, Sere e Matteo per gli innumerabili viaggi in autobus, i viaggi nei manubri delle bici . . . e i voli dai manubri delle bici.
A mis familiares en España por quedar y estar juntos aún mas que antes.
E alla zia Lucia per aver resistito così tanto.
To everybody I could not mention explicitly.
Y por fin a Marina, Fabio, Mamà y Giovanni.

Contributions

The content of this thesis is based on the work made during the three years of PhD at QMATH, the Villum Centre for the Mathematics of Quantum Theory, at the University of Copenhagen. Some of the work is based on pubblications [1] and [2].

[1] M. Christandl and R. Ferrara. “Private States, Quantum Data Hiding, and the Swapping of Perfect Secrecy”. In: *Physical review letters* 119.22 (2017), p. 220506. DOI: 10.1103/PhysRevLett.119.220506. arXiv: 1609.04696 [quant-ph].

[2] M. Christandl, R. Ferrara, and C. Lancien. “Random private quantum states”. In: (2018). arXiv: 1801.02861 [quant-ph].

The thesis also includes unpublished work done in collaboration with Karol Horodecki from the National Quantum Information Centre at the University of Gdańsk, and Māris Ozols from QuSoft and the University of Amsterdam.

Contents

Abstract

Pure maximally entangled states are the most powerful resource provided by quantum mechanics. Entanglement distillation is the process of producing these states with high-fidelity between two distant parties, starting from a source of noisy entanglement. The rate at which this can be done is called distillable entanglement. The maximally entangled states can be used for teleportation of quantum information, Bell inequality violation, and, most importantly from the point of view of this thesis, generation of perfect key. Maximally entangled states are pure and in product with the environment, and thus they guarantee that a simple local measurement done by the parties will produce perfect key, namely perfectly equal and perfectly random strings shared between the two distant parties, with the absolute guarantee that they will be secret to anybody else. The process of using quantum states to share perfect key between distant parties is called quantum key distribution, and the rate at which it can be done is called distillable key. It turns out that there exist noisy entangled states, the private states, that also lead to perfect key just by measurement and, most surprisingly, that the distillable key equals the distillation rate of such states. Proving this equivalence allowed the authors to show that distillable entanglement and distillable key can be very different. There even exists a low-dimensional experimental realization. However, these are very peculiar correlations that have been shown only for parties that interact together directly on the noisy entanglement. In the future of quantum processing, and in the not so distant future of quantum key distribution, parties will be distributed in a network, where the entanglement will have to be generated by parties that are directly connected, and then relayed to arbitrary nodes using quantum teleportation. The intermediate parties are then known as quantum repeater stations, or simply quantum repeaters. In light of this, it is natural to ask how much the separation between distillable key and distillable entanglement extends to general network scenarios, and in particular whether it persists if we insert a repeater station between the two parties.

In this thesis, we provide a new perspective on key distillation, and thus quantum key distribution, by relating private states to quantum data hiding, the phenomenon of having perfect classical correlations that are not accessible by separated parties. This provides a tool for the study of quantum key distribution involving intermediate repeater stations, where for the first time we are able to show a close connection with entanglement distillation. We show the first bounds on the distillable key in quantum repeaters in terms of the distillable entanglement between the nodes, holding in particular for private states, and thus for states that do indeed provide perfect key between the nodes. To develop the tools, we expand the understanding of private states. For them we provide a simplification, that allows us to connect the distillable entanglement and the repeater distillable key to the recoverability of classical information by local parties, when these private states are used as encoders. We also show that in general most private states will have low recoverability of this classical information, which is the intuition behind private states with low distillable entanglement, and show that, under mild assumptions, this implies a low key in some realistic repeater. Our results add toward the intuition that the distillable entanglement is the only relevant resource that survives the relay of quantum information.

Resumé

Rene, maksimalt sammenfiltret stater er den mest kraftfulde ressource ved kvantemekanik. Sammenfiltring destillation er processen med at producere disse stater med høj pålidelighed mellem to fjerntliggende parter, der starter fra en kilde til støjende sammenfiltring. Den hastighed, hvormed dette kan gøres, kaldes destillerbar sammenfiltring. De maksimalt sammenfiltret stater kan bruges til teleportation af kvante oplysninger, Bell-ulighedskrænkelse, og vigtigst af afhandlings synspunkt, generering af perfekt nøgle. Maksimalt sammenfiltret stater er rene og i produkt med miljøet, og dermed garanterer de, at en enkelt lokal måling foretaget af parterne vil producere perfekt nøgle, nemlig helt lige og perfekt tilfældig data delt mellem de to fjerne parter, med absolut garanti for, at den vil være hemmelig for nogen andre. Processen med at bruge kvante stater til, at dele perfekt nøgle mellem fjerne parter kaldes kvante nøglefordeling, og den hastighed, hvormed den kan gøres, kaldes destillerbar nøgle. Det viser sig, at der findes støjende sammenfiltret stater, de private stater, som også fører til en perfekt nøgle bare ved måling, og det mest overraskende den destillerbare nøgle svarer til destillationshastigheden af sådanne tilstande. At bevise denne ækvivalens tillod forfatterne at vise, at den destillerbare sammenfiltring og destillationsnøgle kan være meget anderledes. Der eksisterer endog en lav dimensionel eksperimentel realisering. Men disse er meget ejendommelige korrelationer, der er blevet vist kun for parter, der interagerer direkte sammen på den støjende sammenfiltring. I fremtiden for kvantebehandling, og i den ikke så fjerne fremtid med kvantenøglefordeling, parter vil være fordelt i et netværk, hvor sammenfiltring vil være nødt til at blive genereret af parter, der er direkte forbundet, og derefter videregivet til vilkårlige knudepunkter ved hjælp af quantum teleportation. De mellemliggende parter er så kendt som kvante relæstationer, eller blot kvante relæ. I lyset af dette er det naturligt at spørge, hvor meget adskillelsen mellem destillerbar nøgle og destillerbar sammenfiltring strækker sig til generelle netværks scenarier, og især om det fortsætter hvis vi indsætter en relæstation station mellem de to parter.

I denne afhandling giver vi et nyt perspektiv på nøgledestillation, og dermed kvante nøglefordeling, ved at forbinde private stater til kvante skjult data, fænomenet at have perfekte klassiske korrelationer der ikke er tilgængelige af adskilte parter. Dette giver et værktøj til undersøgelse af kvante nøglefordeling med involveret mellemliggende relæstationer, hvor vi for første gang kan vise en tæt forbindelse med sammenfiltring destillation. Vi viser de første grænser på destillationsnøglen i kvante relæer i form af destillable sammenfiltring mellem knuderne, især for private stater, og dermed for stater, som faktisk giver perfekt nøgle mellem knuderne. For at udvikle værktøjerne udvider vi forståelsen af private stater. For dem giver vi en forenkling, der tillader os at forbinde den destillerbare sammenfiltring, og relæ destillationsnøglen til nyttiggørelse af klassiske oplysninger af lokale parter, når disse private stater bruges som encodere. Vi viser også, at de fleste private stater generelt vil have lav sammenfiltring evne for denne klassiske information, som er intuitionen bag private stater med lavt destillerbar sammenfiltring, og viser, at under milde antagelser, dette indebærer en lav nøgle i nogle realistiske kvante relæer. Vores resultater tilføjer i retning af intuitionen, at destillable sammenfiltring er den eneste relevante ressource, som overlever kvanteinformationsrelæet.

Errata

- Section 1.3.1, page 27: Any global entangled measurement is a separability- and PPT-preserving operation that is not a separable operation. However, there are entangled measurements that are PPT operations.

- Section 1.4.2, page 31: "L is informationally complete if for all operators $K \in \mathcal{L}(\mathrm{AB})$ $\ldots$".

- Conjecture 9: ρ and ρ' do not need to be ε-close.

- Section 2.4.3, page 59, the two equations before Equation (2.18): Because separable states are not close under PPT measurements operators, the correct equations should be:
$$E_{D,\mathsf{L}} \leqslant D^{\infty}_{\mathsf{L}(\underline{\mathrm{A}}:\underline{\mathrm{B}})}(\rho \| \mathcal{P}(\mathrm{A}:\mathrm{B})).$$
$$E_{D,\mathsf{L}} \leqslant D^{\infty}_{\mathsf{L}(\underline{\mathrm{A}}:\underline{\mathrm{B}})}(\rho \| \mathcal{P}(\mathrm{A}:\mathrm{B})) \leqslant D^{\infty}_{\mathsf{L}(\underline{\mathrm{A}}:\underline{\mathrm{B}})}(\rho \| \sigma) \leqslant D^{\infty}_{\mathsf{L}(\underline{\mathrm{A}}:\mathrm{B})}(\rho \| \sigma).$$
Since $\sigma \in \mathcal{P}(\mathrm{A}:\mathrm{B})$, then it applies also for $\sigma \in \mathcal{S}(\mathrm{A}:\mathrm{B})$.

- Section 3.1.3, page 72:
$$E_R(\Phi) \leqslant E_R^{\infty}(\gamma) \leqslant E_R(\gamma).$$
is not straightforward and does not derive from the unitary invariance of the relative entropy, It is a theorem in [Hor+09].

- Section 3.1.3, page 72: $E_R^{\infty}(\gamma^m) = E_R(\gamma^m) = m$ if γ is strictly irreducible but not only if.

- Section 4.5, page 135 after Equation (28): A factor of $1/E$ is consistently missing in the expressions for I and II.

Prologue

Today, most of the internet uses private-key encrypted communication. The messages are encrypted with random strings using modulo addition. The random string, known as the key, needs be known to both communicating parties, so that it can be added at the sender and be removed at the receiver. If the keys are as long as the message and provably uniformly random, then this scheme is provably secure. Namely if the key is secret to anyone else (private), and anybody, usually called an eavesdropper, intercepts the encryption, they will get the same amount of information about the original message as making a random guess. Equivalently, from the eavesdropper point of view, any message will be equally likely, but for this to happen it is of paramount importance for the key to be perfect, namely uniformly random and secret. However, pure provably uniform randomness is hard to come by, thus producing as much randomness as the amount of communication that is secured is extremely impractical, if not impossible. The current solution is to acknowledge that malicious adversaries, like an eavesdropper, have physical limitations and thus limited computational power. A specific type of conjectured hard problem is then used to produce, from a relatively small perfect key, deterministically and quickly a large amount of seemingly random keys. Under the hardness assumption, an adversary will need to spend an impossible amount of computation to break the scheme. However, it is still of paramount importance that the initial key be perfect.

We are thus still left with the problem of sharing the initial perfect key. The solution is again to use another specific type of conjectured hard problems, that allows for public-key encryption schemes. In such schemes, the key for encryption and the key for decryption are different, and it should be impractical for any physical adversary to figure out how to decrypt even knowing the encryption key. The result is that a party Alice can publish the encryption key, so that other parties can communicate private keys to Alice secretly from each other. The security guarantee is quite remarkable, as it allows two distant parties to share a (computationally secure) key from a distance without ever having interacted before. The communication itself to generate the key is assumed to be insecure; the result is that two persons can actually follow such a scheme, literally shout the messages by voice across a room full of people, and nonetheless end sharing a key almost as if it were whispered between each other.

Some of the conjectured hard problems that are currently used to implement classical cryptography are also solvable efficiently by quantum computers [Sho99]. Fortunately, we have learned about this before quantum computers capable of solving current key sizes become a reality, giving time to find a solution. Namely, giving time to find and test problems that we think are also hard for quantum computers, and that we can use to build quantum-secure classical cryptography [Che+16]. At the same time, the same quantum mechanics that makes current classical cryptographic schemes insecure, also give us a way to share perfect keys between two parties, from a distance, without computational assumptions on the eavesdropper [Wie83; Ben+83; BB84; Eke91]. Such schemes are known as quantum key distribution schemes. While classical cryptography with computational security can be thought as allowing the messages to be shouted, quantum key distribution schemes can be thought as going to a noise regime so low, that quantum mechanics allows us to tell if anybody has been listening to the communication, allowing two parties to "whisper" to each other remotely.

Quantum key distribution will be a complementary addition to classical quantum-secure cryptography. Computationally secure schemes suffer from the unavoidable flaw that to be used, a key size needs to be chosen. While a choice can always be made to make the message secure for the foreseeable future, advances in the computational power of computers will eventually break the security of messages encrypted with old instances of cryptographic schemes. Computational security cannot protect forever, but it is the perfect solution to encrypt message for which their confidentiality has an expiry date (like appointments). On the other hand, well implemented quantum key distribution provides perfect security, and thus can secure messages without an expiry date. However, quantum key distribution schemes are affected by costly set-up and high susceptibility to noise, that make the key produced expensive. Once quantum-secure classical cryptography and quantum key distribution will be both available, they will nicely complement each other and allow to choose the best scheme based on the security requirements and cost.

To be able to perform quantum key distribution, two parties need to be able to receive and share entangled states. There are quantum key distribution schemes that use continuous variables, but we will assume that the key is shared by means of finite dimensional quantum systems, the analogue of classical digital systems. To provide its security guarantees, quantum key distribution uses quantum mechanics to bound the information contained in any eavesdropper. The information held by any eavesdropper can then be kept track of, and be shown to be uncorrelated from the output of the scheme. An important contribution to the understanding of quantum key distribution has been the discovery that despite bounding the information of an eavesdropping third party, the problem of distilling key is purely bipartite. In such description the corresponding states of perfect key are a class of entangled states known as private states, an the eavesdropper is simply the environment. This result allowed to determine why some quantum key distribution protocols give larger key rates than then ones achieved by entanglement distillation, even though there are strong connections between key distillation and entanglement distillation. In this thesis we explore this difference further, and investigate what happens to it when performing quantum key distribution in a network.

The first two chapters are dedicated to introducing the basic concepts of finite dimensional quantum systems and of quantum informations. In this chapters we also develop new tools and formalism which are needed to present the results coherently. Most importantly we allow distance measures to be hybrid. Before the present work, all distance measures have been either functions on the states themselves, or functions of outcomes of measurements (where the argument in favour of such measures is that measurement outcomes are the only objects we can access to estimate the distance). We allow distance measures to be computed on states that have a part of measurement outcome and a part of quantum states. We then apply our tools to the study of private states in Chapter 3 where we simplify them, introducing what we call Bell private states and showing that they are a fully equivalent description. This allows to give a precise intuition to all private states as follows. Instead of simply sharing a maximally entangled state, Alice and Bob share a maximally entangled state where a phase error might have been introduced. Usually this would destroy some entanglement, but in private states the information about the error is saved and encoded in orthogonal states also held by Alice and Bob. However these states are bipartite states, and thus do not need to be locally accessible. Assume now that the encoding states are data hiding, meaning that Alice and Bob can barely distinguish them when separated. It is then the intuition that Alice and Bob cannot distill entanglement, since they have poor access to the phase information to be corrected in the entanglement distillation process. We will show cases where this is true and others where the opposite happens, even if the states are data hiding the phase is still correctable. However we show that if the distillable entanglement is reduced then the states must be data hiding in a strong sense, and by choosing the encoding states at random, we can show that the private states must be indistinguishable from states with no entanglement.

Finally in the last chapter, Chapter 4, we apply our findings to quantum repeaters. In a repeater, noisy entanglement is distributed independently between the end points and the repeater, and the repeater station is used to mediate the entanglement between end points. We consider the problem of extending the distance of quantum key distribution with help of intermediate repeater stations. We show that for both private states and protocols that first distill private states, it is essentially optimal to use the standard quantum repeater protocol based on entanglement distillation and entanglement swapping, and thus any excess key in the bipartite setting will be lost. We then distribute private states constructed randomly between the nodes and show that, despite the fact that the private states contain readily extractable perfect key, a broad class of repeater protocols will fail to extract secrecy: the key rate is vanishing for large dimensions, indicating that the states might have small distillable entanglement. These result are indications that the separation between the distillable key and distillable entanglement does not survive in all general network scenarios.

Chapter 1

Preliminaries

In quantum mechanics the state of a complete (or isolated) system is a unit vector in a complex separable Hilbert space. This state is not by itself "observable" (it cannot be seen or measured), it just describes the system and its evolution. The observable objects, commonly just observables (any measurable variable like momentum, energy, mass, position, . . .), are described by different constructs on the Hilbert space.

It needs to be mentioned that, while in classical mechanics the value of observables can be well defined from the state of the system, in general within quantum mechanics it is possible to define the value of an observable only on specific orthogonal states. While observables behave very differently in classical and quantum mechanics, in both cases most observables are continuous, meaning in particular, that they can take uncountably many values. This implies that the dimension of the Hilbert space describing most physical systems is itself uncountably infinite to host uncountably many orthogonal states.

However, in this thesis all quantum systems are finite dimensional. Finite dimensional quantum systems are the quantum counterpart of classical systems with finitely many state, like controllable digital devices. We expect the state of classical digital devices to change only under the control we provide and not under external physical evolution. Similarly we will expect a quantum digital device to change only under the unitaries we provide and not under the physical unitary evolution. Therefore in our quantum systems there is no Hamiltonian, which would usually describe the evolution of the system.

1.1 Linear Algebra

Since quantum systems are described by Hilbert spaces under unitary evolution, the study of finite dimensional quantum systems reduces mathematically to the study of finite dimensional Hilbert spaces and their linear transformations. Thus, many concepts are not exclusive to quantum information theory. Before starting we will recall some of them, many of which can be found in [NC02], and introduce some notation.

1.1.1 Complex Hilbert spaces

For any finite dimensional complex Hilbert space H, we denote with $\langle\cdot|\cdot\rangle$ the complex inner product linear in the second argument, and with $|\mathrm{H}|$ its dimension. Given any subset $\mathcal{K} \subseteq \mathrm{H}$, we denote by $\mathrm{conv}(\mathcal{K})$ the convex hull and by $\overline{\mathrm{conv}}(\mathcal{K})$ its closure. We also define $\mathbb{R}_+\mathcal{K}$ as the cone generated by multiplying the elements of $\mathcal{K}$ with the positive real numbers $\mathbb{R}_+$. Given any symmetric convex subset $\mathcal{K} \subseteq \mathrm{H}$, we denote by $\|x\|_{\mathcal{K}} := \inf\{t : x \in t\mathcal{K}\}$ its gauge (a norm), also known as Minkowski's functional, and by $\mathcal{K}^\circ := \{y : \forall\, x \in \mathcal{K},\ \langle x|y\rangle \leqslant 1\}$ its polar, which satisfies the relations $\|x\|_{\mathcal{K}} = \sup_{y\in\mathcal{K}^\circ} |\,\langle x|y\rangle\,|$ and $\|x\|_{\mathcal{K}^\circ} = \sup_{y\in\mathcal{K}} |\,\langle x|y\rangle\,|$.

1.1.2 Linear operators

The set of linear operators from H to another finite dimensional complex Hilbert space H' is denoted $\mathcal{L}(\mathrm{H},\mathrm{H}')$. For an operator $K \in \mathcal{L}(\mathrm{H},\mathrm{H}')$ we denote with $K^\dagger \in \mathcal{L}(\mathrm{H}',\mathrm{H})$ the adjoint. If $\mathrm{H} = \mathrm{H}'$ we write more concisely $\mathcal{L}(\mathrm{H}) \equiv \mathcal{L}(\mathrm{H},\mathrm{H})$. We denote with $\mathbb{1}_\mathrm{H} \in \mathcal{L}(\mathrm{H})$ the identity operator, but might omit the subscript from time to time, and use simply $\mathbb{1}$ if the system is clear from the context.

An operator $K \in \mathcal{L}(\mathrm{H})$ is said to be normal if $K^\dagger K = KK^\dagger$. It is said to be unitary if $K^{-1} = K^\dagger$, namely $KK^\dagger = K^\dagger K = \mathbb{1}$, and the set of all unitary operators forming the unitary group is denoted $\mathcal{U}(\mathrm{H})$. K is said to be Hermitian if $K^\dagger = K$, and anti-Hermitian if $K^\dagger = -K$; we denote the set of Hermitian operators with $\mathcal{H}(\mathrm{H})$. Finally, an Hermitian operator $K \in \mathcal{H}(\mathrm{H})$ is said to be positive semidefinite (or just positive if clear from the context) if $K = \tilde{K}^\dagger \tilde{K}$ for some $\tilde{K} \in \mathcal{L}(\mathrm{H})$, and we denote the convex cone of positive operators with $\mathcal{H}_+(\mathrm{H})$. As is customary, we write $K \geqslant \tilde{K}$ to denote $K - \tilde{K} \in \mathcal{H}_+(\mathrm{H})$ for any operators $K, \tilde{K} \in \mathcal{L}(\mathrm{H})$. We reserve the term projector for Hermitian projectors, namely positive operators K such that $K^2 = K$.

An Hermitian operator $K \in \mathcal{H}(\mathrm{H})$ can be decomposed into a difference $K = K_+ - K_-$ where $K_\pm$ are positive and orthogonal ($K_+K_- = K_-K_+ = 0$). For any operator $K \in \mathcal{L}(\mathrm{H})$, we denote with $|K|$ the positive operator satisfying $K^\dagger K = |K|^2$. If K is Hermitian then $|K| = K_+ + K_-$.

On $\mathcal{L}(\mathrm{H})$ we have the trace norm $\|K\|_1 = \operatorname{tr}|K|$, the Hilbert-Schmidt norm $\|K\|_2 = \sqrt{\operatorname{tr}|K|^2}$ and the operator norm $\|K\|_\infty = \inf\{t : |K| \leqslant t\mathbb{1}_\mathrm{H}\}$, which are all unitary invariant norms. Furthermore, we write $B_1 \equiv B_1(\mathrm{H})$, $B_2 \equiv B_1(\mathrm{H})$, $B_\infty \equiv B_\infty(\mathrm{H})$ for the corresponding unit balls.

1.1.3 Linear maps (superoperators)

We denote the set of linear operators from $\mathcal{L}(\mathrm{H})$ to $\mathcal{L}(\mathrm{H}')$, from now on called linear maps (but also known as superoperators), with $\mathsf{LM}(\mathrm{H}\rangle\mathrm{H}')$, which is otherwise already defined by $\mathcal{L}(\mathcal{L}(\mathrm{H}),\mathcal{L}(\mathrm{H}'))$. Depending on convenience, like when composing subsets of channels, we will write the same set of linear maps as $\mathsf{LM}(\mathrm{H}'\langle\mathrm{H})$, a convention that we apply also to any subset of the linear maps we define[1]. We denote with $\mathrm{id}_\mathrm{H} \in \mathsf{LM}(\mathrm{H}\rangle\mathrm{H})$ the identity map and with $\operatorname{tr}_\mathrm{H} \in \mathsf{LM}(\mathrm{H}\rangle\mathbb{C})$ the trace, and again we might simply write "id" and "tr" if clear from the context. The space $\mathcal{L}(\mathrm{H})$ is itself a finite dimensional Hilbert space with inner product $\operatorname{tr}\big(K^\dagger\tilde{K}\big)$ for $K, \tilde{K} \in \mathcal{L}(\mathrm{H})$. Any linear map $\Lambda \in \mathsf{LM}(\mathrm{H}\rangle\mathrm{H}')$ can be written (not uniquely) as $\Lambda(Y) = \sum_{i\in I} K_i Y \tilde{K}_i$ for some finite index set I, and operators $K_i \in \mathcal{L}(\mathrm{H},\mathrm{H}')$ and $\tilde{K}_i \in \mathcal{L}(\mathrm{H}',\mathrm{H})$. Its adjoint $\Lambda^\dagger \in \mathsf{LM}(\mathrm{H}'\rangle\mathrm{H})$ is then $\Lambda^\dagger(Y) = \sum_{i\in I} \tilde{K}_i Y K_i$.

We introduce the following notations. Given a linear map $\Lambda \in \mathsf{LM}(\mathrm{H}\rangle\mathrm{H}')$ and a subset $\mathcal{K} \subseteq \mathcal{L}(\mathrm{H})$, we write $\Lambda(\mathcal{K}) \subseteq \mathcal{L}(\mathrm{H}')$ for the set of $\Lambda(K)$ where $K \in \mathcal{K}$. Furthermore, given two subsets of channels $\mathsf{L} \subseteq \mathsf{LM}(\mathrm{H}\rangle\mathrm{H}'')$ and $\mathsf{L}' \subseteq \mathsf{LM}(\mathrm{H}''\rangle\mathrm{H}')$, we define their composition as

$$\mathsf{L}' \circ \mathsf{L} := \left\{\Lambda' \circ \Lambda : \Lambda' \in \mathsf{L}',\ \Lambda \in \mathsf{L}\right\}.$$

For singleton sets, we will write $\Lambda \circ \mathsf{L}'$ and $\mathsf{L} \circ \Lambda'$, instead of $\{\Lambda\} \circ \mathsf{L}'$ and $\mathsf{L} \circ \{\Lambda'\}$. Since we can also think of an operator K as a constant linear map, we abuse our notation and define for $K \in \mathcal{L}(\mathrm{H})$ and any subset of operators $\mathcal{K} \subseteq \mathcal{L}(\mathrm{H})$

$$\begin{aligned}\mathsf{L} \circ K &= \{\Lambda(K) : \Lambda \in \mathsf{L}\}\\ \mathsf{L} \circ \mathcal{K} &= \{\Lambda(K) : \Lambda \in \mathsf{L},\ K \in \mathcal{K}\}.\end{aligned}$$

[1] So we can write the composition, defined in the next paragraph, as $\mathsf{L}(\mathrm{H}''\langle\mathrm{H}') \circ \mathsf{L}(\mathrm{H}'\langle\mathrm{H})$ instead of $\mathsf{L}(\mathrm{H}'\rangle\mathrm{H}'') \circ \mathsf{L}(\mathrm{H}\rangle\mathrm{H}')$. The arrow always points from the input spaces to the output spaces.

A linear map $\Lambda \in \mathsf{LM}(\mathrm{H}\rangle\mathrm{H}')$ is said to be positive if $\Lambda(\mathcal{H}_+(\mathrm{H})) \subseteq \mathcal{H}_+(\mathrm{H}')$. Furthermore, it is said to be trace preserving if $\operatorname{tr}\Lambda(Y) = \operatorname{tr} Y$ for all operators $Y \in \mathcal{L}(\mathrm{H})$. Λ is trace preserving if and only if, when written in the form $\Lambda(Y) = \sum_{i\in I} K_i Y \tilde{K}_i$, it satisfies $\sum_{i\in I} \tilde{K}_i K_i = \mathbb{1}_{\mathrm{H}}$.

1.1.4 Tensor product

For two finite dimensional Hilbert spaces H and H$'$, the finite dimensional Hilbert space $\mathrm{HH}' \equiv \mathrm{H} \otimes \mathrm{H}'$ is the span of $\{x \otimes y : x \in \mathrm{H}, y \in \mathrm{H}'\}$ modded as to make $x \otimes y$ a bilinear operation, with inner product the linear extension of $\langle a \otimes b | x \otimes y\rangle = \langle a|x\rangle \langle b|y\rangle$. Sometimes it will be necessary to make the tensor product of $x \in \mathrm{H}_1\mathrm{H}_1'$ and $y \in \mathrm{H}_2\mathrm{H}_2'$ in $\mathcal{L}(\mathrm{H}_1\mathrm{H}_2\mathrm{H}_1'\mathrm{H}_2')$, we will then add subscripts to make the reordering of $x \otimes y$ implicit, and simply write $x_{\mathrm{H}_1\mathrm{H}_1'} \otimes y_{\mathrm{H}_2\mathrm{H}_2'} \in \mathcal{L}(\mathrm{H}_1\mathrm{H}_2\mathrm{H}_1'\mathrm{H}_2')$.

Because the linear operators also form a Hilbert space, this also defines the tensor product of linear operators $\mathcal{L}(\mathrm{H}_1, \mathrm{H}_1') \otimes \mathcal{L}(\mathrm{H}_2, \mathrm{H}_2')$, and it coincides with the linear operators $\mathcal{L}(\mathrm{H}_1\mathrm{H}_2, \mathrm{H}_1'\mathrm{H}_2')$ on the tensor product spaces. More specifically, the tensor product $K \otimes \tilde{K}$ of two operators coincides with the operator Y defined by the linear extension of $Y(x \otimes y) = Kx \otimes \tilde{K}y$. A particular operator we define on any two spaces H and H$'$ is the swap $S_{\mathrm{HH}'} \in \mathcal{L}(\mathrm{HH}', \mathrm{H}'\mathrm{H})$ as the unique linear extension of $x \otimes y \mapsto y \otimes x$. Again, we might simply write S when clear from the context. Notice that $S_{\mathrm{HH}'}$ is unitary according to the broader definition of $S^\dagger S = \mathbb{1}_{\mathrm{HH}'}$ and $S^\dagger S = \mathbb{1}_{\mathrm{H}'\mathrm{H}}$.

Similarly, we have $\mathsf{LM}(\mathrm{H}_1\rangle\mathrm{H}_1') \otimes \mathsf{LM}(\mathrm{H}_2\rangle\mathrm{H}_2') = \mathsf{LM}(\mathrm{H}_1\mathrm{H}_2\rangle\mathrm{H}_1'\mathrm{H}_2')$, and the tensor product $\Lambda \otimes \Lambda'$ of two linear maps coincides with the linear map Π defined by the linear extension of $\Pi(K \otimes \tilde{K}) = \Lambda(K) \otimes \Lambda'(\tilde{K})$. Most notably, we have $\mathrm{id}_{\mathrm{HH}'} = \mathrm{id}_{\mathrm{H}} \otimes \mathrm{id}_{\mathrm{H}'}$ and $\mathrm{tr}_{\mathrm{HH}'} = \mathrm{tr}_{\mathrm{H}} \otimes \mathrm{tr}_{\mathrm{H}'}$. Extending the notation for the composition of subsets of linear maps, we define the tensor product of $\mathsf{L} \subseteq \mathsf{LM}(\mathrm{H}\rangle\mathrm{H}'')$ and $\mathsf{L}' \subseteq \mathsf{LM}(\mathrm{H}''\rangle\mathrm{H}')$ as

$$\mathsf{L} \otimes \mathsf{L}' := \left\{\Lambda \otimes \Lambda' : \Lambda \in \mathsf{CPTP}, \Lambda' \in \mathsf{L}'\right\},$$

We will also just write $\Lambda \otimes \mathsf{L}'$ and $\mathsf{L} \otimes \Lambda'$ instead of $\{\Lambda\} \otimes \mathsf{L}'$ and $\mathsf{L} \otimes \{\Lambda'\}$.

1.1.5 Complete positivity

Finally, a map $\Lambda \in \mathsf{LM}(\mathrm{H}\rangle\mathrm{H}')$ is said to be completely positive if $\mathrm{id}_{\mathrm{H}''} \otimes \Lambda$ is positive for any other finite dimensional Hilbert space H$''$, which holds if and only if Λ can be written (again not uniquely) in the form $\Lambda(Y) = \sum_{i\in I} K_i Y K_i^\dagger$, where the operators K_i are known as Kraus operators. We denote with $\mathsf{CP}(\mathrm{H}\rangle\mathrm{H}')$ the maps that are completely positive. We also denote with $\mathsf{CPTP}(\mathrm{H}\rangle\mathrm{H}')$ the maps that are completely positive and trace preserving, which notably includes the identity map and trace map. Again, these properties are preserved under composition, which we can write using our notation as $\mathsf{CP}(\mathrm{H}'\langle\mathrm{H}) \supseteq \mathsf{CP}(\mathrm{H}'\langle\mathrm{H}'') \circ \mathsf{CP}(\mathrm{H}''\langle\mathrm{H})$ and $\mathsf{CPTP}(\mathrm{H}'\langle\mathrm{H}) \supseteq \mathsf{CPTP}(\mathrm{H}'\langle\mathrm{H}'') \circ \mathsf{CPTP}(\mathrm{H}''\langle\mathrm{H})$. Stinespring's dilation theorem shows that Λ is completely positive and trace preserving (CPTP) if and only if there exist another system H$''$ with an isometry $\tilde{K} \in \mathcal{L}(\mathrm{H}, \mathrm{H}''\mathrm{H}')$ (namely $\tilde{K}^\dagger\tilde{K} = \mathbb{1}_{\mathrm{H}}$) such that $\Lambda(Y) = \mathrm{tr}_{\mathrm{H}''} \otimes \mathrm{id}_{\mathrm{H}'}(\tilde{K}Y\tilde{K}^\dagger)$; the choice of H$''$ and the isometry are not unique.

1.1.6 Bra-ket notation

In Dirac's bra-ket notation for a Hilbert space H, the vectors are written as $|\alpha\rangle \in \mathrm{H}$, and they are actually treated as linear operators, more precisely as elements of $\mathcal{L}(\mathbb{C}, \mathrm{H})$ acting as $z \mapsto z\,|\alpha\rangle$. Then $\langle\alpha| = (|\alpha\rangle)^\dagger$, namely they are the adjoint of $|\alpha\rangle$, equivalently they are the linear operators $\langle\alpha| : |\beta\rangle \to \langle\alpha|\beta\rangle$. The outer product is then denoted with $|\alpha\rangle\langle\beta|$, and the set of linear operators $\mathcal{L}(\mathrm{H}, \mathrm{H}')$ is then conveniently described by the span of all

$|\alpha\rangle\langle\beta|$ for $|\alpha\rangle \in \mathrm{H}'$ and $|\beta\rangle \in \mathrm{H}$. The inner product then gains an additional notation as $\langle\beta|\,|\alpha\rangle \equiv \langle\beta|\alpha\rangle$, which is especially useful when considering other linear operators. Let $|\alpha\rangle \in \mathrm{H}$, $|\beta\rangle \in \mathrm{H}'$ and $K \in \mathcal{L}(\mathrm{H}, \mathrm{H}')$. Then in particular $\langle\beta| K |\alpha\rangle$ is both the inner product in H' of $|\beta\rangle$ and $K|\alpha\rangle$, and the innerproduct in H of $K^\dagger |\beta\rangle$ and $|\alpha\rangle$. The tensor product can additionally be denoted as $|\alpha\beta\rangle \equiv |\alpha\rangle\,|\beta\rangle \equiv |\alpha\rangle \otimes |\beta\rangle$.

Any operator $K \in \mathcal{L}(\mathrm{H}, \mathrm{H}')$ admits a singular value decomposition, namely there exist an integer $k \leqslant \min\{|\mathrm{H}|, |\mathrm{H}'|\}$ (the rank), singular values $s_i > 0$ for $i = 1, \dots, k$, and orthonormal vectors $\{|r_i\rangle\}_{i=1}^{k} \subset \mathrm{H}$, $\{|l_i\rangle\}_{i=1}^{k} \subset \mathrm{H}'$, so that the operator can be written as $K = \sum_{i=1}^{k} s_i\, |l_i\rangle\langle r_i|$.

We denote by $\mathbb{Z}_n = \mathbb{Z}/n\mathbb{Z} \equiv \{0, \dots, n-1\}$ the integers with addition modulo n, which we use to index orthonormal bases in H. An operator K is unitary diagonalizable if and only if it is normal, in which case there exist eigenvalues $e_i \in \mathbb{C}$ for $i \in \mathbb{Z}_{|\mathrm{H}|}$ and a basis $\{|v_i\rangle\}_{i\in\mathbb{Z}_{|\mathrm{H}|}} \subset \mathrm{H}$, such that $T = \sum_{i\in\mathbb{Z}_{|\mathrm{H}|}} e_i\, |v_i\rangle\langle v_i|$. In particular, Hermitian operators are diagonalizable with real eigenvalues, positive operators are diagonalizable with positive eigenvalues, and unitary operators are diagonalizable with phase eigenvalues (eigenvalues in the complex unit circle). If two normal operators commute, then they can be diagonalized using a common basis.

1.1.7 Computational basis

Some definitions are basis dependent. We call the computational basis the choice of orthonormal basis $\{|i\rangle_{\mathrm{H}}\}_{i\in\mathbb{Z}_{|\mathrm{H}|}} \subset \mathrm{H}$, as customary in quantum information, with bra denoted $\langle i|_{\mathrm{H}}$ and operators denoted $|i\rangle\langle j|_{\mathrm{H}}$. For an operator $K \in \mathcal{L}(\mathrm{H})$ we denote with $\overline{K}$ the complex conjugate on this basis. We also denote with $K^{\mathrm{T_H}}$ the transpose, also basis dependent. When needed, we will denote the transposition map $\vartheta_{\mathrm{H}} \in \mathcal{L}(\mathrm{H})$. In the case of multiple systems like HH' we fix the computational basis to be the tensor product of the computational basis of H and H', namely $\{|ij\rangle_{\mathrm{HH}'} = |i\rangle_{\mathrm{H}} \otimes |j\rangle_{\mathrm{H}}\}_{i\in\mathbb{Z}_{|\mathrm{H}|}, j\in\mathbb{Z}_{|\mathrm{H}'|}}$. Then the transposition also satisfies $\vartheta_{\mathrm{H}} \otimes \vartheta_{\mathrm{H}'} = \vartheta_{\mathrm{HH}'}$, is its own adjoint and is positive but not completely positive. Also, for the computational basis and the transposition, it is common to omit the subscript when clear from the context.

1.2 Quantum systems

A finite quantum system is any finite dimensional complex Hilbert space H. If the system is completely isolated and has never interacted with any other system, then quantum mechanics dictates that system state is completely described by a unit vector in H. Multiple systems are described by unit vectors in the tensor product of their Hilbert spaces. As it is customary in quantum mechanics, we follow Dirac's bra-ket notation (only) for the Hilbert space of the quantum system. However, the definition is not flexible, in the sense that it only describes such complete systems. If $|\psi\rangle \in \mathrm{HH}'$ is the state describing the joint system of H and H', it is in general not possible to find an $|\alpha\rangle \in \mathrm{H}$ and $|\beta\rangle \in \mathrm{H}'$ that describes the state of the single systems. For this reason, the unit vector states like $|\psi\rangle$ are usually called pure states. At the same time, the only physically relevant quantities in quantum mechanics (because they are the only testable ones) are expectation values, namely quantities of the form $\langle\psi| K |\psi\rangle = \mathrm{tr}(|\psi\rangle\langle\psi|\, K)$ for some operator $K \in \mathcal{L}(\mathrm{HH}')$. The expectation value is said to be only of system H', if $K = \mathbb{1}_{\mathrm{H}} \otimes \tilde{K}$ for some $\tilde{K} \in \mathcal{L}(\mathrm{H}')$. Then $\mathrm{tr}_{\mathrm{HH}'}(|\psi\rangle\langle\psi|\, K) = \mathrm{tr}_{\mathrm{H}'}((\mathrm{tr}_{\mathrm{H}} \otimes \mathrm{id}_{\mathrm{H}'}\, |\psi\rangle\langle\psi|)\tilde{K})$, and therefore $\mathrm{tr}_{\mathrm{H}} \otimes \mathrm{id}_{\mathrm{H}'}\, |\psi\rangle\langle\psi|$ is regarded as the state of the single system H', as it gives the correct expectation values even if there is no pure state. Since the trace is a completely-positive trace-preserving map, $\mathrm{tr}_{\mathrm{H}} \otimes \mathrm{id}_{\mathrm{H}'}\, |\psi\rangle\langle\psi|$ is a positive operator of unit trace, just like $|\psi\rangle\langle\psi|$, also known as a density matrix.

The evolution of a pure quantum state predicted by quantum mechanics is always the action of a unitary operator. Computing the density matrix of the new system results in a state which is the output of a CPTP map on the original state. At the same time, we can interpret Stinespring's dilation theorem, as saying that any CPTP map is the result of the unitary evolution of the density matrix in a larger system. Thus CPTP maps produce valid states that correctly describe the statistics resulting from possibly unknown interaction in a larger system.

Therefore we lose no power in describing quantum systems with density matrices, while at the same time we gain the ability of describing incomplete systems[2]. As is common in quantum information, we thus abstract from the physical axioms of quantum mechanics using the formalism described below.

1.2.1 States and channels

A state on a quantum system H is a positive-semidefinite operator with unit trace. The set of all states is closed and convex and is denoted by $\mathcal{D}(\mathrm{H}) \subset \mathcal{H}_+(\mathrm{H})$. A state is said to be pure if it is a rank-1 projector $|\psi\rangle\langle\psi|$, otherwise it is said to be mixed. We say that a state is a uniform mixture on a subspace if it is of the form $P/\operatorname{tr} P$, where P is the projector on said subspace. The uniform mixture on the whole space $\mathbb{1}/|\mathrm{H}|$ is called the maximally mixed state. As customary, on multiple systems like HH' we may omit the identity map when tensored with the trace, namely we can write $\operatorname{tr}_{\mathrm{H}} \equiv \operatorname{tr}_{\mathrm{H}} \otimes \mathbb{1}_{\mathrm{H}'}$. Similarly, we can omit the system of the trace when acting on a whole state, so that $\operatorname{tr}(\operatorname{tr}_{\mathrm{H}}(\rho)) \equiv \operatorname{tr}_{\mathrm{H}'}(\operatorname{tr}_{\mathrm{H}} \otimes \mathrm{id}_{\mathrm{H}'}(\rho))$ for a state $\rho \in \mathcal{D}(\mathrm{HH}')$. The state $\operatorname{tr}_{\mathrm{H}} \rho$ is called the reduced state, or marginal, of ρ on H', and, when useful, we use the convention of writing it even more concisely as $\rho_{\mathrm{H}'} \equiv \operatorname{tr}_{\mathrm{H}} \rho$; the original state can then be recovered by explicitly specifying all the systems as $\rho_{\mathrm{HH}'}$.

A channel is a completely-positive trace-preserving map, and the set $\mathsf{CPTP}(\mathrm{H}\rangle\mathrm{H}')$ of channels is all the allowed operations from system H to system H'. Because channels are then linear maps, the first consequence is that there exists no channel that implements $\rho \mapsto \rho \otimes \rho$ for all states $\rho \in \mathcal{D}(\mathrm{H})$, which is clearly quadratic and not linear in ρ; this is known as the no-cloning theorem [Par70; WZ82]. When defining channels, we will usually define them implicitly only on states, but because the span of $\mathcal{D}(\mathrm{H})$ is $\mathcal{H}(\mathrm{H})$, by linear extension this defines it on all Hermitian operators.

A mixed unitary channel $\Lambda \in \mathsf{CPTP}(\mathrm{H}\rangle\mathrm{H})$ is a channel of the form

$$\Lambda(\rho) = \frac{1}{|I|} \sum_{i \in I} U_i \rho U_i^\dagger$$

where I is some finite index set and $\{U_i\}_{i \in I}$ is a set of unitaries that are sampled uniformly at random. If the set of unitaries forms a subgroup, then the channel is called a twirl and it maps into the set of states invariant under the twirl. Let now $\rho \in \mathcal{D}(\mathrm{CT})$, we say that a unitary $U \in \mathcal{U}(\mathrm{CT})$ is a controlled unitary if it is of the form $U = \sum_{i \in \mathbb{Z}_{|\mathrm{C}|}} |i\rangle\langle i|_{\mathrm{C}} \otimes U_{i,\mathrm{T}}$. We then call C the control system and T the target system, and call controlled unitary channel the resulting unitary conjugation.

A dephasing is a channel $\Lambda \in \mathsf{CPTP}(\mathrm{H}\rangle\mathrm{H})$, also known as pinching, of the form

$$\Lambda(\rho) = \sum_{i \in I'} P_i \rho P_i$$

where $\{P_i\}_{i \in I'}$ are projectors (and thus an orthogonal partition of the identity). Finally, when dealing with multiple systems $\mathrm{H}_1\mathrm{H}_2$, we call local channels those channels that act of the systems independently, namely any $\Lambda \in \mathsf{CPTP}(\mathrm{H}_1\rangle\mathrm{H}_1') \otimes \mathsf{CPTP}(\mathrm{H}_2\rangle\mathrm{H}_2')$.

[2] A pure state is described equivalently by $e^{i\phi}\,|\psi\rangle$ as well as $|\psi\rangle$ (because the expectations values are the same), even though these might be different vectors in the Hilbert space. This is automatically taken care of in the density matrix, as the resulting density matrix is $|\psi\rangle\langle\psi|$ in both cases.

1.2.2 Instruments

Consider systems H, H′, M, and let $\{|i\rangle_{\mathrm{M}}\}_{i\in\mathbb{Z}_{|\mathrm{M}|}}$ be a choice of computational basis. A quantum instrument is a channel $\Lambda \in \mathsf{CPTP}(\mathrm{H}\rangle\mathrm{H}'\mathrm{M})$ of the form

$$\Lambda(\rho) = \sum_{i\in\mathbb{Z}_{|\mathrm{M}|}} \Lambda_i(\rho) \otimes |i\rangle\langle i|_{\mathrm{M}}$$

where Λ_i are completely positive maps, namely $\Lambda_i \in \mathsf{CP}(\mathrm{H}\rangle\mathrm{H}')$, that can be thought as partial channels. The constraint of Λ being a channel, imposes the sum of Λ_i being also a channel (as it is easily checked by tracing out M). The projectors on the computational basis are to be thought of as those states that can be stored in a classical memory instead of the quantum system itself. Instruments are hybrid channels, in the sense that they contain a quantum and a classical part.

We denote the set of all instruments as $\mathsf{IN}(\mathrm{H}\rangle\mathrm{H}'\underline{\mathrm{M}})$[3] indicating with $\underline{\mathrm{M}}$ the fact that the output on M for any channel in IN is always classical, irrespective of the input to the channel (in particular $\underline{\cdot}$ does not denote a different system, but merely a property of the subset of channels IN).

1.2.3 Purification and extensions

Given a state $\rho \in \mathcal{D}(\mathrm{H})$, we call an extension of ρ any state $\rho' \in \mathcal{D}(\mathrm{HH}')$ such that $\rho = \mathrm{tr}_{\mathrm{H}'}\,\rho'$. If the extension is a pure state, we call it a purification. A purification can always be made by taking the eigenvalue decomposition $\rho = \sum_{i\in\mathbb{Z}_{|\mathrm{H}|}} p_i\,|v_i\rangle\langle v_i|$ and constructing $|\rho\rangle\langle\rho| \in \mathrm{HH}$ as $|\rho\rangle\langle\rho| = \sum_{ij\in\mathbb{Z}_{|\mathrm{H}|}} \sqrt{p_i p_k}\,|v_i\,v_i\rangle\langle v_k\,v_k|$. Purifications are not unique, but they are related by local isometries on the purifying system. Furthermore, any purification of a pure state is a product state.

Given a channel $\Lambda \in \mathsf{CPTP}(\mathrm{H}\rangle\mathrm{H}')$, we call a purification of Λ, any channel $\Pi \in \mathsf{CPTP}(\mathrm{H}\rangle\mathrm{H}'\mathrm{H}'')$ such that $\Lambda = \mathrm{tr}_{\mathrm{H}''} \circ \Pi$ and $\Pi(\rho) = K\rho K^\dagger$ for some isometry $K \in \mathcal{L}(\mathrm{H}, \mathrm{H}'\mathrm{H}'')$. A purification can always be made using Stinespring's dilation theorem.

1.2.4 Unspecified output systems

Quite often we will be optimizing over channels that do not have a specific output dimension, or equivalently we will be optimizing both over channels and over the output systems. To deal with such different channels together (when possible), we will often define unions of channels with different outputs spaces. This will of course lose any linear structure on these sets, and we will not be able to sensibly define their composition. Our first such definition are the set of all channels on H

$$\mathsf{CPTP}(\mathrm{H}) := \bigcup_{\mathrm{H}'} \mathsf{CPTP}(\mathrm{H}\rangle\mathrm{H}').$$

Notice that defining the set of all instruments makes no sense, as it will include all channels already for $|\mathrm{M}| = 1$.

1.3 Bipartite systems

While we have already seen multiple quantum systems in the previous section, the purpose of this section is to introduce the models for the interaction of systems that are spatially separated. If two systems A and B are spatially separated from each other then

[3] Note that in [2], we use H for systems that are measured and $\underline{\mathrm{H}}$ for systems that stay quantum.

there are limitations on the operations that can be performed, coming from the physical principle that all interactions must be local. To perform any global operations, the two systems must be brought together first, which requires quantum communication. We will see that global operations can also be mediated by a third quantum system, however this only shifts the problem of the quantum communication to the third system. Even without quantum communication though, classical communication can allow for non trivial interactions, while being much easier to implement (the outcome of a measurement can be communicated to condition an operation on the other systems, or the operations could depend on some common classical information). As mentioned, the purpose of this section is to introduce the models of interaction that we will use, together with the operations that can be done by spatially separated systems using classical communication. To do this, it is convenient to introduce a layer of abstraction through the concept of parties, among which all the systems are distributed. The parties are usually labelled with names such as Alice, Bob, Charlie, etc. We then separate the systems from different parties in a colon separated list, as A:B, AA′:BB′, A:C:B, etc. In this section we consider systems of two parties (bipartite systems), but most of the contents are straightforwardly generalized to multipartite systems.

1.3.1 Separability and the PPT criterion

Let A:B be a bipartite system of Alice and Bob. We have already classified the computational basis state $|i\rangle\langle i|$ as the classical ones. A state (up to normalization) like $\sum_i |ii\rangle\langle ii|_{\mathrm{AB}}$ thus represents shared classical information between Alice and Bob. If Alice and Bob are separated and decide to produce a new quantum state, they will only be able to perform local channels to produce the quantum state, but they can do this conditioned on some shared classical state like the above. The resulting state will not be classical, but we say it will be separable. The separable states of a bipartite system are defined as

$$\mathcal{S}(\mathrm{A}{:}\mathrm{B}) := \operatorname{conv}\{\rho\otimes\sigma : \rho\in\mathcal{D}(\mathrm{A}),\ \sigma\in\mathcal{D}(\mathrm{B})\},$$

If a state is not separable, it is said to be entangled. If Alice and Bob start with a separable state, then any operation they do while spatially separated and unable to send each other quantum states will result in another separable state. We call separable operations [Rai99b] the channels with this property, namely a channel $\Lambda \in \mathsf{CPTP}(\mathrm{A_{in}}{:}\mathrm{B_{in}}\rangle \mathrm{A_{out}}{:}\mathrm{B_{out}})$, is separable if it maps separable states into separable states even when applied to arbitrary subsystems [Rai01], namely if for all systems A and B it satisfies

$$\mathrm{id}_{\mathrm{AB}}\otimes\Lambda(\mathcal{S}(\mathrm{AA_{in}}{:}\mathrm{BB_{in}}))\subseteq\mathcal{S}(\mathrm{AA_{out}}{:}\mathrm{BB_{out}}).$$

Such a condition is satisfied if and only if Λ can be written with product Kraus operators, namely if and only if there exist $K_i \in \mathcal{L}(\mathrm{A_{in}},\mathrm{A_{out}})$ and $\tilde{K}_i \in \mathcal{L}(\mathrm{B_{in}},\mathrm{B_{out}})$ over some finite index $i\in I$ such that

$$\Lambda(\rho) = \sum_{i\in I}(K\otimes\tilde{K})\rho(K\otimes\tilde{K})^\dagger.$$

We denote the set of separable operations with $\mathsf{SEP}(\mathrm{A_{in}}{:}\mathrm{B_{in}}\rangle\mathrm{A_{out}}{:}\mathrm{B_{out}})$, and the set of all separable operations on A and B as

$$\mathsf{SEP}(\mathrm{A}{:}\mathrm{B}) := \bigcup_{\mathrm{A_{out}},\mathrm{B_{out}}}\mathsf{SEP}(\mathrm{A}{:}\mathrm{B}\rangle\mathrm{A_{out}}{:}\mathrm{B_{out}}).$$

Generally checking whether any state is separable is a hard problem [Gur03], meaning that we have no easily computable criterion that is necessary and sufficient to determine separability. However, positive maps that are not completely positive are easily computable necessary criterions: for any separable state $\rho \in \mathcal{S}(\mathrm{A{:}B})$ it is easily checked that any positive map $\Lambda \in \mathsf{P}(\mathrm{A}\rangle\mathrm{A}')$ produces a positive-semidefinite operator, namely $\Lambda \otimes \mathrm{id}_{\mathrm{B}}(\mathcal{S}(\mathrm{A{:}B})) \subseteq \mathcal{H}_+(\mathrm{A'B})$. Checking for separability of ρ is actually equivalent to checking that all positive maps produce a valid state [HHH01], however already determining separability within an error that decreases with the dimension of the system requires intractably many positive maps [AS17]. Still, knowing as many positive maps outside of CPTP as possible can only improve our understanding of separable states.

Within such positive maps, the partial transpose is of particular interest because, among other reasons that will be mentioned later, it gives a criterion that is preserved by separable operations. For bipartite systems we reserve a special notation to denote the action of the transposition. For any bipartite system A:B we define the partial transpose of a state $\rho \in \mathcal{D}(\mathrm{AB})$ as $\rho^{\Gamma} := \mathrm{id}_{\mathrm{A}} \otimes \vartheta_{\mathrm{B}}(\rho)$. Because the transpose is an involution, the choice to act on the second party is without loss of generality. If $\rho^{\Gamma} \geqslant 0$ then we say that ρ is PPT (positive under partial transposition), otherwise we say it is NPT (non-positive under partial transposition), and we define $\mathcal{P}(\mathrm{A{:}B})$ to be the set of all PPT states. Since if ρ is PPT then ρ^{Γ} is also a valid state, we have:

$$\mathcal{P}(\mathrm{A{:}B}) = \{\rho \in \mathcal{D}(\mathrm{AB}) : \rho^{\Gamma} \in \mathcal{D}(\mathrm{AB})\} = \mathcal{D}(\mathrm{AB}) \cap \mathcal{D}(\mathrm{AB})^{\Gamma},$$

which is also a closed convex set. Most importantly, while the partial transpose itself is basis dependent, the PPT condition, and thus the set of PPT states, is not. The condition is also preserved under the trace. As for separable states, we define PPT operations [Rai99b] as those channels $\Lambda{\in}\mathsf{CPTP}(\mathrm{A_{in}{:}B_{in}}\rangle\mathrm{A_{out}{:}B_{out}})$ that preserve the PPT property, even when applied to arbitrary subsystems, namely if for any systems A and B they satisfy

$$\mathrm{id}_{\mathrm{AB}} \otimes \Lambda(\mathcal{P}(\mathrm{AA_{in}{:}BB_{in}})) \subseteq \mathcal{P}(\mathrm{AA_{out}{:}BB_{out}}).$$

We denote with $\mathsf{PPT}(\mathrm{A_{in}{:}B_{in}}\rangle\mathrm{A_{out}{:}B_{out}})$ the set of PPT operations, and with

$$\mathsf{PPT}(\mathrm{A{:}B}) := \bigcup_{\mathrm{A_{out}},\mathrm{B_{out}}} \mathsf{PPT}(\mathrm{A{:}B}\rangle\mathrm{A_{out}{:}B_{out}})$$

the set of all PPT operations on A and B, like for separable operations.

In low dimensions the PPT criterion is also sufficient, more precisely it was shown in [HHH01] that $\mathcal{P}(\mathbb{C}^2{:}\mathbb{C}^2) = \mathcal{S}(\mathbb{C}^2{:}\mathbb{C}^2)$ and $\mathcal{P}(\mathbb{C}^2{:}\mathbb{C}^3) = \mathcal{S}(\mathbb{C}^2{:}\mathbb{C}^3)$. However in other dimensions they are different, with the first examples of entangled PPT states being presented in [Tan86; Hor97]. Since then we have learned of situations, some of which will be explained in this thesis, where separability and PPTness behave much alike, and others in which they behave in disparate ways.

Other bipartite channels

Notice that the definitions of separable and PPT operations have the same structure as the definition of completely positive maps. In [BP10] the authors defined the larger sets of separability-preserving operations and PPT-preserving operations as those channels $\Lambda{\in}\mathsf{CPTP}(\mathrm{A_{in}{:}B_{in}}\rangle\mathrm{A_{out}{:}B_{out}})$ satisfying

$$\Lambda(\mathcal{S}(\mathrm{A_{in}{:}B_{in}})) \subseteq \mathcal{S}(\mathrm{A_{out}{:}B_{out}})$$

and

$$\Lambda(\mathcal{P}(\mathrm{A_{in}{:}B_{in}})) \subseteq \mathcal{P}(\mathrm{A_{out}{:}B_{out}}),$$

respectively. Separable and PPT operations can then be thought as the completely separability-preserving and the completely PPT-preserving operations. An example of separability- and PPT-preserving operations that are neither separable or PPT, are the swap and any global entangled measurement.

If in the conditions above we impose that all input states are to be mapped into separable or PPT states, we get entanglement-annihilating [MZ10] and PPT-inducing channels [Fil14]. More precisely, entanglement-annihilating channels are those channels $\Lambda \in \mathsf{CPTP}(\mathrm{A_{in}{:}B_{in}} \rangle \mathrm{A_{out}{:}B_{out}})$ for which

$$\Lambda(\mathcal{D}(\mathrm{A_{in}{:}B_{in}})) \subseteq \mathcal{S}(\mathrm{A_{out}{:}B_{out}}).$$

and PPT-inducing channels are the ones for which

$$\Lambda(\mathcal{D}(\mathrm{A_{in}{:}B_{in}})) \subseteq \mathcal{P}(\mathrm{A_{out}{:}B_{out}}).$$

Notice that there are no "completely" entanglement-annihilating or PPT-inducing channels. Entanglement-annihilating and PPT-inducing are not to be confused with entanglement-breaking [HSR03] and PPT-binding channels [HHH00]. Entanglement-breaking channels are those channels $\Lambda \in \mathsf{CPTP}(\mathrm{B_{in}} \rangle \mathrm{B_{out}})$ for which

$$\mathrm{id_A} \otimes \Lambda(\mathcal{D}(\mathrm{AB_{in}})) \subseteq \mathcal{S}(\mathrm{A{:}B_{out}})$$

and PPT-inducing channels are the ones for which

$$\mathrm{id_A} \otimes \Lambda(\mathcal{D}(\mathrm{AB_{in}})) \subseteq \mathcal{P}(\mathrm{A{:}B_{out}}).$$

Namely, these channels impose the separability and PPT condition with respect to outer systems.

1.3.2 Local Operations and Classical Communication

Separable operations, though, are not the channels that describe what operations spatially separated parties can do. The difference is subtle but it has physical consequences. Without any interactions, two parties can always perform their operations independently of one another, we call this local operations, which is simply all the local channels on the system of the two parties:

$$\mathsf{LO}(\mathrm{A_{in}{:}B_{in}} \rangle \mathrm{A_{out}{:}B_{out}}) := \mathsf{CPTP}(\mathrm{A_{in}} \rangle \mathrm{A_{out}}) \otimes \mathsf{CPTP}(\mathrm{B_{in}} \rangle \mathrm{B_{out}}),$$

and

$$\mathsf{LO}(\mathrm{A{:}B}) := \bigcup_{\mathrm{A_{out}B_{out}}} \mathsf{LO}(\mathrm{A{:}B} \rangle \mathrm{A_{out}{:}B_{out}}).$$

We will now allow Alice and Bob to send classical messages to each other. We just need a description of channels that have a classical part, and those are the quantum instruments.

One-way LOCC

First we allow Alice to send messages to Bob after his local operations, and allow Bob to do his local operations conditioned on the messages, but we will not allow Bob to send messages to Alice; in this case the communication is said to be one way. To introduce the classical communication between two parties, we simply give to one party the classical part of an instrument on the other. We can thus define a single round of local operation with classical communication (LOCC) from Alice to Bob as:

$$\mathsf{LOCC}_{\rightarrow}(\mathrm{A_{out}{:}B_{out}} \langle \mathrm{A_{in}{:}B_{in}}) := \bigcup_{\mathrm{M}} [\mathrm{id}_{\mathrm{A_{out}}} \otimes \mathsf{CPTP}(\mathrm{B_{out}} \langle \mathrm{MB_{in}})] \circ [\mathsf{IN}(\mathrm{A_{out}}\underline{\mathrm{M}} \langle \mathrm{A_{in}}) \otimes \mathrm{id}_{\mathrm{B_{in}}}],$$

Similarly if the message is from Bob to Alice we define

$$\mathsf{LOCC}_{\leftarrow}(\mathrm{A}_{\mathrm{out}}{:}\mathrm{B}_{\mathrm{out}}\langle \mathrm{A}_{\mathrm{in}}{:}\mathrm{B}_{\mathrm{in}}) := \bigcup_{\mathrm{M}} [\mathsf{CPTP}(\mathrm{A}_{\mathrm{out}}\langle \mathrm{A}_{\mathrm{in}}\mathrm{M}) \otimes \mathrm{id}_{\mathrm{B}_{\mathrm{out}}}] \circ [\mathrm{id}_{\mathrm{A}_{\mathrm{in}}} \otimes \mathsf{IN}(\underline{\mathrm{M}}\mathrm{B}_{\mathrm{out}}\langle \mathrm{B}_{\mathrm{in}})].$$

We then have the sets of all one-way LOCC operations on A and B:

$$\begin{aligned} \mathsf{LOCC}_{\rightarrow}(\mathrm{A}{:}\mathrm{B}) &:= \bigcup_{\mathrm{A}_{\mathrm{out}},\mathrm{B}_{\mathrm{out}}} \mathsf{LOCC}_{\rightarrow}(\mathrm{A}{:}\mathrm{B}\rangle \mathrm{A}_{\mathrm{out}}{:}\mathrm{B}_{\mathrm{out}}) \\ \mathsf{LOCC}_{\leftarrow}(\mathrm{A}{:}\mathrm{B}) &:= \bigcup_{\mathrm{A}_{\mathrm{out}},\mathrm{B}_{\mathrm{out}}} \mathsf{LOCC}_{\leftarrow}(\mathrm{A}{:}\mathrm{B}\rangle \mathrm{A}_{\mathrm{out}}{:}\mathrm{B}_{\mathrm{out}}). \end{aligned}$$

Other than to define general LOCC below, without loss of generality it is normally enough to deal only with $\mathsf{LOCC}_{\rightarrow}$.

We could envision the parties sending multiple rounds of messages, but if the communication is restricted to be only one-way, then a single round actually defines all the one-way LOCC channels [DW05]. The reason is, since the operations of the sender do not depend on the ones of the receiver, the receiver can just wait until all the messages have arrived. This also means that there is no reason for the sender to perform its instruments in multiple steps. This is particularly relevant practically, because it means that if the sender's final outcome is purely classical, there is no requirement of keeping the quantum part of the instrument alive. Because keeping the quantum information alive is currently challenging, having to wait for two-way communication is sometimes undesirable, and current proposals for quantum repeaters are shifting toward purely one-way operations [Mur+16].

Two-way LOCC

If we allow two-way classical communication, then there is no such simplification on the number of rounds. First we define a single round of two-way LOCC as:

$$\mathsf{LOCC}^{1}(\mathrm{A}_{\mathrm{out}}{:}\mathrm{B}_{\mathrm{out}}\langle \mathrm{A}_{\mathrm{in}}{:}\mathrm{B}_{\mathrm{in}}) := \bigcup_{\mathrm{A},\mathrm{B}} \mathsf{LOCC}_{\leftarrow}(\mathrm{A}_{\mathrm{out}}{:}\mathrm{B}_{\mathrm{out}}\langle \mathrm{A}{:}\mathrm{B}) \circ \mathsf{LOCC}_{\rightarrow}(\mathrm{A}{:}\mathrm{B}\langle \mathrm{A}_{\mathrm{in}}{:}\mathrm{B}_{\mathrm{in}}).$$

and then define n-rounds of two-way LOCC recursively as

$$\mathsf{LOCC}^{n}(\mathrm{A}_{\mathrm{out}}{:}\mathrm{B}_{\mathrm{out}}\langle \mathrm{A}_{\mathrm{in}}{:}\mathrm{B}_{\mathrm{in}}) := \bigcup_{\mathrm{A},\mathrm{B}} \mathsf{LOCC}^{1}(\mathrm{A}_{\mathrm{out}}{:}\mathrm{B}_{\mathrm{out}}\langle \mathrm{A}{:}\mathrm{B}) \circ \mathsf{LOCC}^{n-1}(\mathrm{A}{:}\mathrm{B}\langle \mathrm{A}_{\mathrm{in}}{:}\mathrm{B}_{\mathrm{in}}).$$

Notice that this construction in no way imposes Alice or Bob to be the first one to perform a non-trivial operation, because the identity channel is included in the instruments. Having Bob be the first instead of Alice simply changes the number of rounds by at most one. Finally, LOCC can be defined as [Rai99b]

$$\mathsf{LOCC}(\mathrm{A}_{\mathrm{in}}{:}\mathrm{B}_{\mathrm{in}}\rangle \mathrm{A}_{\mathrm{out}}{:}\mathrm{B}_{\mathrm{out}}) := \bigcup_{n\in\mathbb{N}} \mathsf{LOCC}^{n}(\mathrm{A}_{\mathrm{in}}{:}\mathrm{B}_{\mathrm{in}}\rangle \mathrm{A}_{\mathrm{out}}{:}\mathrm{B}_{\mathrm{out}}),$$

together with

$$\mathsf{LOCC}(\mathrm{A}{:}\mathrm{B}) := \bigcup_{\mathrm{A}_{\mathrm{out}},\mathrm{B}_{\mathrm{out}}} \mathsf{LOCC}(\mathrm{A}{:}\mathrm{B}\rangle \mathrm{A}_{\mathrm{out}}{:}\mathrm{B}_{\mathrm{out}}).$$

Hierarchy. By construction we have:

$$\mathsf{LO}(\mathrm{A}{:}\mathrm{B}) \subseteq \mathsf{LOCC}_{\rightarrow}(\mathrm{A}{:}\mathrm{B}) \subseteq \mathsf{LOCC}(\mathrm{A}{:}\mathrm{B}) \subseteq \mathsf{SEP}(\mathrm{A}{:}\mathrm{B}) \subseteq \mathsf{PPT}(\mathrm{A}{:}\mathrm{B}). \tag{1.1}$$

1.4 Measurements

We have defined LOCC, SEP, and PPT in the form of operations, suitable for describing the transformations that can be performed by Alice and Bob. Usually when defining measurements, the whole system is measured, however to prove our results we need a more flexible definition of measurement, where only some parties are measured, as otherwise some of the statements will be false. The purpose of this section is to give a framework to work with partial measurements in LOCC, SEP, and PPT.

Consider again systems H and H′, and let $\{|i\rangle_{\mathrm{H}'}\}_{i\in\mathbb{Z}_{|\mathrm{H}'|}}$ be a choice of computational basis. A measurement on H is a channel $\mathcal{M} \in \mathsf{CPTP}(\mathrm{H}\rangle\mathrm{H}')$ such that

$$\mathcal{M}(\rho) := \sum_{i\in\mathbb{Z}_{|\mathrm{H}'|}} \mathrm{tr}(M_i\rho)\, |i\rangle\langle i|_{\mathrm{H}'}\,,$$

for some positive operators $M_i \in \mathcal{H}_+(\mathrm{H})$. To satisfy the trace-preserving condition these operators will satisfy $\sum_{i\in\mathbb{Z}_{|\mathrm{H}'|}} M_i = \mathbb{1}_{\mathrm{H}}$, and thus $\{\mathrm{Tr}\, M_i\rho\}_{i\in\mathbb{Z}|\mathrm{H}'|}$ forms a probability distribution. The i are known as the measurement outcomes, Tr $M_i\rho$ as the measurement probabilities, the M_i as the measurement operators, $\{M_i\}_{i\in\mathbb{Z}_{|\mathrm{H}'|}}$ as the positive operator-valued measure (POVM), and $|\mathrm{H}'|$ as the number of outcomes. A measurement is called projective if its measurement operators are projectors. A measurement on a basis of H, also known as a von Neumann measurement, is a projective measurement with rank-one measurement operators, namely they are the projectors onto an orthonormal basis $\{|v_i\rangle\}_{i\in\mathbb{Z}_{|\mathrm{H}|}}$ of H. Finally, the dephasing on the computational basis is the same channel as the von Neumann measurement $\mathcal{M} \in \mathsf{CPTP}(\mathrm{H}\rangle\mathrm{H})$ on the computational basis[4]. We denote by $\mathsf{M}(\mathrm{H}\rangle\underline{\mathrm{H}}')$ the set of measurements. For the set of all measurements on H

$$\mathsf{M}(\underline{\mathrm{H}}) := \bigcup_{\mathrm{H}'} \mathsf{M}(\mathrm{H}\rangle\underline{\mathrm{H}}').$$

where we use $\underline{\cdot}$ on the input, to denote that any output system has classical output.

Let now $\mathsf{L}(\mathrm{H}\rangle\mathrm{H}') \subseteq \mathsf{CPTP}(\mathrm{H}\rangle\mathrm{H}')$ be any subset of channels. We define $\mathsf{L}(\mathrm{H}\rangle\underline{\mathrm{H}}') \subseteq \mathsf{L}(\mathrm{H}\rangle\mathrm{H}')$ as those channels that ends with a measurement on H′. By definition, this is the intersection with the measurements on H:

$$\mathsf{L}(\mathrm{H}\rangle\underline{\mathrm{H}}') := \mathsf{L}(\mathrm{H}\rangle\mathrm{H}') \cap \mathsf{M}(\mathrm{H}\rangle\underline{\mathrm{H}}').$$

Notice that $\mathsf{L}(\mathrm{H}\rangle\underline{\mathrm{H}}')$ could be empty even if $\mathsf{L}(\mathrm{H}\rangle\mathrm{H}')$ is not, as is the case for the set of unitary conjugations. $\mathsf{L}(\mathrm{H}\rangle\underline{\mathrm{H}}')$ is a subset of the measurements achieved by composing H′ with any measurement, and in general the two sets are not the same:

$$\mathsf{L}(\underline{\mathrm{H}}'\langle\mathrm{H}) \subseteq \mathsf{M}(\underline{\mathrm{H}}'\langle\mathrm{H}') \circ \mathsf{L}(\mathrm{H}'\langle\mathrm{H}).$$

The inclusion is due to the fact that there always exist a non-disturbing measurement in $\mathsf{M}(\underline{\mathrm{H}}'\langle\mathrm{H}')$. However, while all the measurements in L are preserved, $\mathsf{M}(\underline{\mathrm{H}}'\langle\mathrm{H}') \circ \mathsf{L}(\mathrm{H}'\langle\mathrm{H})$ will in general also include new measurements and the inclusion will be strict. In the example of unitary conjugations, we do not get the empty set. If $\mathsf{L}(\mathrm{H}\rangle\mathrm{H}')$ is defined for all systems H′, then we have the set of all measurements in L given by

$$\mathsf{L}(\underline{\mathrm{H}}) := \bigcup_{\mathrm{H}'} \mathsf{L}(\mathrm{H}\rangle\underline{\mathrm{H}}').$$

A subset $\mathsf{L} \in \mathsf{L}(\underline{\mathrm{H}})$ is said to be informationally complete [MWW09], if for all operators $K \in \mathcal{L}(\mathrm{H})$ there exist $\Lambda \in \mathsf{L}$ such that $\Lambda(K) \neq 0$.

[4] We could make a definition of instrument and measurements where the classical registry is any basis, however this would let us call measurements also any dephasing channel on a basis, which we want to avoid.

1.4.1 Partial measurements

It is a straightforward observation to see that measurements are obtained by tracing out the quantum part of an instrument, thus we could have also defined the set of measurements as

$$\mathsf{M}(\mathrm{H}\rangle\underline{\mathrm{M}}) = \mathsf{IN}(\mathrm{H}\rangle\underline{\mathrm{M}}).$$

We can think of instruments as measuring a subsystem, and thus we can call them also partial measurements. Let $\mathrm{H}, \mathrm{H}', \mathrm{H}''$ be any systems and let $\mathsf{L}(\mathrm{H}\rangle\mathrm{H}'\mathrm{H}'') \subseteq \mathsf{CPTP}(\mathrm{H}\rangle\mathrm{H}'\mathrm{H}'')$, we thus define the subset L of partial measurements on H'', like in the case of measurements, as the intersection with the instruments that are classical on H''

$$\mathsf{L}\big(\mathrm{H}\rangle\mathrm{H}'\underline{\mathrm{H}}''\big) := \mathsf{L}\big(\mathrm{H}\rangle\mathrm{H}'\mathrm{H}''\big) \cap \mathsf{IN}\big(\mathrm{H}\rangle\mathrm{H}'\underline{\mathrm{H}}''\big),$$

where now we view H'' as being the classical output system of the instruments (the $\underline{\mathrm{M}}$ in $\mathsf{IN}(\mathrm{H}\rangle\mathrm{H}'\underline{\mathrm{M}})$). Like for measurement, $\mathsf{L}\big(\mathrm{H}\rangle\mathrm{H}'\underline{\mathrm{H}}''\big)$ is in general a non trivial subset of the partial measurements achieved by composing with any local measurement on H'' [2],

$$\mathsf{L}\big(\mathrm{H}'\underline{\mathrm{H}}''\langle\mathrm{H}\big) \subseteq \big[\mathrm{id}_{\mathrm{H}'} \otimes \mathsf{M}\big(\underline{\mathrm{H}}''\langle\mathrm{H}''\big)\big] \circ \mathsf{L}\big(\mathrm{H}'\mathrm{H}''\langle\mathrm{H}\big).$$

In the particular case where L is invariant under local channels, so we have

$$\mathsf{L}\big(\mathrm{H}'\mathrm{H}''\langle\mathrm{H}\big) \supseteq \big[\mathrm{id}_{\mathrm{H}'} \otimes \mathsf{CPTP}\big(\mathrm{H}''\langle\mathrm{H}''\big)\big] \circ \mathsf{L}\big(\mathrm{H}'\mathrm{H}''\langle\mathrm{H}\big),$$

then the opposite inclusion becomes trivial and we get the equality:

$$\mathsf{L}\big(\mathrm{H}'\underline{\mathrm{H}}''\langle\mathrm{H}\big) = \big[\mathrm{id}_{\mathrm{H}'} \otimes \mathsf{M}\big(\underline{\mathrm{H}}''\langle\mathrm{H}''\big)\big] \circ \mathsf{L}\big(\mathrm{H}'\mathrm{H}''\langle\mathrm{H}\big). \tag{1.2}$$

Finally, while $\mathsf{M}\big(\mathrm{H}\rangle\underline{\mathrm{H}}''\big)$ itself is not invariant under local channels, we still have:

$$\mathsf{M}\big(\underline{\mathrm{H}}''\langle\mathrm{H}\big) = \mathsf{M}\big(\underline{\mathrm{H}}''\langle\mathrm{H}''\big) \circ \mathsf{M}\big(\underline{\mathrm{H}}''\langle\mathrm{H}\big),$$

and even

$$\mathsf{M}\big(\underline{\mathrm{H}}''\langle\mathrm{H}\big) = \mathsf{M}\big(\underline{\mathrm{H}}''\langle\mathrm{H}''\big) \circ \mathsf{CPTP}\big(\mathrm{H}''\langle\mathrm{H}\big).$$

Notice that it makes no sense, just like for instruments, to define for example $\mathsf{L}\big(\mathrm{H}\underline{\mathrm{H}}''\big)$, namely to define partial measurements for input systems, unless the are some restriction between the input and the output systems. This becomes possible for bipartite channels due to the extra structure imposed by bipartite systems, which is the reason for introducing partial measurements.

1.4.2 Bipartite classes

If $\mathcal{K}$ defines a subset of states $\mathcal{K}(\mathrm{H}) \subseteq \mathcal{D}(\mathrm{H})$ for any system H, then we also call it a class states. Similarly, if L defines a subset of channels $\mathsf{L}(\mathrm{H}\rangle\mathrm{H}') \subseteq \mathsf{CPTP}(\mathrm{H}\rangle\mathrm{H}')$ for any systems H and H', then we also call it a class of channels. Separable and PPT states share some common properties as classes of states, and, similarly, LOCC, SEP, and PPT also share some common properties as classes of channels. The purpose of this section is to define these properties for channels and for states, so that we can use them in more generality. The definitions for bipartite systems also induce these definitions on single party systems by taking one of the parties to be trivial. However giving the definitions for single party only, does not imply the definition for bipartite systems, because it does not fix the freedom in grouping the parties. In principle, we should therefore reiterate these definitions for multipartite systems when we get there, instead we will leave this implicit.

We call L a class of bipartite channels, if it defines

$$\mathsf{L}(\mathrm{A_{in}{:}B_{in}}\rangle \mathrm{A_{out}{:}B_{out}}) \subseteq \mathsf{CPTP}(\mathrm{A_{in}{:}B_{in}}\rangle \mathrm{A_{out}{:}B_{out}})$$

for any systems $\mathrm{A_{in}}$, $\mathrm{B_{in}}$, $\mathrm{A_{out}}$, $\mathrm{B_{out}}$, and define the sets of all its channels on AB as

$$\mathsf{L}(\mathrm{A{:}B}) := \bigcup_{\mathrm{A_{out}{:}B_{out}}} \mathsf{L}(\mathrm{A{:}B}\rangle \mathrm{A_{out}{:}B_{out}}).$$

We then define the following properties, where it is implicit that they need to hold for all mentioned systems.

- L is closed under tensor products if

$$\mathsf{L}\big(\mathrm{A_1{:}B_1}\big\rangle \mathrm{A_1'{:}B_1'}\big) \otimes \mathsf{L}\big(\mathrm{A_2{:}B_2}\big\rangle \mathrm{A_2'{:}B_2'}\big) \subseteq \mathsf{L}\big(\mathrm{A_1A_2{:}B_1B_2}\big\rangle \mathrm{A_1'A_2'{:}B_1'B_2'}\big).$$

- L is close under composition if

$$\mathsf{L}(\mathrm{A_{out}{:}B_{out}}\langle \mathrm{A_{in}{:}B_{in}}) \supseteq \mathsf{L}(\mathrm{A_{out}{:}B_{out}}\langle \mathrm{A{:}B}) \circ \mathsf{L}(\mathrm{A{:}B}\langle \mathrm{A_{in}{:}B_{in}}).$$

- L is closed under local operations if

$$\mathsf{L}(\mathrm{A_{out}{:}B_{out}}\langle \mathrm{A_{in}{:}B_{in}}) \supseteq \mathsf{LO}(\mathrm{A_{out}{:}B_{out}}\langle \mathrm{A{:}B}) \circ \mathsf{L}(\mathrm{A{:}B}\langle \mathrm{A_{in}{:}B_{in}}).$$

  In particular this property implies that we can write those channels where a party is removed simply by tracing the party, namely:

$$\mathsf{L}(\mathbb{C}{:}\mathrm{B_{out}}\langle \mathrm{A_{in}{:}B_{in}}) = \mathrm{tr}_{\mathrm{A_{out}}} \circ \mathsf{L}(\mathrm{A_{out}{:}B_{out}}\langle \mathrm{A_{in}{:}B_{in}}). \tag{1.3}$$

- L is a class of operations if it contains local operations and is closed under tensor product and composition. In particular, measurement classes are not classes of operations.

- L contains classical communication from Alice to Bob if

$$\mathsf{L}(\mathrm{A_{in}{:}B_{in}}\rangle \mathrm{A_{out}}\underline{\mathrm{M}}{:}\mathrm{B_{out}}) \subseteq \mathsf{L}(\mathrm{A_{in}{:}B_{in}}\rangle \mathrm{A_{out}}{:}\underline{\mathrm{M}}\mathrm{B_{out}}), \tag{1.4}$$

  or from Bob to Alice if

$$\mathsf{L}(\mathrm{A_{in}{:}B_{in}}\rangle \mathrm{A_{out}}{:}\underline{\mathrm{M}}\mathrm{B_{out}}) \subseteq \mathsf{L}(\mathrm{A_{in}{:}B_{in}}\rangle \mathrm{A_{out}}\underline{\mathrm{M}}{:}\mathrm{B_{out}}).$$

- L is informationally complete if for all operators $K \in \mathcal{L}(\mathrm{AB})$ there exist $\Lambda \in \mathsf{L}(\mathrm{A{:}B})$ such that $\Lambda(K) \neq 0$.

Some comments about these properties are due. Local operations are informationally complete, closed under composition and tensor product, and trivially closed under themselves. It implies that classes of operations, their measurements and partial measurements are informationally complete. A class of channels can be informationally complete, without its measurements being too (again the class of unitary conjugations is an example, having an empty subset of measurements). However if L is informationally complete and closed under local operations, then its measurements will be informationally complete. If L is a class of operations containing classical communication from Alice to Bob, then it will contain $\mathsf{LOCC}_{\rightarrow}$, and if it contains communication both ways, then it will contain also LOCC. Therefore LOCC are the smallest classes satisfying these properties and they are informationally complete.

We now abstract the properties of classes of states. We call $\mathcal{K}$ a class of bipartite states, if it defines $\mathcal{K}(\mathrm{A{:}B}) \subseteq \mathcal{D}(\mathrm{AB})$ for all systems A and B. We then have the following natural properties.

- $\mathcal{K}$ is convex if $\mathcal{K}(\mathrm{A{:}B})$ is convex.
- $\mathcal{K}$ is closed under tensor products if

$$\mathcal{K}(\mathrm{A{:}B}) \otimes \mathcal{K}(\mathrm{A'{:}B'}) \subseteq \mathcal{K}(\mathrm{AA'{:}BB'}).$$

- $\mathcal{K}$ is closed under the trace if

$$\mathrm{tr}_{\mathrm{AB}}\, \mathcal{K}(\mathrm{AA'{:}BB'}) \subseteq \mathcal{K}(\mathrm{A'{:}B'}).$$

- $\mathcal{K}$ is closed under L if

$$\mathcal{K}(\mathrm{A_{out}{:}B_{out}}) \supseteq \mathsf{L}(\mathrm{A_{in}{:}B_{in}}\rangle \mathrm{A_{out}{:}B_{out}}) \circ \mathcal{K}(\mathrm{A_{in}{:}B_{in}}).$$

- $\mathcal{K}$ is closed under measurement operators in L, if for any $\sigma \in \mathcal{K}(\mathrm{A'A{:}B'B})$ and any measurement operator M of a measurement in $\mathsf{L}(\underline{\mathrm{A}}'{:}\underline{\mathrm{B}}')$, we have

$$\frac{\mathrm{tr}_{\mathrm{A'B'}}[(M \otimes \mathbb{1}_{\mathrm{AB}})\sigma]}{\mathrm{tr}[(M \otimes \mathbb{1}_{\mathrm{AB}})\sigma]} \in \mathcal{K}(\mathrm{A{:}B}),$$

  namely if $\mathcal{K}$ is closed under conditioning on a measurement outcome in L.

The class of pure states is an example of a class closed under tensor product but not under the trace, while the Werner states (presented in Section 2.5) are a class that is closed under the trace but not under tensor product. Notice that the class of PPT states is closed under PPT measurement operators and consequently under measurements operators for all subclasses of PPT. Similarly the class of separable states is closed under separable measurement operators, and consequently under LOCC measurement operators.

1.4.3 Bipartite measurements

For a bipartite class of channels L, the sets of measurements in L are simply $\mathsf{L}(\mathrm{A{:}B}\rangle\underline{\mathrm{A}}'{:}\underline{\mathrm{B}}')$, and the sets of partial measurement are $\mathsf{L}(\mathrm{A{:}B}\rangle\underline{\mathrm{A}}'{:}\mathrm{B}')$ and $\mathsf{L}(\mathrm{A{:}B}\rangle\mathrm{A}'{:}\underline{\mathrm{B}}')$, where we have by construction

$$\mathsf{L}(\mathrm{A{:}B}\rangle\underline{\mathrm{A}}'{:}\underline{\mathrm{B}}') \subseteq \mathsf{L}(\mathrm{A{:}B}\rangle\underline{\mathrm{A}}'{:}\mathrm{B}'), \mathsf{L}(\mathrm{A{:}B}\rangle\mathrm{A}'{:}\underline{\mathrm{B}}').$$

When L is closed under local operations, then we can use Equation (1.2) and think of these measurement sets as the ones achieved by composing the channels with local measurements. Finally, the restrictions imposed by bipartite channels allows us to define measurements and partial measurements with unspecified output (something that otherwise we said does not make sense). We define

$$\begin{aligned}
\mathsf{L}(\underline{\mathrm{A}}{:}\underline{\mathrm{B}}) &:= \bigcup_{\mathrm{A_{out}},\mathrm{B_{out}}} \mathsf{L}(\mathrm{A{:}B}\rangle\underline{\mathrm{A}}_{\mathrm{out}}{:}\underline{\mathrm{B}}_{\mathrm{out}})\\
\mathsf{L}(\underline{\mathrm{A}}{:}\mathrm{B}) &:= \bigcup_{\mathrm{A_{out}},\mathrm{B_{out}}} \mathsf{L}(\mathrm{A{:}B}\rangle\underline{\mathrm{A}}_{\mathrm{out}}{:}\mathrm{B}_{\mathrm{out}})\\
\mathsf{L}(\mathrm{A}{:}\underline{\mathrm{B}}) &:= \bigcup_{\mathrm{A_{out}},\mathrm{B_{out}}} \mathsf{L}(\mathrm{A{:}B}\rangle\mathrm{A}_{\mathrm{out}}{:}\underline{\mathrm{B}}_{\mathrm{out}})
\end{aligned}$$

for which we again have $\mathsf{L}(\underline{\mathrm{A}}{:}\underline{\mathrm{B}}) \subset \mathsf{L}(\underline{\mathrm{A}}{:}\mathrm{B}), \mathsf{L}(\mathrm{A}{:}\underline{\mathrm{B}})$. It still makes no sense though to have some systems measured and some not within a single party, for example something like $\mathsf{L}(\mathrm{A{:}B}\underline{\mathrm{M}})$, especially for classes closed under local operations. The above definitions preserve inclusion and thus they preserve Equation (1.1) under measurements.

Simplifications

As LO, LOCC, LOCC$_{\rightarrow}$, SEP and PPT are all closed under composition and tensor products, to work with their measurements and their partial measurements we can use Equation (1.2), which says (omitting $\mathrm{id}_{\mathrm{B}_{\mathrm{out}}}$)

$$\mathsf{L}(\underline{\mathrm{A}}_{\mathrm{out}}{:}\mathrm{B}_{\mathrm{out}}\rangle \mathrm{A}_{\mathrm{in}}{:}\mathrm{B}_{\mathrm{in}}) = \mathsf{M}(\underline{\mathrm{A}}_{\mathrm{out}}\langle \mathrm{A}_{\mathrm{out}}) \circ \mathsf{L}(\mathrm{A}_{\mathrm{out}}{:}\mathrm{B}_{\mathrm{out}}\rangle \mathrm{A}_{\mathrm{in}}{:}\mathrm{B}_{\mathrm{in}}).$$

and similarly for partial measurements on Bob, and for full measurements. Because of this local measurements have the simple characterization

$$\mathsf{LO}(\mathrm{A}{:}\mathrm{B}\rangle \underline{\mathrm{A}}_{\mathrm{out}}{:}\underline{\mathrm{B}}_{\mathrm{out}}) = \mathsf{M}(\mathrm{A}\rangle \underline{\mathrm{A}}_{\mathrm{out}}) \otimes \mathsf{M}(\mathrm{B}\rangle \underline{\mathrm{B}}_{\mathrm{out}}).$$

The set of all local measurements on A:B is then:

$$\mathsf{LO}(\underline{\mathrm{A}}{:}\underline{\mathrm{B}}) = \bigcup_{\mathrm{A}_{\mathrm{out}},\mathrm{B}_{\mathrm{out}}} \mathsf{LO}(\mathrm{A}{:}\mathrm{B}\rangle \underline{\mathrm{A}}_{\mathrm{out}}{:}\underline{\mathrm{B}}_{\mathrm{out}}) = \mathsf{M}(\underline{\mathrm{A}}) \otimes \mathsf{M}(\underline{\mathrm{B}})$$

while the local partial measurements are

$$\begin{aligned} \mathsf{LO}(\underline{\mathrm{A}}{:}\mathrm{B}) &= \mathsf{M}(\underline{\mathrm{A}}) \otimes \mathsf{CPTP}(\mathrm{B}) \\ \mathsf{LO}(\mathrm{A}{:}\underline{\mathrm{B}}) &= \mathsf{CPTP}(\mathrm{A}) \otimes \mathsf{M}(\underline{\mathrm{B}}). \end{aligned}$$

We get a simplification also for the one-way LOCC channels, and without loss of generality let us consider only $\mathsf{LOCC}_{\rightarrow}(\underline{\mathrm{A}}{:}\mathrm{B})$. If Alice is measured then we can relax the channel on Bob to global channels, which is formalized by the following equation

$$\begin{aligned} \mathsf{LOCC}_{\rightarrow}(\underline{\mathrm{A}}{:}\mathrm{B}) &= \bigcup_{\mathrm{B}_{\mathrm{out}}} \mathsf{LOCC}_{\rightarrow}(\mathbb{C}{:}\mathrm{B}_{\mathrm{out}}\langle \mathrm{A}{:}\mathrm{B}) \\ &= \bigcup_{\mathrm{B}_{\mathrm{out}},\mathrm{M}} \mathsf{CPTP}(\mathrm{B}_{\mathrm{out}}\langle \mathrm{MB}) \circ [\mathsf{M}(\underline{\mathrm{M}}\langle \mathrm{A}) \otimes \mathrm{id}_{\mathrm{B}}], \end{aligned} \tag{1.5}$$

The second equality follows from the first by definition, while the first one is readily verified using the classical communication from Alice to Bob in the form of Equation (1.4)

$$\begin{aligned} \mathsf{LOCC}_{\rightarrow}(\underline{\mathrm{A}}{:}\mathrm{B}) &= \bigcup_{\mathrm{A}_{\mathrm{out}},\mathrm{B}_{\mathrm{out}}} \mathsf{LOCC}_{\rightarrow}(\underline{\mathrm{A}}_{\mathrm{out}}{:}\mathrm{B}_{\mathrm{out}}\langle \mathrm{A}{:}\mathrm{B}) \\ &\subseteq \bigcup_{\mathrm{A}_{\mathrm{out}},\mathrm{B}_{\mathrm{out}}} \mathsf{LOCC}_{\rightarrow}(\mathbb{C}{:}\underline{\mathrm{A}}_{\mathrm{out}}\mathrm{B}_{\mathrm{out}}\langle \mathrm{A}{:}\mathrm{B}) \\ &\subseteq \bigcup_{\mathrm{B}_{\mathrm{out}}} \mathsf{LOCC}_{\rightarrow}(\mathbb{C}{:}\mathrm{B}_{\mathrm{out}}\langle \mathrm{A}{:}\mathrm{B}) \\ &\subseteq \mathsf{LOCC}_{\rightarrow}(\underline{\mathrm{A}}{:}\mathrm{B}). \end{aligned}$$

Namely, we can think of $\mathsf{LOCC}_{\rightarrow}(\underline{\mathrm{A}}{:}\mathrm{B})$ both as Alice sending all the measurement outcomes to Bob, or as Alice being completely removed. This implies that, in completing to full measurements, we can relax the measurement on Bob to global measurements, namely

$$\mathsf{LOCC}_{\rightarrow}(\underline{\mathrm{A}}{:}\underline{\mathrm{B}}) = \bigcup_{\mathrm{B}_{\mathrm{out}},\mathrm{M}} \mathsf{M}(\underline{\mathrm{B}}_{\mathrm{out}}\langle \mathrm{MB}) \circ [\mathsf{M}(\underline{\mathrm{M}}\langle \mathrm{A}) \otimes \mathrm{id}_{\mathrm{B}}].$$

Classical communication

Next, inspired by the previous comments about LOCC$_{\rightarrow}$ measurements, we will make two similar remarks to simplify the use of general partial measurements. These appeared stated imprecisely in [2][5]. The first one is fairly simple, if L is closed under local operations and contains classical communication from Alice to Bob, then indeed a measured Alice can always be removed by sending the measurement outcomes to Bob. Namely, let Alice be measured without loss of generality, then we have

$$\mathsf{L}(\underline{\mathrm{A}}{:}\mathrm{B}) = \bigcup_{\mathrm{B}_{\mathrm{out}}} \mathsf{L}(\mathrm{A}{:}\mathrm{B}\rangle\mathbb{C}{:}\mathrm{B}_{\mathrm{out}}). \tag{1.6}$$

as it is shown by

$$\begin{aligned}\mathsf{L}(\underline{\mathrm{A}}{:}\mathrm{B}) &= \bigcup_{\mathrm{A}_{\mathrm{out}},\mathrm{B}_{\mathrm{out}}} \mathsf{L}(\mathrm{A}{:}\mathrm{B}\rangle\underline{\mathrm{A}}_{\mathrm{out}}{:}\mathrm{B}_{\mathrm{out}}) \subseteq \bigcup_{\mathrm{B}_{\mathrm{out}}} \mathsf{L}(\mathrm{A}{:}\mathrm{B}\rangle\mathbb{C}{:}\underline{\mathrm{A}}_{\mathrm{out}}\mathrm{B}_{\mathrm{out}}) \\ &\subseteq \bigcup_{\mathrm{B}_{\mathrm{out}}} \mathsf{L}(\mathrm{A}{:}\mathrm{B}\rangle\mathbb{C}{:}\mathrm{B}_{\mathrm{out}}) \subseteq \bigcup_{\mathrm{A}_{\mathrm{out}},\mathrm{B}_{\mathrm{out}}} \mathsf{L}(\mathrm{A}{:}\mathrm{B}\rangle\underline{\mathrm{A}}_{\mathrm{out}}{:}\mathrm{B}_{\mathrm{out}}) = \mathsf{L}(\underline{\mathrm{A}}{:}\mathrm{B}).\end{aligned}$$

Because of Equation (1.3) we can think of this as a form of "$\mathsf{L}(\underline{\mathrm{A}}{:}\mathrm{B}) = \mathrm{tr}_{\mathrm{A}}\, \mathsf{L}(\mathrm{A}{:}\mathrm{B})$". This can be straightforwardly generalized to any multipartite setting, as long as the measured party has classical communication to at least one other party.

The second remark is of complementary flavour, and says that when Alice is already measured and we want to complete the measurement on Bob, we can relax to global measurements.

Lemma 1. *Let L be a class of bipartite channels that contains classical communication from* A *to* B *and that is closed under local operations. Then*

$$\mathsf{L}(\underline{\mathrm{A}}{:}\underline{\mathrm{B}}) = \bigcup_{\mathrm{M},\mathrm{A}_{\mathrm{out}},\mathrm{B}_{\mathrm{out}}} \mathsf{M}(\underline{\mathrm{M}}\langle\mathrm{A}_{\mathrm{out}}\mathrm{B}_{\mathrm{out}}) \circ \mathsf{L}(\underline{\mathrm{A}}_{\mathrm{out}}{:}\mathrm{B}_{\mathrm{out}}\langle\mathrm{A}{:}\mathrm{B}).$$

Lemma 1 should be thought as a kind of "$\mathsf{L}(\underline{\mathrm{A}}{:}\underline{\mathrm{B}}) = \mathsf{M}(\underline{\mathrm{A}}\underline{\mathrm{B}}) \circ \mathsf{L}(\underline{\mathrm{A}}{:}\mathrm{B})$". This formulation will be enough in the next chapter to lift measurements in the trace norms. Again this should generalize to the multipartite setting, the condition being that there is only one party that is not yet measured and, directly or indirectly, it can receive classical communication from any other party.

Proof. By the assumption that L is closed under local operations we can use Equation (1.2), which says that we can write the partial measurements as composition with the all the measurements on the measured system, and by simply relaxing to global measurements we get

$$\begin{aligned}\mathsf{L}(\underline{\mathrm{A}}{:}\underline{\mathrm{B}}) &= \bigcup_{\mathrm{A}_{\mathrm{out}},\mathrm{B}_{\mathrm{out}}} \mathsf{L}(\underline{\mathrm{A}}_{\mathrm{out}}{:}\underline{\mathrm{B}}_{\mathrm{out}}\langle\mathrm{A}{:}\mathrm{B}) \\ &= \bigcup_{\mathrm{A}_{\mathrm{out}},\mathrm{B}_{\mathrm{out}}} (\mathsf{M}(\underline{\mathrm{A}}_{\mathrm{out}}\langle\mathrm{A}_{\mathrm{out}}) \otimes \mathsf{M}(\underline{\mathrm{B}}_{\mathrm{out}}\langle\mathrm{B}_{\mathrm{out}})) \circ \mathsf{L}(\mathrm{A}_{\mathrm{out}}{:}\mathrm{B}_{\mathrm{out}}\langle\mathrm{A}{:}\mathrm{B}) \\ &\subseteq \bigcup_{\mathrm{M},\mathrm{A}_{\mathrm{out}},\mathrm{B}_{\mathrm{out}}} \mathsf{M}(\underline{\mathrm{M}}\langle\mathrm{A}_{\mathrm{out}}\mathrm{B}_{\mathrm{out}}) \circ \mathsf{L}(\underline{\mathrm{A}}_{\mathrm{out}}{:}\mathrm{B}_{\mathrm{out}}\langle\mathrm{A}{:}\mathrm{B}).\end{aligned}$$

[5] In [2] we would allow to compose two classes of channels L and L' by simply composing any channels with matching input/output. The flaw in this definition is that, while we are allowed to split the Hilbert space in tensor product spaces to define the channel, two Hilbert spaces of the same dimension, like $\mathbb{C}^d \otimes \mathbb{C}^n$ and $\mathbb{C}^n \otimes \mathbb{C}^d$, are actually the same Hilbert space independently of how they where built. Therefore composing (e.g.) separable channels, gives separable channels only if we keep track of the input/output systems, and make sure that the bipartitions match in the composition.

where we also made use of $\mathsf{M}(\underline{\mathrm{A}}_{\mathrm{out}}\langle \mathrm{A}_{\mathrm{out}}) = \mathsf{M}(\underline{\mathrm{A}}_{\mathrm{out}}\langle \mathrm{A}_{\mathrm{out}}) \circ \mathsf{M}(\underline{\mathrm{A}}_{\mathrm{out}}\langle \mathrm{A}_{\mathrm{out}})$. By assumption on the presence of classical communication we then have

$$\begin{aligned} \mathsf{L}(\underline{\mathrm{A}}{:}\underline{\mathrm{B}}) &\subseteq \bigcup_{\mathrm{M},\mathrm{A}_{\mathrm{out}},\mathrm{B}_{\mathrm{out}}} \mathsf{M}(\underline{\mathrm{M}}\langle \mathrm{A}_{\mathrm{out}}\mathrm{B}_{\mathrm{out}}) \circ \mathsf{L}(\underline{\mathrm{A}}_{\mathrm{out}}{:}\mathrm{B}_{\mathrm{out}}\langle \mathrm{A}{:}\mathrm{B}) \\ &\subseteq \bigcup_{\mathrm{M},\mathrm{A}_{\mathrm{out}},\mathrm{B}_{\mathrm{out}}} \mathsf{M}(\underline{\mathrm{M}}\langle \mathrm{A}_{\mathrm{out}}\mathrm{B}_{\mathrm{out}}) \circ \mathsf{L}(\mathbb{C}{:}\underline{\mathrm{A}}_{\mathrm{out}}\mathrm{B}_{\mathrm{out}}\langle \mathrm{A}{:}\mathrm{B}) \\ &\subseteq \bigcup_{\mathrm{M}} \mathsf{L}(\mathbb{C}{:}\underline{\mathrm{M}}\langle \mathrm{A}{:}\mathrm{B}) \\ &\subseteq \mathsf{L}(\underline{\mathrm{A}}{:}\underline{\mathrm{B}}). \end{aligned}$$

□

1.5 Qudit systems

So far we have reviewed the formalism and the basics concepts of quantum information, while at the same time only talking generally about "quantum systems". But who am I kidding? There is no way I can slip through this thesis without the mentioning "quantum computer". We have hinted at the fact that finite dimensional quantum systems are the quantum analogue of digital systems, while carefully dodging mentioning that classical computers are the emblematic example of digital systems. The finite systems handled by classical computers are called bits, and while there are many examples of everyday finite systems (coins, alphabets, etc) that are not called bits, it is natural to call them bits when the topic involves some form of computation. In this section we will review some more specific aspects of finite dimensional quantum systems and call them quantum bits, or quantum dits, to highlight the more computational nature of these properties.

1.5.1 One qudit - Bit flips and phase flips

A single finite dimensional quantum system H is also known as a qudit, or qubit if the dimension is two[6]. Let $d = |\mathrm{H}|$ be its dimension. We define the d-th root of unity $\omega = e^{i\frac{2\pi}{d}}$ and, given the choice of computational basis $\{|i\rangle_{\mathrm{H}}\}_{i\in\mathbb{Z}_d}$, we define the following unitary operators

$$X_{\mathrm{H}} := \sum_{j\in\mathbb{Z}_d} |j+1\rangle\langle j|_{\mathrm{H}} \qquad\qquad Z_{\mathrm{H}} := \sum_{j\in\mathbb{Z}_d} \omega^j\, |j\rangle\langle j|_{\mathrm{H}}\,, \tag{1.7}$$

which satisfy $X_{\mathrm{H}}^d = Z_{\mathrm{H}}^d = \mathbb{1}_{\mathrm{H}}$. We define also their corresponding channels:

$$\mathcal{X}_{\mathrm{H}}^j(\rho) := X_{\mathrm{H}}^j \rho X_{\mathrm{H}}^{-j} \qquad\qquad \mathcal{Z}_{\mathrm{H}}^j(\rho) := Z_{\mathrm{H}}^j \rho Z_{\mathrm{H}}^{-j}$$

As usual we allow the system subscript to be implicit. For $d = 2$ we will allow the notation $\mathcal{Z}^+ = \mathcal{Z}^0$ and $\mathcal{Z}^- = \mathcal{Z}^1$. The operators obey the commutation relation $ZX = \omega XZ$. Since commuting them picks up only a phase which cancels out under conjugation, we have that $\mathcal{X}$ and $\mathcal{Z}$ commute. For $d = 2$ they reduce to the Pauli X and the Pauli Z

$$\mathrm{X} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \qquad\qquad Z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},$$

also known as the bit flip and the phase flip, we will thus keep calling them bit flip and phase flip for arbitrary d.

[6] "qubits are neato" cit. mAlexander

We also define the discrete Fourier transform with unitary normalization:

$$F_{\mathrm{H}} := \frac{1}{\sqrt{d}} \sum_{ij \in \mathbb{Z}_d} \omega^{-ij} |i\rangle\langle j|_{\mathrm{H}} . \tag{1.8}$$

or just F when then system is implicit. The Fourier transform is a change of basis from the computational basis to the eigenbasis of X, also known as the conjugate basis. Furthermore, X and Z have the same spectrum and F is the transform between them, as displayed by $XF|i\rangle = \omega^i F|i\rangle = FZ|i\rangle$ for $i \in \mathbb{Z}_d$, namely we have the transformation rules

$$XF = FZ \qquad\qquad XF^{-1} = F^{-1}Z^{-1}.$$

If we denote the channel of the Fourier transform with

$$\mathcal{F}_{\mathrm{H}}(\rho) := F_{\mathrm{H}} \rho F_{\mathrm{H}}^{\dagger}$$

then we can rewrite the relations as $\mathcal{X} \circ \mathcal{F} = \mathcal{F} \circ \mathcal{Z}$ and $\mathcal{X} \circ \mathcal{F}^{-1} = \mathcal{F}^{-1} \circ \mathcal{Z}^{-1}$.

The bit flips and phase flips generate the unitary subgroup known as the Weyl or Weyl-Heisenberg group [Wey27]. Up to phases, all the elements of the group are given by $X^i Z^j$ for $i, j \in \mathbb{Z}_d$, known as the Weyl operators. Therefore $\mathcal{X}^i \mathcal{Z}^j \equiv \mathcal{X}^i \circ \mathcal{Z}^j$ are all the channels generated by the Weyl group on a single qudit. The fourier transform then transforms these channels nicely as

$$\mathcal{X}^i \mathcal{Z}^j \circ \mathcal{F} = \mathcal{F} \circ \mathcal{X}^j \mathcal{Z}^i.$$

We will denote the Weyl operators also as Σ^{ij}, and they satisfy the following basic properties:

$$\begin{aligned} \Sigma^{ij} &:= X^i Z^j & \qquad \Sigma^{ijT} &= \omega^{-ij} \Sigma^{-i,j} \\ \overline{\Sigma}^{ij} &= \Sigma^{i,-j} & \qquad \Sigma^{ij\dagger} &= \omega^{ij} \Sigma^{-i,-j} \end{aligned}$$

The Weyl operators satisfy the following commutation relation

$$\Sigma^{ij} \Sigma^{kl} = \omega^{jk} \Sigma^{i+k,j+l} = \omega^{jk-il} \Sigma^{kl} \Sigma^{ij}.$$

and all except the identity are traceless, namely $\mathrm{tr}(\Sigma^{ij\dagger}) = \delta_{i0}\delta_{j0}$, where δ is the Kronecker delta. More precisely, the Weyl operators form an orthogonal basis for $\mathcal{H}(\mathrm{H})$, namely

$$\mathrm{tr}\left(\Sigma^{ij\dagger} \Sigma^{kl}\right) = d\delta_{ik}\delta_{jl},$$

so scaling them by $1/\sqrt{d}$ makes them into an orthonormal basis.

Finally, we can rewrite the measurement on the computational basis as a twirl on the phase flips, namely we have:

$$\sum_{i \in \mathbb{Z}_d} |i\rangle\langle i| \rho |i\rangle\langle i| = \frac{1}{d} \sum_{j \in \mathbb{Z}_d} \mathcal{Z}^i(\rho)$$

and similarly the measurement on the conjugate basis can be rewritten as a twirling over bit flips. If we twirl over bit flips and phase flips, then the result is always the maximally mixed state.

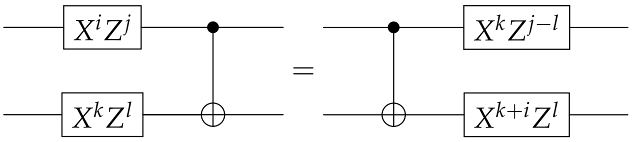


Figure 1.1: Effect of commuting CNOT with arbitrary bit/phase flips.

1.5.2 Two qudits - Bell states

Consider the system HH, or equivalently a system HH′ with $|\mathrm{H}'| = |\mathrm{H}|$. We call this a two qudit system. Let again $d = |\mathrm{H}|$. Having fixed a choice of computational product basis on the two qudit system $\{|i\rangle_{\mathrm{H}} \otimes |j\rangle_{\mathrm{H}'}\}_{ij\in\mathbb{Z}_d}$ we define the maximally entangled state[7]

$$\Phi_{\mathrm{HH}'} := \frac{1}{d}\sum_{ij\in\mathbb{Z}_d} |ii\rangle\langle jj|_{\mathrm{HH}'} = |\Phi\rangle\langle\Phi|_{\mathrm{HH}'}$$
$$|\Phi\rangle_{\mathrm{HH}'} := \frac{1}{\sqrt{d}}\sum_{i\in\mathbb{Z}_d} |ii\rangle_{\mathrm{HH}'}\,.$$

We will also use $\Phi^{\log d} \equiv \Phi_{\mathbb{C}^d\otimes\mathbb{C}^d}$ to denote the maximally entangled state on the anonymous qudits $\mathbb{C}^d \otimes \mathbb{C}^d$. For any operator K we have the following two useful properties, the second one known as the mirror lemma, that we can use to move operators and channels around

$$\mathrm{tr}((K\otimes\mathbb{1})\Phi) = \mathrm{tr}(K), \qquad\qquad (K\otimes\mathbb{1})\,|\Phi\rangle = (\mathbb{1}\otimes K^{\mathsf{T}})\,|\Phi\rangle\,.$$

We also have $d\Phi = S^{\Gamma}$ (where S is the swap operator defined in the preliminaries).

Acting with bit flips and phase flips (the Weyl operators) on Φ we can generate a basis for HH. We define for $i, j \in \mathbb{Z}_d$:[8]

$$\phi_{ij,\mathrm{HH}'} := \mathrm{id}_{\mathrm{H}} \otimes \mathcal{X}^i_{\mathrm{H}'}\mathcal{Z}^j_{\mathrm{H}'}(\Phi_{\mathrm{HH}'}) = |\phi_{ij}\rangle\langle\phi_{ij}|_{\mathrm{HH}'}\,. \tag{1.9}$$
$$|\phi_{ij}\rangle_{\mathrm{HH}'} := (\mathbb{1}_{\mathrm{H}} \otimes X^i_{\mathrm{H}'}Z^j_{\mathrm{H}'})\,|\Phi\rangle_{\mathrm{HH}'}\,.$$

In $d = 2$ these are known as the Bell states, for which we have the common notation

$$\Phi^+ \equiv \phi_{00} \qquad\qquad \Phi^- \equiv \phi_{01}$$
$$\Psi^+ \equiv \phi_{10} \qquad\qquad \Psi^- \equiv \phi_{11},$$

We will thus continue to call ϕ_{ij} Bell states in arbitrary dimension. Notice that we can write $\Phi^{\pm} = \mathcal{Z}^{\pm}(\Phi)$. A nice property that the Bell states satisfy, and that we will use further on, is that the Fourier transform can swap the indexes of the Bell states

$$[\overline{\mathcal{F}}\otimes\mathcal{F}](\phi_{ij}) = [\mathrm{id}\otimes(\mathcal{F}\circ\mathcal{X}^i\mathcal{Z}^j\circ\mathcal{F}^{\dagger})](\Phi) = [\mathrm{id}\otimes\mathcal{X}^j\mathcal{Z}^i](\Phi) = \phi_{ji}, \tag{1.10}$$

using the mirror lemma and the relations from the previous section. It is worth to remark, that in $d = 2$ there are additional product unitaries that make it possible to achieve any permutations of the Bell states [Ben+96c], however to the best of my knowledge, we do not know if it is possible in arbitrary dimension.

[7] where the reason for "maximally" is left for later.

[8] The decision of acting on the first or second system is in principle arbitrary, but it does generate a permuted basis.

On two qudits, the controlled-not is the controlled unitary defined as

$$\begin{aligned} CNOT_{\mathrm{HH}'} &\coloneqq \sum_{ij\in\mathbb{Z}_d} |i,i+j\rangle\langle ij|_{\mathrm{HH}'} = \sum_{i\in\mathbb{Z}_d} |i\rangle\langle i|_{\mathrm{H}} \otimes X^i_{\mathrm{H}'} \\ \mathcal{CNOT}_{\mathrm{HH}'}(\rho) &\coloneqq CNOT_{\mathrm{HH}'}\rho CNOT^\dagger_{\mathrm{HH}'}, \end{aligned}$$

namely the quantum analogue of addition in $\mathbb{Z}_d$. The $CNOT$ manifestly commutes with $Z\otimes\mathbb{1}$ and $\mathbb{1}\otimes X$, and it can be checked easily that $CNOT(X\otimes\mathbb{1}) = (X\otimes X)CNOT$ and $CNOT(\mathbb{1}\otimes Z) = (Z^{-1}\otimes Z)CNOT$, which put together give the commutation relation for the Weyl operators (see Figure 1.1 for the corresponding quantum circuit):

$$CNOT(X^iZ^j \otimes X^kZ^l) = (X^iZ^{j-l}\otimes X^{i+k}Z^l)CNOT. \tag{1.11}$$

The $CNOT$ can thus be used to transform between the Bell basis and the computational basis as follows:[9]

$$\begin{aligned} \phi_{ij} &= (\mathrm{id}\otimes\mathcal{X}^i\mathcal{Z}^j)\circ\mathcal{CNOT}\circ(\mathcal{F}^{-1}\otimes\mathrm{id})(|00\rangle\langle 00|) \\ &= \mathcal{CNOT}\circ(\mathcal{Z}^j\otimes\mathcal{X}^i\mathcal{Z}^j)\circ(\mathcal{F}^{-1}\otimes\mathrm{id})(|00\rangle\langle 00|) \\ &= \mathcal{CNOT}\circ(\mathcal{F}^{-1}\otimes\mathrm{id})\circ(\mathcal{X}^j\otimes\mathcal{X}^i\mathcal{Z}^j)(|00\rangle\langle 00|) \\ &= \mathcal{CNOT}\circ(\mathcal{F}^{-1}\otimes\mathrm{id})(|ji\rangle\langle ji|). \end{aligned}$$

We define $\hat{\Phi}_{\mathrm{HH}'}$, which we call the maximally correlated state, to be the state after measuring $\Phi_{\mathrm{HH}'}$ in the computational basis. If $\mathcal{M}$ is the measurement on the computational basis of HH′, and $\mathcal{M}'$ the measurement on the computational basis of H or H′, then:

$$\hat{\Phi} = \mathcal{M}(\Phi) = \mathrm{id}\otimes\mathcal{M}'(\Phi) = \mathcal{M}'\otimes\mathrm{id}(\Phi) = \frac{1}{d}\sum_{i\in\mathbb{Z}_d}|ii\rangle\langle ii|\,.$$

Furthermore, we define $\hat{\phi}_{ij,\mathrm{HH}'}$ to be the outcome of measuring $\phi_{ij,\mathrm{HH}'}$ on the computational basis. For the Bell states with no bit flips, the outcome is also $\hat{\Phi}$, namely we have for all $j\in\mathbb{Z}_d$

$$\begin{aligned} \hat{\phi}_{0j} &= \mathcal{M}(\phi_{0j}) = \mathrm{id}\otimes\mathcal{M}'(\phi_{0j}) = \mathcal{M}'\otimes\mathrm{id}(\phi_{0j}) \\ &= \hat{\Phi} = \frac{1}{d}\sum_{i\in\mathbb{Z}_d}|ii\rangle\langle ii| = \frac{1}{d}\sum_{k\in\mathbb{Z}_d}\phi_{0k}. \end{aligned}$$

The maximally correlated state is the uniform mixture on what we call the maximally correlated subspace; we denote the projector onto this subspace by

$$\mathbb{1}_{\hat{\Phi},\mathrm{HH}'} \coloneqq \sum_{i\in\mathbb{Z}_d}|ii\rangle\langle ii|_{\mathrm{HH}'} = \sum_{j\in\mathbb{Z}_d}\phi_{0j,\mathrm{HH}'}.$$

Finally, the twirl on $\Sigma^{ab}\otimes\overline{\Sigma}^{ab}$, which we can call the Weyl or Pauli twirl, can be used to implement the dephasing channel on the Bell basis. Namely:

$$\frac{1}{d^2}\sum_{ab\in\mathbb{Z}_d}\left(\Sigma^{ab}\otimes\overline{\Sigma}^{ab}\right)\rho\left(\Sigma^{ab}\otimes\overline{\Sigma}^{ab}\right)^\dagger = \sum_{kl\in\mathbb{Z}_d}\phi_{kl}\rho\phi_{kl}.$$

Proof. The statement can be checked by first writing:

$$X^a\otimes X^a = \sum_{kl\in\mathbb{Z}_d}\omega^{-al}\phi_{kl} \qquad\qquad Z^b\otimes Z^{-b} = \sum_{kl\in\mathbb{Z}_d}\omega^{bk}\phi_{kl},$$

[9] With X, Z, F and $CNOT$ defined as we did, acting on the first system would map $|\phi_{ij}\rangle$ to $|j,-i\rangle$ instead.

and using them to verify that

$$\Sigma^{ab} \otimes \overline{\Sigma}^{ab} = \sum_{kl \in \mathbb{Z}_d} \omega^{-al+bk} \phi_{kl} \qquad\qquad \phi_{kl} = \frac{1}{d^2} \sum_{ab \in \mathbb{Z}_d} \omega^{-bk+al} \Sigma^{ab} \otimes \overline{\Sigma}^{ab}.$$

Then:

$$\frac{1}{d^2} \sum_{ab \in \mathbb{Z}_d} (\Sigma^{ab} \otimes \overline{\Sigma}^{ab}) \rho (\Sigma^{ab} \otimes \overline{\Sigma}^{ab})^\dagger = \frac{1}{d^2} \sum_{abkl \in \mathbb{Z}_d} \omega^{bk-al} \phi_{kl} \rho \phi_{\tilde{k}\tilde{l}} \omega^{-b\tilde{k}+a\tilde{l}} = \sum_{kl \in \mathbb{Z}_d} \phi_{kl} \rho \phi_{kl}. \square$$

1.5.3 Three qudits - Quantum teleportation

One of the reasons for calling Φ "maximally" entangled is that it can be used to implement the identity channel, through the protocol of quantum teleportation [Ben+93]. In this protocol Alice holds a qudit system A, while sharing with Bob also the maximally entangled state $\Phi_{\mathrm{A_{in}B_{out}}}$, meaning we have the bipartite system $\mathrm{AA_{in}}$:$\mathrm{B_{out}}$. In quantum teleportation, it is possible, using only LOCC, for Alice to send the quantum state of A to Bob. The steps of the LOCC protocol are

1. Alice measures the Bell basis ϕ_{ij} on $\mathrm{AA_{in}}$;
2. Alice sends the measurement outcome ij (two classical dits) to Bob;
3. Bob applies $\Sigma^{ab\intercal}$ to $\mathrm{B_{out}}$;
4. The measurement outcome can be discarded.

The resulting channel Λ has separable Kraus operators

$$\langle \phi_{ab} | \otimes \Sigma^{ab\intercal} = \left[\langle \Phi | \left(\mathbb{1} \otimes \Sigma^{ab\dagger} \right) \right] \otimes \Sigma^{ab\intercal}.$$

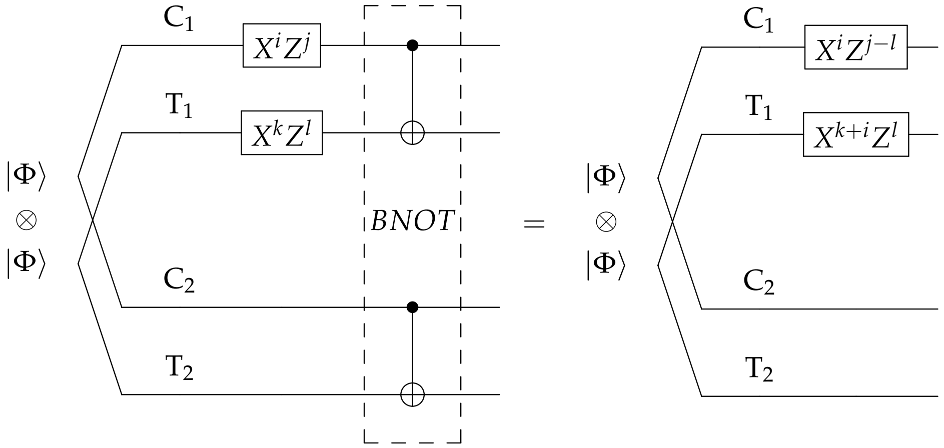
Verifying that Λ is the identity channel is a simple application of the mirror lemma. Let $K \in \mathcal{L}(\mathrm{H})$, then

$$\begin{aligned}
\Lambda(K \otimes \Phi_{\mathrm{A_{in}B_{out}}}) &:= \sum_{a,b \in \mathbb{Z}_d} \left(\langle \phi_{ab} |_{\mathrm{AA_{in}}} \otimes \Sigma^{ab\intercal}_{\mathrm{B_{out}}} \right) (K \otimes \Phi_{\mathrm{A_{in}B_{out}}}) \left(| \phi_{ab} \rangle_{\mathrm{AA_{in}}} \otimes \overline{\Sigma}^{ab}_{\mathrm{B_{out}}} \right) \\
&= \sum_{a,b \in \mathbb{Z}_d} \left(\langle \Phi |_{\mathrm{AA_{in}}} \otimes \mathbb{1}_{\mathrm{B_{out}}} \right) (K \otimes \Phi_{\mathrm{A_{in}B_{out}}}) \left(| \Phi \rangle_{\mathrm{AA_{in}}} \otimes \mathbb{1}_{\mathrm{B_{out}}} \right) \\
&= \sum_{a,b,i,j \in \mathbb{Z}_d} \left(\langle \Phi |_{\mathrm{AA_{in}}} \otimes \mathbb{1}_{\mathrm{B_{out}}} \right) (K_{ij} |i\rangle\langle j|_{\mathrm{A}} \otimes \Phi_{\mathrm{A_{in}B_{out}}}) \left(| \Phi \rangle_{\mathrm{AA_{in}}} \otimes \mathbb{1}_{\mathrm{B_{out}}} \right) \\
&= K_{ij} |i\rangle\langle j|_{\mathrm{B_{out}}}.
\end{aligned}$$

The teleportation protocol works with any Bell state, the only difference is that the correcting Weyl operator will need to be adjusted. The implication of the quantum teleportation protocol is that quantum channels need not be tangible (as an optical fiber, or a system ready to receive photons at any time) but it can be stored in memory like any other quantum state. Maximally entangled state can be shared and then stored at a later time, effectively storing identity channels for later use.

1.5.4 Four qudits - *BNOT*

We conclude our background on Bell states with the bilateral *CNOT*, the unitary on four qudits tensor product of two *CNOT*'s. We let d be again the dimension of the four systems,



(1.12)

Figure 1.2: Effect of the BNOT on Bell states.

and to keep track of them we will name $\mathrm{C} = \mathrm{C}_1\mathrm{C}_2$ the two control qudits, and $\mathrm{T} = \mathrm{T}_1\mathrm{T}_2$ the two target qudits. Then the bilateral *CNOT*, or *BNOT* for short, is defined as:

$$BNOT_{\mathrm{CT}} := CNOT_{\mathrm{C}_1\mathrm{T}_1} \otimes CNOT_{\mathrm{C}_2\mathrm{T}_2}$$
$$\mathcal{BNOT}_{\mathrm{CT}}(\rho) := BNOT_{\mathrm{CT}} \cdot \rho \cdot BNOT^{\dagger}_{\mathrm{CT}}.$$

This unitary was introduced for $d = 2$ in [Ben+96b], where it was called Bilateral XOR (BXOR). Like in [Ben+96c], our interest in this gate lies in its effect on Bell states.

Lemma 2 ([1]). *For all $i, j, k, l \in \mathbb{Z}_d$:*

$$BNOT_{\mathrm{CT}}\Big(|\phi_{ij}\rangle_{\mathrm{C}} \otimes |\phi_{kl}\rangle_{\mathrm{T}}\Big) = |\phi_{i,j-l}\rangle_{\mathrm{C}} \otimes |\phi_{k+i,l}\rangle_{\mathrm{T}} \tag{1.13}$$

Lemma 2 follows because $\Phi \otimes \Phi$ is invariant under the action of the *BNOT*, namely

$$BNOT_{\mathrm{CT}} \cdot (|\Phi\rangle_{\mathrm{C}} \otimes |\Phi\rangle_{\mathrm{T}}) = |\Phi\rangle_{\mathrm{C}} \otimes |\Phi\rangle_{\mathrm{T}}\,. \tag{1.14}$$

Applying Equation (1.14) and the commutation relation of the *CNOT* Equation (1.11) to the definition of the Bell states in Equation (1.9), proves Lemma 2, as displayed in Figure 1.2. From Lemma 2, it follows in particular that

$$\mathcal{BNOT}(\phi_{00} \otimes \phi_{0j}) = \phi_{0j} \otimes \phi_{0j}. \tag{1.15}$$

This will be our crucial tool in the simplification of private states in Chapter 3.

Chapter 2

Quantum Information

Often in quantum communication problems, such as key distillation considered in this thesis, the goal of Alice and Bob is to share some specific target states. However, because of noise in the input, the target states cannot be produced exactly in general. Instead, the target states must be approximated as well as possible from what is given as input using LOCC. To even define what it means to approximate these states, we need some distance measure. However, there are many distance measures. For distillation the relevant distance needs to measure how different two states are in the space of quantum states. But other times we only need to measure how different the observations, and thus the outcomes of measurements, are on the states. Sometimes the measurements are restricted, like in the case of bipartite systems, or they are only partial measurements. The purpose of this section is to present the distances induced by the trace norm, the Holevo information and the relative entropy, and to present their relevant properties. These will be used extensively throughout the thesis, to describe the connection between distinguishability, entanglement distillation and distillation of key across repeaters in the upcoming chapters.

2.1 Trace norm

As a norm, the trace norm induces the trace distance $\frac{1}{2}\|\rho - \sigma\|_1$ for $\rho, \sigma \in \mathcal{D}(\mathrm{H})$. By definition the norm of any state is one, and thus $\|\rho - \sigma\|_1 \leqslant 2$, which is saturated for orthogonal states. The trace distance of ρ from a subset of states $\mathcal{K} \subseteq \mathcal{D}(\mathrm{H})$ is defined as:

$$\|\rho - \mathcal{K}\|_1 := \inf_{\sigma \in \mathcal{K}} \|\rho - \sigma\|_1$$

and the distance of two subsets $\mathcal{K}, \mathcal{K}' \subseteq \mathcal{D}(\mathrm{H})$ as

$$\|\mathcal{K}' - \mathcal{K}\|_1 := \inf_{\rho \in \mathcal{K}'} \|\rho - \mathcal{K}\|_1.$$

Let $\varepsilon > 0$, we say that two states ρ and σ are ε close if their trace distance is lower than ε, for which use the notation

$$\rho \approx_\varepsilon \sigma \qquad \Leftrightarrow \qquad \frac{1}{2}\|\rho - \sigma\|_1 < \varepsilon.$$

Similarly we say that ρ and $\mathcal{K}$ are ε close if their distance it lower than ε:

$$\rho \approx_\varepsilon \mathcal{K} \qquad \Leftrightarrow \qquad \frac{1}{2}\|\rho - \mathcal{K}\|_1 < \varepsilon$$

$$\mathcal{K}' \approx_\varepsilon \mathcal{K} \qquad \Leftrightarrow \qquad \frac{1}{2}\|\mathcal{K}' - \mathcal{K}\|_1 < \varepsilon.$$

The trace norm is a monotone under quantum channels, namely for all channels $\Lambda \in \mathsf{CPTP}(\mathrm{H}\rangle\mathrm{H}')$ we have:

$$\|\Lambda(\rho) - \Lambda(\sigma)\|_1 \leqslant \|\rho - \sigma\|_1,$$

and similarly for $\|\Lambda(\rho) - \Lambda(\mathcal{K})\|_1$.

2.1.1 Restricted norms

Because of the monotonicity of the trace distance, if we look from the outcomes of a measurement at two states, they might look closer than they are. However the outcomes of a measurement are the only thing we can physically access to estimate their distance. If we then have a restriction on our measurements, it makes sense to compute the maximum distance achieved under such restriction. Therefore we have the following definition, generalized to include partial measurements.

Definition 3 (L norm [MWW09; 2]). *Let L be a subset of channels on H. For any Hermitian operator $X \in \mathcal{H}(\mathrm{H})$, the trace norm in restriction to L, or L norm, is defined as:*

$$\|X\|_{\mathsf{L}} := \sup_{\Lambda \in \mathsf{L}} \|\Lambda(X)\|_1. \tag{2.1}$$

Let now $\mathcal{K} \subseteq \mathcal{D}(\mathrm{H})$ be a subset of states. For any state $\rho \in \mathcal{D}(\mathrm{H})$, the trace distance from $\mathcal{K}$ in restriction to L, or L distance from $\mathcal{K}$, is defined as:

$$\|\rho - \mathcal{K}\|_{\mathsf{L}} := \inf_{\sigma \in \mathcal{K}} \|\rho - \sigma\|_{\mathsf{L}}.$$

Notice that we defined the L norm on general Hermitian operators, but for our purposes they can be thought of as just the difference of two states. $\|\cdot\|_{\mathsf{L}}$ is also convex, and if L is informationally complete, then $\|X\|_{\mathsf{L}} = 0$ if and only if $X = 0$, and the L norm is indeed a norm. If L contains the identity channel, then the latter is always the optimal channel. In such case the above definition is not very interesting. However, if we consider partial measurements as defined in Section 1.4, then they all exclude the identity. When L is some form of PPT, SEP or LOCC measurement or partial measurement, then we will generally refer to $\|\cdot\|_{\mathsf{L}}$ as a "local norm".

Since the trace norm is just the trace of the positive part and the negative part, namely $\|X\|_1 = \operatorname{tr}|X| = \operatorname{tr} X_+ + \operatorname{tr} X_-$, and because said parts are orthogonal, it turns out that the trace norm can always be achieved with a projective measurement (which depends on X), see [MWW09] for further details. Formally we have

$$\|X\|_1 = \|X\|_{\mathsf{M}(\underline{\mathrm{H}})} = \|X\|_{\mathsf{M}\left(\mathrm{H}\rangle\underline{\mathbb{C}^2}\right)}. \tag{2.2}$$

Let us emphasize that, in restricting to measurements, the trace distance on two states $\|\mathcal{M}(\rho) - \mathcal{M}(\sigma)\|_1$ is equal to its classical counterpart (the statistical distance) on the measurement outcomes.

If L is a subset of measurements, then the L norm can always be expressed in the following convenient form [MWW09]:

$$\|X\|_{\mathsf{L}} = \sup_{T \in K_{\mathsf{L}}} \operatorname{tr}(TX), \tag{2.3}$$

where $K_{\mathsf{L}} \subseteq B_\infty$ is a symmetric convex body, generated by the course graining of the measurement operators of measurements in L. The set K_{L} always contains $\pm\mathbb{1}_{\mathrm{H}}$, since unless L is empty, the value of the L norm on states is always one ($\|\rho\|_{\mathsf{L}} = 1$). By construction, the polar K_{L}° is then the unit ball for $\|\cdot\|_{\mathsf{L}}$ and we have $\|\cdot\|_{K_{\mathsf{L}}^\circ} = \|\cdot\|_{\mathsf{L}}$. When there exists a positive semidefinite closed convex cone $\mathbb{R}_+\mathcal{C} \subseteq \mathcal{H}_+(\mathrm{H})$ that generates L, namely such that L is all the measurements with measurement operators in $\mathbb{R}_+\mathcal{C}$, then the symmetric convex body K_{L} simplifies to

$$K_{\mathsf{L}} = \{\mathbb{R}_+\mathcal{C} - \mathbb{1}_{\mathrm{H}}\} \cap \{\mathbb{1}_{\mathrm{H}} - \mathbb{R}_+\mathcal{C}\} \tag{2.4}$$

2.1.2 Bipartite norms

If L is a class of bipartite channels closed under tensor product, we then find that the local norms are super-multiplicative, namely for any Hermitian operators $X \in \mathcal{H}(\mathrm{A'B'})$ and $Y \in \mathcal{H}(\mathrm{AB})$ we have first

$$\|Y\|_{\mathsf{L}(\underline{\mathrm{A}}:\underline{\mathrm{B}})}\|X\|_{\mathsf{L}\left(\underline{\mathrm{A}}':\underline{\mathrm{B}}'\right)} \leqslant \|Y \otimes X\|_{\mathsf{L}\left(\underline{\mathrm{A}}\underline{\mathrm{A}}':\underline{\mathrm{B}}\underline{\mathrm{B}}'\right)}$$

from which it follows that (using $\|\rho\|_{\mathsf{L}} = 1$)

$$\|X\|_{\mathsf{L}\left(\underline{\mathrm{A}}':\underline{\mathrm{B}}'\right)} \leqslant \|\rho \otimes X\|_{\mathsf{L}\left(\underline{\mathrm{A}}\underline{\mathrm{A}}':\underline{\mathrm{B}}\underline{\mathrm{B}}'\right)} \tag{2.5}$$

and that

$$K_{\mathsf{L}(\underline{\mathrm{A}}:\underline{\mathrm{B}})} \otimes K_{\mathsf{L}\left(\underline{\mathrm{A}}':\underline{\mathrm{B}}'\right)} \subset K_{\mathsf{L}\left(\underline{\mathrm{A}}\underline{\mathrm{A}}':\underline{\mathrm{B}}\underline{\mathrm{B}}'\right)}.$$

If L is a class of bipartite channels that contains classical communication from Alice to Bob, then an immediate consequence of Equation (2.2), namely the fact that measurements achieve the trace norm, is that by Lemma 1 we have [2]

$$\|X\|_{\mathsf{L}(\underline{\mathrm{A}}:\underline{\mathrm{B}})} = \|X\|_{\mathsf{L}(\underline{\mathrm{A}}:\mathrm{B})}. \tag{2.6}$$

as showed by

$$\begin{aligned}
\|X\|_{\mathsf{L}(\underline{\mathrm{A}}:\mathrm{B})} &= \sup_{\Lambda \in \mathsf{L}(\underline{\mathrm{A}}:\mathrm{B})} \|\Lambda(X)\|_{\mathsf{M}(\underline{\mathrm{H}})} \\
&= \sup_{\mathrm{A'B'M}} \sup_{\Lambda \in \mathsf{L}\left(\mathrm{A}:\mathrm{B}\rangle \underline{\mathrm{A}}':\mathrm{B}'\right)} \sup_{\mathcal{M} \in \mathsf{M}\left(\underline{\mathrm{A}}'\underline{\mathrm{B}}'\right)|\underline{\mathrm{M}})} \|\mathcal{M} \circ \Lambda(X)\|_1 \\
&= \sup\left\{\|\Lambda(X)\|_1 : \Lambda \in \bigcup\nolimits_{\mathrm{M,A',B'}} \mathsf{M}\left(\underline{\mathrm{M}}\langle \mathrm{A'B'}\right) \circ \mathsf{L}\left(\underline{\mathrm{A}}':\mathrm{B}'\langle \mathrm{A}:\mathrm{B}\right)\right\} \\
&= \sup_{\Lambda \in \mathsf{L}(\underline{\mathrm{A}}:\underline{\mathrm{B}})} \|\Lambda(X)\|_1 \\
&= \|X\|_{\mathsf{L}(\underline{\mathrm{A}}:\underline{\mathrm{B}})}.
\end{aligned}$$

There would thus be no need to define restricted trace distances for channels other than measurements. The necessity of restricting beyond sets of measurements is imposed by the relative entropy, where the analogue of Equation (2.2) does not hold. There again, the definition will be interesting when the identity channel is excluded.

2.1.3 Separable and PPT norms

The separable and PPT measurements $\mathsf{SEP}(\underline{\mathrm{A}}:\underline{\mathrm{B}})$ and $\mathsf{PPT}(\underline{\mathrm{A}}:\underline{\mathrm{B}})$ can indeed be characterized as the measurements with measurement operators in the cones generated by $\mathcal{S}(\mathrm{A}:\mathrm{B})$ and $\mathcal{P}(\mathrm{A}:\mathrm{B})$, respectively. By Equation (2.4), the associated convex bodies $K_{\mathsf{SEP}(\underline{\mathrm{A}}:\underline{\mathrm{B}})}$ and $K_{\mathsf{PPT}(\underline{\mathrm{A}}:\underline{\mathrm{B}})}$ are

$$\begin{aligned}
K_{\mathsf{SEP}(\underline{\mathrm{A}}:\underline{\mathrm{B}})} &= \{\mathbb{R}_+\mathcal{S}(\mathrm{A}:\mathrm{B}) - \mathbb{1}_{\mathrm{AB}}\} \cap \{\mathbb{1}_{\mathrm{AB}} - \mathbb{R}_+\mathcal{S}(\mathrm{A}:\mathrm{B})\}, \\
K_{\mathsf{PPT}(\underline{\mathrm{A}}:\underline{\mathrm{B}})} &= \{\mathbb{R}_+\mathcal{P}(\mathrm{A}:\mathrm{B}) - \mathbb{1}_{\mathrm{AB}}\} \cap \{\mathbb{1}_{\mathrm{AB}} - \mathbb{R}_+\mathcal{P}(\mathrm{A}:\mathrm{B})\} \\
&= B_\infty(\mathrm{AB}) \cap B_\infty(\mathrm{AB})^{\Gamma}.
\end{aligned}$$

As we would expect, since separable and PPT operations are defined as those operations that preserve separability and PPTness, separable and PPT states do not increase the respective local norms.

Fact 4. *Let* $(\mathsf{L},\mathcal{K})$ *be either* $(\mathsf{PPT},\mathcal{P})$ *or* $(\mathsf{SEP},\mathcal{S})$*. For any state* $\rho \in \mathcal{K}(\mathrm{A{:}B})$ *and any Hermitian operator* $X \in \mathcal{H}(\mathrm{A'B'})$ *we have*

$$\|X\|_{\mathsf{L}(\underline{\mathrm{A}}':\underline{\mathrm{B}}')} = \|\rho\otimes X\|_{\mathsf{L}(\underline{\mathrm{A}}\underline{\mathrm{A}}':\underline{\mathrm{B}}\underline{\mathrm{B}}')}.$$

Proof. One inequality is given by Equation (2.5). For the opposite inequality, we have that the state preparation channel $\Lambda(X) = \rho\otimes X$ is in $\mathsf{L}(\mathrm{A'{:}B'}\rangle\mathrm{AA'{:}BB'})$, namely it is a local channel, and therefore for any Hermitian operator X on $\mathrm{A'B'}$,

$$\|\rho\otimes X\|_{\mathsf{L}(\underline{\mathrm{A}}\underline{\mathrm{A}}':\underline{\mathrm{B}}\underline{\mathrm{B}}')} = \|\Lambda(X)\|_{\mathsf{L}(\underline{\mathrm{A}}':\underline{\mathrm{B}}')} \leqslant \|X\|_{\mathsf{L}(\underline{\mathrm{A}}':\underline{\mathrm{B}}')}$$

which proves the claim. □

Since not all states increase the local norms, it is then interesting to investigate how much of an increase is possible. Namely, we study how tensoring a Hermitian operator X with a state ρ changes the PPT or SEP norm. More precisely, we are now interested in finding upper bounds on $\|\rho\otimes X\|_{\mathsf{L}(\underline{\mathrm{A}}\underline{\mathrm{A}}':\underline{\mathrm{B}}\underline{\mathrm{B}}')}$ in terms of $\|X\|_{\mathsf{L}(\underline{\mathrm{A}}':\underline{\mathrm{B}}')}$. For our statements we need the robustness of entanglement [VT99], which for any state ρ on AB is defined as

$$\mathcal{R}_{\mathrm{A:B}}(\rho) := \inf_{\sigma\in\mathcal{S}(\mathrm{A:B})} \mathcal{R}_{\mathrm{A:B}}(\rho\|\sigma). \tag{2.7}$$

where

$$\mathcal{R}_{\mathrm{A:B}}(\rho\|\sigma) := \inf\left\{ s : \frac{1}{1+s}(\rho + s\sigma) \in \mathcal{S}(\mathrm{A:B})\right\}.$$

This is not to be confused with the global robustness of entanglement where σ is allowed to vary over all states in $\mathcal{D}(\mathrm{AB})$ [Dat09].

Proposition 5 ([2]). *For any Hermitian operator* X *on* $\mathrm{A'B'}$ *and any state* ρ *on* AB*, we have*

$$\begin{aligned}\|\rho\otimes X\|_{\mathsf{SEP}(\underline{\mathrm{A}}\underline{\mathrm{A}}':\underline{\mathrm{B}}\underline{\mathrm{B}}')} &\leqslant (2\mathcal{R}_{\mathrm{A:B}}(\rho)+1)\|X\|_{\mathsf{SEP}(\underline{\mathrm{A}}':\underline{\mathrm{B}}')},\\ \|\rho\otimes X\|_{\mathsf{PPT}(\underline{\mathrm{A}}\underline{\mathrm{A}}':\underline{\mathrm{B}}\underline{\mathrm{B}}')} &\leqslant \|\rho^{\Gamma}\|_1 \|X\|_{\mathsf{PPT}(\underline{\mathrm{A}}':\underline{\mathrm{B}}')}.\end{aligned}$$

Setting $k = \min(|\mathrm{A}|,|\mathrm{B}|)$*, we therefore have*

$$\begin{aligned}\|\rho\otimes X\|_{\mathsf{SEP}(\underline{\mathrm{A}}\underline{\mathrm{A}}':\underline{\mathrm{B}}\underline{\mathrm{B}}')} &\leqslant (2k-1)\|X\|_{\mathsf{SEP}(\underline{\mathrm{A}}':\underline{\mathrm{B}}')},\\ \|\rho\otimes X\|_{\mathsf{PPT}(\underline{\mathrm{A}}\underline{\mathrm{A}}':\underline{\mathrm{B}}\underline{\mathrm{B}}')} &\leqslant k\|X\|_{\mathsf{PPT}(\underline{\mathrm{A}}':\underline{\mathrm{B}}')}.\end{aligned}$$

Proof. The second set of inequalities in the proposition is easily derived from the first one, after upper bounding the maximal value of $\mathcal{R}(\rho)$ and $\|\rho^{\Gamma}\|_1$. The fact that $\|\rho^{\Gamma}\|_1 \leqslant k$ is well-known, see also [VW02], while it was shown in [VT99, Theorem C.2] that $\mathcal{R}(\rho)\leqslant k-1$. In both cases the maximal value is achieved by Φ_{AB} (and its generalization to A and B with different dimensions).

For the SEP norm, we follow an argument inspired by [LPW18, Theorem 16]. We know that there exists a separable state σ which is such that, setting $r := \mathcal{R}_{\mathrm{A:B}}(\rho)$, the following state is also separable:

$$\rho' = \frac{1}{1+r}\rho + \frac{r}{1+r}\sigma.$$

Because the SEP norm is left unchanged under tensoring with a separable state, we have

$$\begin{aligned}\|X\|_{\mathsf{SEP}(\underline{\mathrm{A}}':\underline{\mathrm{B}}')} &= \|\rho'\otimes X\|_{\mathsf{SEP}(\underline{\mathrm{A}}\underline{\mathrm{A}}':\underline{\mathrm{B}}\underline{\mathrm{B}}')}\\ &\geqslant \frac{1}{1+r}\|\rho\otimes X\|_{\mathsf{SEP}(\underline{\mathrm{A}}\underline{\mathrm{A}}':\underline{\mathrm{B}}\underline{\mathrm{B}}')} - \frac{r}{1+r}\|\sigma\otimes X\|_{\mathsf{SEP}(\underline{\mathrm{A}}\underline{\mathrm{A}}':\underline{\mathrm{B}}\underline{\mathrm{B}}')}\\ &= \frac{1}{1+r}\|\rho\otimes X\|_{\mathsf{SEP}(\underline{\mathrm{A}}\underline{\mathrm{A}}':\underline{\mathrm{B}}\underline{\mathrm{B}}')} - \frac{r}{1+r}\|X\|_{\mathsf{SEP}(\underline{\mathrm{A}}':\underline{\mathrm{B}}')},\end{aligned}$$

where we used the triangle inequality and the separability of ρ' and σ. Hence, we obtain as claimed that $(2r+1)\|X\|_{\mathsf{SEP}(\underline{A}':\underline{B}')} \geqslant \|\rho \otimes X\|_{\mathsf{SEP}(\underline{A}\underline{A}':\underline{B}\underline{B}')}$.

For the PPT, norm notice first that the polar of $B_\infty \cap B_\infty^\Gamma$ is simply

$$\mathrm{conv}\left(B_1 \cup B_1^\Gamma\right) = \{\lambda Y + (1-\lambda)Z : \ \|Y\|_1, \|Z^\Gamma\|_1 \leqslant 1,\ \lambda \in [0,1]\}.$$

Therefore we have

$$\begin{aligned}
\|\rho \otimes X\|_{\mathsf{PPT}(\underline{A}\underline{A}':\underline{B}\underline{B}')} &= \inf\left\{\mu : \rho \otimes X \in \mu\,\mathrm{conv}\left(B_1(\mathrm{AA'BB'}) \cup B_1(\mathrm{AA'BB'})^\Gamma\right)\right\} \\
&= \inf\{\mu : \rho \otimes X = \lambda Y + (1-\lambda)Z \wedge \|Y\|_1, \|Z^\Gamma\|_1 \leqslant \mu \wedge \lambda \in [0,1]\} \\
&= \inf\{\max(\|Y\|_1, \|Z^\Gamma\|_1) : \rho \otimes X = \lambda Y + (1-\lambda)Z \wedge \lambda \in [0,1]\},
\end{aligned}$$

where $Y, Z \in \mathcal{L}(\mathrm{AA'BB'})$. Now, let $X = \lambda_0 Y_0 + (1-\lambda_0)Z_0$ such that $\|X\|_{\mathsf{PPT}(\underline{A}':\underline{B}')} = \max\left(\|Y_0\|_1, \|Z_0^\Gamma\|_1\right)$ with $Y_0, Z_0 \in \mathcal{L}(\mathrm{A'B'})$, as just derived. Since $\rho \otimes X = \lambda_0 \rho \otimes Y_0 + (1-\lambda_0)\rho \otimes Z_0$, we then have

$$\begin{aligned}
\|\rho \otimes X\|_{\mathsf{PPT}(\underline{A}\underline{A}':\underline{B}\underline{B}')} &\leqslant \max\left(\|\rho \otimes Y_0\|_1, \|(\rho \otimes Z_0)^\Gamma\|_1\right) \\
&= \max\left(\|\rho\|_1\|Y_0\|_1, \|\rho^\Gamma\|_1\|Z_0^\Gamma\|_1\right) \\
&\leqslant \|\rho^\Gamma\|_1 \max\left(\|Y_0\|_1, \|Z_0^\Gamma\|_1\right) \\
&= \|\rho^\Gamma\|_1 \|X\|_{\mathsf{PPT}(\underline{A}':\underline{B}')},
\end{aligned}$$

the first equality being by multiplicativity of the trace norm under tensoring and the second inequality being because $\|\rho^\Gamma\|_1 \geqslant \|\rho\|_1$. □

A legitimate question at this point is whether the inequalities in Proposition 5 are tight. In the next proposition we show that at least the scaling is optimal, in the sense that we can always find a Hermitian operator that can take full advantage of Φ_{AB} and achieve an almost optimal increase in local norm.

Proposition 6. *For any $k = |\mathrm{A}| = |\mathrm{B}|$ there exists a Hermitian operator $X \in \mathcal{H}(\mathrm{A'B'})$ such that for* L *being either* SEP *or* PPT *it holds that*

$$\|\Phi_{\mathrm{AB}} \otimes X\|_{\mathsf{L}(\underline{A}\underline{A}':\underline{B}\underline{B}')} \geqslant \frac{k+1}{2}\|X\|_{\mathsf{L}(\underline{A}':\underline{B}')}$$

Proof. Let L be either SEP or PPT, and let $|\mathrm{A}'| = |\mathrm{B}'| = k$. Then, we know from [DLT02] that the following holds: the symmetric and antisymmetric states ρ_s and ρ_a are orthogonal, but they are close under PPT and thus separable measurements. However, $\Phi \otimes \rho_s$ and $\Phi \otimes \rho_a$ are orthogonal under one-way LOCC measurements, and thus under separable and PPT measurements, because Φ can be used to teleport the state wholly to one of the parties, where a global projective measurement can be performed yielding orthogonal outcomes. Quantitatively, setting $X = \rho_s - \rho_a$, we have $\|\Phi \otimes X\|_{\mathsf{L}} = 2$ and $\|X\|_{\mathsf{L}} = 4/(k+1)$, which concludes the proof. □

2.2 Holevo information

We have just mentioned an example of what are called data hiding states in Proposition 6, namely the symmetric and antisymmetric states as an example of states that are orthogonal, but $O(1/d)$-close under separable or PPT measurements. Since they are orthogonal, they can be used to encode a bit of classical information that can be recovered using a global measurement. We can think of somebody wanting to send a classical bit as a message to Alice and doing it by sending either the symmetric or the antisymmetric state (the encoding). If Alice receives the whole state, she can read the message with a projective measurement (the decoding). Symmetric and antisymmetric states are two qudit states, and thus we can split the message into two parties Alice and Bob, but since they are close under PPT measurements, most of the time Alice and Bob will get outcomes that are uncorrelated with the original classical message. (In particular the best they can do is perform a measurement in the computational basis and conclude that they shared the symmetric state if the outcomes are correlates, something that happens only with probability $O(1/d)$). Thus Alice and Bob cannot read the message unless they join together or use quantum communication. This is called quantum data hiding [TDL01; DLT02; EW02]. In this section we study the relevant measure of information and relate it to the trace distance.

2.2.1 Entropies

We define the following commonly used real functions: $\eta(x) := -x \log x$ defined for $x \geqslant 0$ by setting $\eta(0) = 0$, the binary entropy $h(x) := \eta(x) + \eta(1-x)$ defined for $x \in [0,1]$, and $g(\varepsilon) := \eta(x) - \eta(1+x) = (1+x)h(x/(1+x))$ defined for $x \geqslant 0$; notice that g is monotonically increasing. The logarithm log will always be in base two unless otherwise stated. Let $K \in \mathcal{H}_+(\mathrm{A})$ with eigenvalue decomposition $K = \sum_{i \in \mathbb{Z}_{|\mathrm{A}|}} p_i |v_i\rangle\langle v_i|$. We define $\eta(K)$ by defining it on the eigenvalues:

$$\eta(K) = \sum_{i \in \mathbb{Z}_{|\mathrm{A}|}} \eta(p_i) |v_i\rangle\langle v_i| \,.$$

The von Neumann entropy, or simply entropy, of a state $\rho \in \mathcal{D}(\mathrm{A})$ is then defined as

$$H(\rho) := \operatorname{tr}_{\mathrm{A}} \eta(\rho) = \sum_{i \in \mathbb{Z}_{|\mathrm{A}|}} \eta(p_i)$$

which is always positive, and in particular zero only for pure states, and maximal for the maximally mixed state. It is also additive on tensor products, namely $H(\rho \otimes \sigma) = H(\rho) + H(\sigma)$, and unitary invariant. A common notation for the entropy of states on multiple systems is $H(\mathrm{A})_\rho \equiv H(\rho_{\mathrm{A}})$, that allows to make statements for the entropy function $H(\mathrm{A})$ independently of the actual state. If $\rho \in \mathcal{D}(\mathrm{AB})$ is a pure state, then $H(\mathrm{A})_\rho = H(\mathrm{B})_\rho$. Let now $\rho \in \mathcal{D}(\mathrm{ABC})$; using the entropy we then have:

- The conditional entropy $H(\mathrm{A}|\mathrm{B})$ and coherent information $I(\mathrm{A}\rangle\mathrm{B})$

$$H(\mathrm{A}|\mathrm{B})_\rho := H(\mathrm{AB})_\rho - H(\mathrm{B})_\rho =: -I(\mathrm{A}\rangle\mathrm{B})_\rho.$$

- The mutual information

$$I(\mathrm{A{:}B})_\rho := H(\mathrm{A})_\rho + H(\mathrm{B})_\rho - H(\mathrm{AB})_\rho.$$

  The mutual information is always positive (sub-additivity) and is manifestly symmetric on A and B.

- The conditional mutual information

$$\begin{aligned} I(\mathrm{A{:}B}|\mathrm{C})_\rho &:= H(\mathrm{AC})_\rho + H(\mathrm{BC})_\rho - H(\mathrm{C})_\rho - H(\mathrm{ABC})_\rho \\ &= H(\mathrm{A}|\mathrm{C})_\rho + H(\mathrm{B}|\mathrm{C})_\rho - H(\mathrm{AB}|\mathrm{C})_\rho \\ &= I(\mathrm{A{:}BC})_\rho - I(\mathrm{A{:}C})_\rho \\ &= I(\mathrm{A}\rangle\mathrm{BC})_\rho - I(\mathrm{A}\rangle\mathrm{C})_\rho. \end{aligned}$$

> The conditional information is also always positive (strong sub-additivity), and is again manifestly symmetric on A and B. The purifying system is always an equivalent choice for conditioning, namely for pure states $|\psi\rangle\langle\psi| \in \mathcal{D}(\mathrm{ABCE})$, it holds $I(\mathrm{A{:}B|C})_{|\psi\rangle\langle\psi|} = I(\mathrm{A{:}B|E})_{|\psi\rangle\langle\psi|}$ [DY08]. In particular for pure states $|\psi\rangle\langle\psi| \in \mathcal{D}(\mathrm{ABC})$, the conditional mutual information reduces to the mutual information, $I(\mathrm{A{:}B|C})_{|\psi\rangle\langle\psi|} = I(\mathrm{A{:}B})_{|\psi\rangle\langle\psi|}$.

Strong sub-additivity implies that mutual information and conditional mutual information are monotone (decreasing) under tracing out systems at Alice or Bob, while the coherent information is monotone under tracing out systems at Bob:

$$\begin{aligned} I(\mathrm{A}\rangle\mathrm{BB'}) - I(\mathrm{A}\rangle\mathrm{B}) &= I(\mathrm{A{:}B|B}) \geqslant 0 \\ I(\mathrm{A{:}BB'}) - I(\mathrm{A{:}B}) &= I(\mathrm{A{:}B'|B}) \geqslant 0 \\ I(\mathrm{A{:}BB'|E}) - I(\mathrm{A{:}B|E}) &= I(\mathrm{A{:}B'|BE}) \geqslant 0. \end{aligned}$$

Together with Stinespring's dilation, the additivity of the entropy on tensor products, and the unitary invariance of the entropy, the above implies monotonicity under local operations at Alice and Bob for the mutual information and the conditional mutual information, and at Bob for the coherent information. Namely, for any local channel at Bob $\Lambda \in \mathrm{id}_{\mathrm{A}} \otimes \mathrm{CPTP}(\mathrm{B_{in}}\rangle\mathrm{B_{out}})$, we have

$$\begin{aligned} I(\mathrm{A}\rangle\mathrm{B_{out}})_{\Lambda(\rho)} &\leqslant I(\mathrm{A}\rangle\mathrm{B_{in}})_{\rho} \\ I(\mathrm{A{:}B_{out}})_{\Lambda(\rho)} &\leqslant I(\mathrm{A{:}B_{in}})_{\rho} \\ I(\mathrm{A{:}B_{out}|E})_{\Lambda(\rho)} &\leqslant I(\mathrm{A{:}B_{in}|E})_{\rho}. \end{aligned}$$

The coherent information $I(\mathrm{A}\rangle\mathrm{B})$ is not monotone under local operations at A, as can be verified by adding local randomness.

The conditional information and the coherent information satisfy

$$|H(\mathrm{A|B})| = |I(\mathrm{A}\rangle\mathrm{B})| \leqslant \log|\mathrm{A}|,$$

which can be proven using a purifying system E. The extremes are achieved by the maximally entangled state and the maximally mixed state. The consequence for both the mutual and conditional mutual information is

$$I(\mathrm{A{:}B}), I(\mathrm{A{:}B|E}) \leqslant 2\min\{\log|\mathrm{A}|, \log|\mathrm{B}|\}\,.$$

which is saturated for the maximally entangled state on Alice and Bob.

2.2.2 Restricted Holevo Information

Let M be the system for a classical message, so that $|\mathrm{M}|$ is the size of the message, and let H be a system for encoding the message into. More precisely, we have now a classical source $\pi = \sum_i p_i |i\rangle\langle i|$, where p_i are probabilities, and a quantum encoding $\mathcal{E}$ that maps $|i\rangle\langle i| \mapsto \sigma_i \in \mathcal{D}(\mathrm{H})$. The amount of information about the classical source contained by the quantum encoding is the mutual information between the message and the encoding states, namely the mutual information of:

$$\xi := \sum_{i\in\mathbb{Z}_{|\mathrm{M}|}} p_i |i\rangle\langle i| \otimes \mathcal{E}(|i\rangle\langle i|) = \sum_{i\in\mathbb{Z}_{|\mathrm{M}|}} p_i |i\rangle\langle i| \otimes \sigma_i \in \mathcal{D}(\mathrm{MH}).$$

This is known as the Holevo information χ [Hol73] of the encoded source ensemble:

$$\chi\Big(\{p_i,\sigma_i\}_{i\in\mathbb{Z}_{|\mathrm{M}|}}\Big) := H\Big(\textstyle\sum_{i \mathrm{in} \mathbb{Z}_{|\mathrm{M}|}} p_i\sigma_i\Big) - \displaystyle\sum_{i\in\mathbb{Z}_{|\mathrm{M}|}} p_i H(\sigma_i) = I(\mathrm{M{:}H})_{\xi}.$$

The Holevo information is an upper bound on the accessible information, which measures the amount of mutual information that can be actually achieved after decoding back into a classical message, namely after a measurement. The accessible information is thus

$$\chi_{\mathsf{M}(\underline{\mathrm{H}})}\left(\{p_i,\sigma_i\}_{i\in\mathbb{Z}_{|\mathrm{M}|}}\right):=\sup_{\mathcal{M}\in\mathsf{M}(\underline{\mathrm{H}})}\chi\left(\{p_i,\mathcal{M}(\sigma_i)\}_{i\in\mathbb{Z}_{|\mathrm{M}|}}\right)$$

and since the mutual information is monotone under local operations, we have

$$\chi_{\mathsf{M}(\underline{\mathrm{H}})}(\{p_i,\sigma_i\})\leqslant\chi(\{p_i,\sigma_i\}).$$

In the case we are interested in, however, we have orthogonal states, which in principle achieve perfect encoding: a global measurement can recover the message and thus the accessible information and the mutual information are the same. Namely, when the encoding states are orthogonal we have

$$\chi_{\mathsf{M}(\underline{\mathrm{H}})}(\{p_i,\sigma_i\})=\chi(\{p_i,\sigma_i\}).$$

Our restrictions come from bipartite systems H = A:B, namely, we will have restrictions in our measurements coming from a class of bipartite channels L such as LOCC. The relevant quantity is then the L locally accessible information [Bad+03], which is defined as

$$\chi_{\mathsf{L}(\underline{\mathrm{A}}:\underline{\mathrm{B}})}(\{p_i,\sigma_i\})[i\in\mathbb{Z}_{|\mathrm{M}|}]:=\sup_{\mathcal{M}\in\mathsf{L}(\underline{\mathrm{A}}:\underline{\mathrm{B}})}\chi(\{p_i,\mathcal{M}(\sigma_i)\})[i\in\mathbb{Z}_{|\mathrm{M}|}]$$

which is again upper bounded by the Holevo information:

$$\chi_{\mathsf{L}(\underline{\mathrm{A}}:\underline{\mathrm{B}})}(\{p_i,\sigma_i\})\leqslant\chi(\{p_i,\sigma_i\}).$$

Whenever the inequality is strict we say that some of the information is hidden to local observers. The term data hiding though, is usually reserved for families of states for which the locally accessible information decreases as we increase the local dimension of the states, like is the case for the symmetric and antisymmetric states. If we allow any class of maps L instead of just measurements, then we define the L Holevo information as

$$\chi_{\mathsf{L}(\mathrm{H})}(\{p_i,\sigma_i\}):=\sup_{\Lambda\in\mathsf{L}(\mathrm{H})}\chi(\{p_i,\Lambda(\sigma_i)\})$$

that again is a sensible definition only when the identity channel is excluded, which would otherwise always be the optimal channels, by monotonicity of the mutual information under local operations.

2.2.3 Asymptotic continuity and Indistinguishability

Recall that the way we stated the indistinguishability of the symmetric and antisymmetric states was in terms of their local norms. To see that this implies indistinguishability in the Holevo information, it is enough to use the asymptotic continuity of the mutual information. Asymptotic continuity is a strong continuity property of entropic quantities, stating that their change is at most linear in the distance between two states and in the number of qubits (namely logarithmically in the dimension of the system). This property has many applications, and the first example we will see now is the asymptotic continuity of the mutual information, which is derived from the asymptotic continuity of the entropy and of the coherent information (conditional entropy) as follows. Let H be any system, let ρ and $\tilde{\rho}$ be two states on H, and let $\varepsilon:=\frac{1}{2}\|\rho-\tilde{\rho}\|_1$, namely let ρ and $\tilde{\rho}$ be two ε-close states.

For the entropy we have the sharp version of Fannes inequality [Fan73; Aud07; Zha07; Pet07]

$$\left|H(\mathrm{H})_\rho - H(\mathrm{H})_{\tilde\rho}\right| \leqslant f(\varepsilon) \leqslant \varepsilon \log |\mathrm{H}| + h(\varepsilon),$$

where

$$f(\varepsilon) = \begin{cases} \varepsilon \log(1 - |M|) + h(\varepsilon) & \varepsilon \leqslant 1 - \frac{1}{|\mathrm{H}|} \\ \log |\mathrm{H}| & \varepsilon \geqslant 1 - \frac{1}{|\mathrm{H}|} \end{cases}$$

and we recall that $h(x) = -x \log x - (1-x)\log(1-x)$ is the binary entropy. We mention the bound on $f(\varepsilon)$ because it is monotone. The second bound is not monotone in ε but it is easier to deal with. Since the actual bound $f(\varepsilon)$ is monotone, we can always use its monotonicity first, and then relax to the second bound, to the effect that it will seem that we are claiming monotonicity of the second. Let now ρ and $\tilde\rho$ be ε-close states on HH'. For the coherent information we have the refined version of the Alicki-Fannes inequality [AF04; Win16]

$$\left|I(\mathrm{H}\rangle\mathrm{H}')_\rho - I(\mathrm{H}\rangle\mathrm{H}')_{\tilde\rho}\right| \leqslant 2\varepsilon \log |\mathrm{H}| + g(\varepsilon).$$

where we recall that $g(\varepsilon) = (1+\varepsilon)h(\varepsilon/(1+\varepsilon))$ and we point out that this bound is indeed monotone. Together, since $H(\mathrm{H}) + I(\mathrm{H}\rangle\mathrm{H}')$, they give

$$\left|I(\mathrm{H}{:}\mathrm{H}')_\rho - I(\mathrm{H}{:}\mathrm{H}')_{\tilde\rho}\right| \leqslant 3\varepsilon \log |\mathrm{H}| + h(\varepsilon) + g(\varepsilon).$$

Notice that the bounds above depend on only one of the system dimensions (by symmetry we can pick the smallest dimension for the mutual information).

The importance of scaling linearly in the number of qubits, and thus logarithmically in the dimension, becomes evident if we let ρ and $\tilde\rho$ be families of states. If their distance ε goes to zero polynomially in the dimension, then their mutual information will also go to zero, because the asymptotic continuity will not change the limit of the scaling. We thus define two states ρ and $\tilde\rho$, implicitly a family of states on systems H of increasing dimension, to be indistinguishable if ε goes to zero polynomially in the dimension.

2.2.4 Holevo Information and Norm distances

Let us now go back to the Holevo information. Recall that we defined the classical source $\pi = \sum_i p_i |i\rangle\langle i| \in \mathcal{D}(\mathrm{M})$. Let now $\rho \in \mathcal{D}(\mathrm{H})$ be any state, and let us apply the asymptotic continuity bound to $\pi \otimes \rho$. Let $\Lambda \in \mathsf{L}(\mathrm{H}\rangle\mathrm{H}')$. Clearly the mutual information of both $\pi \otimes \rho$ and $\pi \otimes \Lambda(\rho)$ is zero, being product states, and thus we get

$$I(\mathrm{M}{:}\mathrm{H}')_{\mathrm{id}\otimes\Lambda(\xi)} = \left|I(\mathrm{M}{:}\mathrm{H}')_{\mathrm{id}\otimes\Lambda(\xi)} - I(\mathrm{M}{:}\mathrm{H}')_{\pi\otimes\Lambda(\rho)}\right| \leqslant 3\delta \log |M| + h(\delta) + g(\delta)$$

where ξ defined in Section 2.2.2 is the ensemble correlating the message to the encoding, and where

$$\delta = \frac{1}{2}\|\mathrm{id}\otimes\Lambda(\xi) - \pi\otimes\Lambda(\rho)\|_1.$$

We can then easily check that we can expand the norm and get

$$\|\mathrm{id}\otimes\Lambda(\xi) - \pi\otimes\Lambda(\rho)\|_1 = \sum_i p_i \|\Lambda(\sigma_i) - \Lambda(\rho)\|_1 \leqslant \sum_i p_i \|\sigma_i - \rho\|_{\mathsf{L}(\mathrm{H})}.$$

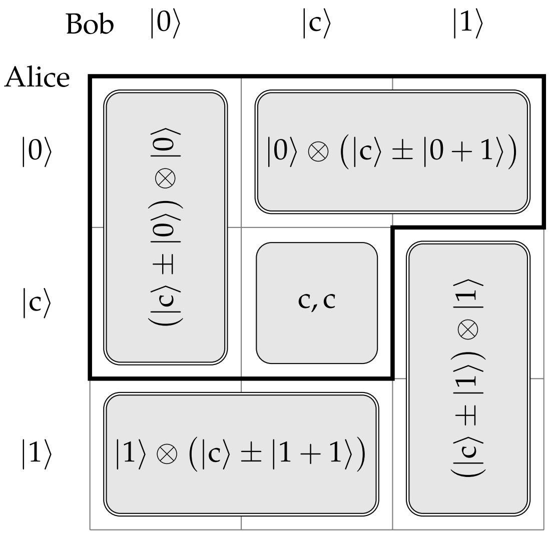


Figure 2.1: Domino states on qutrits.

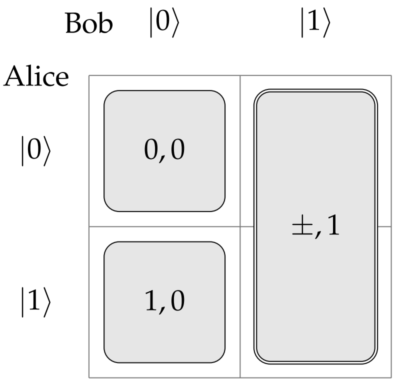


Figure 2.2: Locking states on qubits.

We now define ε to be the minimal average of the trace distances

$$\varepsilon := \inf_{\rho \in \mathcal{D}(\mathrm{H})} \sum_i p_i \frac{1}{2} \|\sigma_i - \rho\|_{\mathsf{L}(\mathrm{H})}$$

and by the monotonicity of the asymptotic continuity bound, and because it holds for any measurement in L, we get

$$\chi_{\mathsf{L}(\mathrm{H})}(\{p_i, \sigma_i\}) \leqslant 3\varepsilon \log |M| + h(\varepsilon) + g(\varepsilon).$$

Thus if there is a state that is on average indistinguishable from all the encoding states, then the L Holevo information, the information that can be recovered using maps in L, must be low. Generally a good guess will be $\sigma = \sum_i p_i \sigma_i$. Notice how the output dimension of the measurement does not affect the bound, because we can make it depend only on the original message size.

In the case of a uniformly-random single bit message $\{\frac{1}{2}, \sigma^{\pm}\}$, like that which is encoded by the symmetric and antisymmetric states with uniform probability, we get a particular simplification. Using $\sigma = \frac{1}{2}\sigma^+ + \frac{1}{2}\sigma^-$ we find

$$\frac{1}{2}\|\sigma^+ - \sigma^-\|_{\mathsf{L}(\underline{\mathrm{A}}:\underline{\mathrm{B}})} = \|\sigma^{\pm} - \sigma\|_{\mathsf{L}(\underline{\mathrm{A}}:\underline{\mathrm{B}})}.$$

The local trace distance of $\sigma^{\pm}$ thus becomes an upper bound on the locally accessible information. This in particular will apply to private bits in Chapter 3.

2.2.5 Domino states

Restricted Holevo informations can be found for example in the proof that LOCC and separable operations are actually different. The fact that at the base of the construction of LOCC we have local operations distinguishes it from separable operations. Indeed, it has been shown that there are encoding bipartite states, for which the Holevo information is achieved by separable measurements, but not by LOCC measurements. The emblematic example of states displaying the distinction, are the so-called domino states [Ben+99]. Let the bipartite system be two qutrits $\mathbb{C}^3{:}\mathbb{C}^3$ and let $\{|0\rangle, |c\rangle, |1\rangle\}$ be a basis for $\mathbb{C}^3$. The domino basis can be defined as

$$\begin{aligned} |\mathrm{c},\mathrm{c}\rangle &:= |c\rangle \otimes |\mathrm{c}\rangle \\ |\psi_{ijk}\rangle &:= S^k\left(\tfrac{1}{\sqrt{2}}(|\mathrm{c}\rangle + (-1)^j\,|i+k\rangle) \otimes |i\rangle\right) \end{aligned}$$

where $i, j, k \in \mathbb{Z}_2$ and S is the swap operator on $\mathbb{C}^3 \otimes \mathbb{C}^3$, which in Figure 2.1 corresponds to a reflection along the main diagonal. The domino states are a product basis and thus the projective measurement in this basis is separable, as the measurement operators are product operators. However what was shown quantitatively in [Ben+99] is that there is no such channel in LOCC, and thus the domino states cannot be perfectly distinguished by two separated parties no matter the amount of communication or number of rounds used. Namely, even ignoring the centre piece $|\mathrm{c},\mathrm{c}\rangle$, we have

$$\chi_{\mathsf{LOCC}(\underline{\mathbb{C}}^3:\underline{\mathbb{C}}^3)}\left(\left\{\tfrac{1}{8},\psi_{ijk}\right\}\right) < \chi_{\mathsf{SEP}(\underline{\mathbb{C}}^3:\underline{\mathbb{C}}^3)}\left(\left\{\tfrac{1}{8},\psi_{ijk}\right\}\right) = \chi\left(\left\{\tfrac{1}{8},\psi_{ijk}\right\}\right). \tag{2.8}$$

2.2.6 Locking states

The LOCC vs SEP result for the Domino states, can be reduced (although only qualitatively) to the difference between $\mathsf{LOCC}_{\rightarrow}$ vs LOCC using simpler states [WH02]. Let Alice and Bob now share a qudit and a qubit respectively, namely the bipartite system $\mathbb{C}^d{:}\mathbb{C}^2$. Let $i \in \mathbb{Z}_d$ and $j \in \mathbb{Z}_2$, we define the locking states, displayed in Figure 2.2 for $d = 2$, as

$$\psi_{ij} := \mathcal{F}^j(|i\rangle\langle i|) \otimes |j\rangle\langle j|$$

where $\mathcal{F}$ is the Fourier transform channel. Essentially Alice has $\log d$ bits of information, but Bob decides whether it is stored in the computational or the conjugate basis. They are called locking states because the lack of a single classical bit makes the loss of information grow boundlessly as d grows [DiV+04; CW05]. Namely we have

$$\chi_{\mathsf{M}(\underline{\mathbb{C}}^d)}\left(\left\{\tfrac{1}{2d},\mathcal{F}^j(|i\rangle\langle i|)\right\}\right) = \chi\left(\left\{\tfrac{1}{2d},\mathcal{F}^j(|i\rangle\langle i|)\right\}\right) = \frac{1}{2}\log d$$

but

$$\chi_{\mathsf{M}(\underline{\mathbb{C}}^{2d})}\left(\left\{\tfrac{1}{2d},\mathcal{F}^j(|i\rangle\langle i|)\otimes|j\rangle\langle j|\right\}\right) = \chi\left(\left\{\tfrac{1}{2d},\mathcal{F}^j(|i\rangle\langle i|)\otimes|j\rangle\langle j|\right\}\right) = 1 + \log d.$$

It was shown in [WH02] for $d = 2$, that the locking states cannot be distinguished perfectly without communication from Bob to Alice. More specifically, Alice cannot do any non-trivial operations, only reversible ones, until she has received the basis information from Bob. Because the domino states contain locking states such as the ones highlighted by the contour in Figure 2.1, it is then argued that neither Alice or Bob can begin the protocol, therefore excluding all LOCC operations. Separable operations thus contain channels that do not begin with any local operation.

Here we compute the $\mathsf{LOCC}_{\rightarrow}$ Holevo information and show that the best Alice and Bob can do is a local measurement.

Lemma 7. *For any choice of* $\mathsf{L} = \mathsf{LO}(\underline{\mathbb{C}}^d{:}\underline{\mathbb{C}}^2), \mathsf{LOCC}_{\rightarrow}(\underline{\mathbb{C}}^d{:}\underline{\mathbb{C}}^2), \mathsf{LOCC}_{\rightarrow}(\underline{\mathbb{C}}^d{:}\mathbb{C}^2)$, *we have*

$$\chi_{\mathsf{L}}\left(\left\{\tfrac{1}{2d},\psi_{ij}\right\}\right) = 1 + \frac{1}{2}\log d$$

Proof. For the lower bound, it is sufficient to check that the mutual information of the local measurement in the computational basis achieves $1 + \frac{1}{2}\log d$. We will thus only show the upper bound for $\mathsf{LOCC}_{\rightarrow}(\underline{\mathbb{C}}^d{:}\mathbb{C}^2)$. Let $\mathrm{A} = \mathbb{C}^d$ and $\mathrm{B} = \mathbb{C}^2$. By the monotonicity of the mutual information under local operations, we can lift the measurement at Bob and by Equation (1.5) we thus have

$$\chi_{\mathsf{LOCC}_{\rightarrow}(\underline{\mathrm{A}}:\mathrm{B})}\left(\left\{\tfrac{1}{2d}, \psi_{ij}\right\}\right) \leqslant \chi_{\mathsf{M}(\underline{\mathrm{A}})\otimes\mathrm{id}_{\mathrm{B}}}\left(\left\{\tfrac{1}{2d}, \psi_{ij}\right\}\right)$$

Let now $\mathcal{M} \in \mathsf{M}(\mathrm{A}\rangle\underline{\mathrm{M}})$, and let $\mathrm{I} = \mathbb{C}^d$ be the system for the $\log d$ bits of message and $\mathrm{J} = \mathbb{C}^d$ the system for the additional bit of message. The Holevo information after the measurement is the mutual information of the state $\xi \in \mathcal{D}(\mathrm{IJMB})$ given by:

$$\xi = \frac{1}{2d} \sum_{i\in\mathbb{Z}_d, j\in\mathbb{Z}_2} |ij\rangle\langle ij|_{\mathrm{IJ}} \otimes \left[\mathcal{M} \circ \mathcal{F}^j(|i\rangle\langle i|_{\mathrm{A}})\right]_{\mathrm{M}} \otimes |j\rangle\langle j|_{\mathrm{B}}$$

It is straightforward to check that tracing out Bob does not change the entropy, as his state is just a repetition of J, therefore $H(\mathrm{IJMB})_\xi = H(\mathrm{IJM})_\xi$. At the same time we have $H(\mathrm{MB})_\xi = H(\mathrm{M})_\xi + H(\mathrm{B})_\xi$, as tracing out the message IJ leaves AB maximally mixed. The result is that we can rewrite the mutual information as

$$\begin{aligned} I(\mathrm{IJ}{:}\mathrm{MB})_\xi &= H(\mathrm{IJ})_\xi + H(\mathrm{MB})_\xi - H(\mathrm{IJMB})_\xi \\ &= H(\mathrm{IJ})_\xi + H(\mathrm{M})_\xi + H(\mathrm{B})_\xi - H(\mathrm{IJM})_\xi \\ &= I(\mathrm{IJ}{:}\mathrm{M})_\xi + H(\mathrm{B})_\xi. \end{aligned}$$

Taking the supremum over measurements we thus have

$$\chi_{\mathsf{M}(\underline{\mathrm{A}})\otimes\mathrm{id}_{\mathrm{B}}}\left(\left\{\tfrac{1}{2d}, \psi_{ij}\right\}\right) = \chi_{\mathsf{M}(\underline{\mathrm{A}})}\left(\left\{\tfrac{1}{2d}, \mathcal{F}^j(|i\rangle\langle i|)\right\}\right) + H(\mathrm{B})_\xi = \frac{1}{2}\log(d) + 1. \qquad \square$$

2.3 Relative entropy

Let now $\rho, \sigma \in \mathcal{D}(\mathrm{A})$ and let supp denote the support; the relative entropy is defined as

$$D(\rho\|\sigma) := \begin{cases} \mathrm{tr}_{\mathrm{A}}[\rho\log\rho - \rho\log\sigma] & \operatorname{supp}\rho \subseteq \operatorname{supp}\sigma \\ +\infty & \text{otherwise} \end{cases}$$

where the logarithm of $K = \sum_{i\in\mathbb{Z}_{|\mathrm{H}|}} p_i\, |v_i\rangle\langle v_i| \in \mathcal{H}_+(\mathrm{H})$ is defined only on the non-zero eigenvalues:

$$\log(K) = \sum_{i:p_i>0} \log(p_i)\, |v_i\rangle\langle v_i|\,.$$

The relative entropy is non-negative, and zero if and only if $\rho = \sigma$. If ρ and σ commute, then it equals the Kullback–Leibler of the eigenvalues, namely

$$D(\rho\|\sigma) = \sum_{i:r_i\neq 0} r_i(\log r_i - \log s_i)$$

where r_i and s_i are the eigenvalues of any decomposition $\rho = \sum_{i\in\mathbb{Z}_{|\mathrm{A}|}} r_i\, |v_i\rangle\langle v_i|$ and $\sigma = \sum_{i\in\mathbb{Z}_{|\mathrm{A}|}} s_i\, |v_i\rangle\langle v_i|$. This is in particular the case of $D(\mathcal{M}(\rho)\|\mathcal{M}(\sigma))$ for a measurement $\mathcal{M}$, as the measurement always produces a state on the computational basis, and thus all the outcomes commute.

The relative entropy is monotonically decreasing under quantum channels, namely for any $\Lambda \in \mathsf{CPTP}(\mathrm{A}\rangle\mathrm{B})$ we have

$$D(\Lambda(\rho)\|\Lambda(\sigma)) \leqslant D(\rho\|\sigma),$$

which is, remarkably, a property equivalent to strong sub-additivity. Monotonicity then implies joint convexity, namely for any quantum states $\rho_i, \sigma_i \in \mathcal{D}(\mathrm{A})$ and probabilities p_i for i over some finite set I, we have

$$D\Big(\sum_{i\in I} p_i\rho_i \Big\| \sum_{i\in I} p_i\sigma_i\Big) \leqslant D\Big(\sum_{i\in I} p_i\,|i\rangle\langle i| \otimes \rho_i \Big\| \sum_{i\in I} p_i\,|i\rangle\langle i| \otimes \sigma_i\Big) = \sum_{i\in I} p_i D(\rho_i\|\sigma_i).$$

The relative entropy can be used to compute the mutual information via $I(\mathrm{A{:}B}) = D(\rho_{\mathrm{AB}}\|\rho_{\mathrm{A}} \otimes \rho_{\mathrm{B}})$, and it is additive on tensor products, namely $D(\rho_1 \otimes \rho_2\|\sigma_1 \otimes \sigma_2) = D(\rho_1\|\sigma_1) + D(\rho_2\|\sigma_2)$.

2.3.1 Restricted relative entropies

We can use the relative entropy expression of the mutual information to get a similar expansion for the Holevo information in terms of the relative entropy. With $\sigma := \sum_i p_i\sigma_i$ and ξ defined in Section 2.2.2 we have

$$\begin{aligned}\chi_{\mathsf{L}(\mathrm{H})}(\{p_i,\sigma_i\}) &= \sup_{\Lambda\in\mathsf{L}(\mathrm{H})} D(\mathrm{id}\otimes\Lambda(\xi)\|\pi\otimes\Lambda(\sigma))\\ &= \sup_{\Lambda\in\mathsf{L}(\mathrm{H})} \sum_i p_i D(\Lambda(\sigma_i)\|\Lambda(\sigma))\\ &\leqslant \sum_i \sup_{\Lambda\in\mathsf{L}(\mathrm{H})} p_i D(\Lambda(\sigma_i)\|\Lambda(\sigma)).\end{aligned}$$

By the monotonicity of the relative entropy, this relaxation is still lower than the unrestricted Holevo information, which is simply

$$\chi(\{p_i,\sigma_i\}) = \sum_i p_i D(\sigma_i\|\sigma).$$

We give the above as a mild justification for the definition we give below, but we will see later that the restricted relative entropy gives rise to various useful bounds and entanglement measures.

Definition 8 ([Pia09; 1]). *Let L be a subset of channels on H. For any states ρ and σ on H, the relative entropy in restriction to L, or L relative entropy, is defined as:*

$$D_{\mathsf{L}}(\rho\|\sigma) := \sup_{\Lambda\in\mathsf{L}} D(\Lambda(\rho)\|\Lambda(\sigma)). \tag{2.9}$$

Let now $\mathcal{K} \subseteq \mathcal{D}(\mathrm{H})$ be a subset of states. For any state $\rho \in \mathcal{D}(\mathrm{H})$, the relative entropy from $\mathcal{K}$ in restriction to L, or the L relative entropy from $\mathcal{K}$, is defined as:

$$D_{\mathsf{L}}(\rho\|\mathcal{K}) := \inf_{\sigma\in\mathcal{K}} D_{\mathsf{L}}(\rho\|\sigma).$$

The unrestricted relative entropy of ρ from $\mathcal{K}$ is defined as [DH99]:

$$D(\rho\|\mathcal{K}) := \inf_{\sigma\in\mathcal{K}} D(\rho\|\sigma).$$

Just like for the trace norm, D_{L} is still jointly convex and non-negative. If L is informationally complete, then $D_{\mathsf{L}}(\rho\|\sigma)$ if and only $\rho = \sigma$. If $\mathcal{K}$ is convex, then $D_{\mathsf{L}}(\rho\|\mathcal{K})$ and $D(\rho\|\mathcal{K})$ are also convex.

These definitions for restricted distinguishability, were introduced in [Pia09], but only using measurements. The reason for needing more than measurements is that for the relative entropy the analogue of Equation (2.2) does not hold: the relative entropy between two states ρ and σ in general is not achieved through the relative entropy of a global measurement outcome. We have by monotonicity

$$D_{\mathsf{M}(\underline{\mathrm{H}})}(\rho\|\sigma) \leqslant D(\rho\|\sigma), \tag{2.10}$$

but equality holds if and only if ρ and σ commute [BFT17]. By Equation (2.10) we still have $D_{\mathsf{L}(\underline{\mathrm{A}}:\underline{\mathrm{B}})}(\rho\|\mathcal{K}) \leqslant D_{\mathsf{L}(\underline{\mathrm{A}}:\mathrm{B})}(\rho\|\mathcal{K})$, but in Chapter 4 it will be possible to prove upper bounds only in terms of some $D_{\mathsf{L}(\underline{\mathrm{A}}:\mathrm{B})}(\rho\|\mathcal{K})$ and not in terms of $D_{\mathsf{L}(\underline{\mathrm{A}}:\underline{\mathrm{B}})}(\rho\|\mathcal{K})$, the bound with the latter being likely false.

The unrestricted relative entropy with respect to $\mathcal{K}$ is asymptotically continuous [DH99], with the following improved bound [Win16, Lemma 7]. For any convex set of states $\mathcal{K} \in \mathcal{D}(\mathrm{H})$ such that $\kappa = \sup_\rho D(\rho\|\mathcal{K}) < +\infty$, and for any two ε-close states ρ and $\tilde{\rho}$ on H, it holds:

$$|D(\rho\|\mathcal{K}) - D(\tilde{\rho}\|\mathcal{K})| \leqslant \varepsilon\kappa + g(\varepsilon). \tag{2.11}$$

Note that, if $\mathcal{K}$ contains the maximally mixed state $\mathbb{1}/|\mathrm{H}|$, then

$$D(\rho\|\mathcal{K}) \leqslant D(\rho\|\mathbb{1}/|\mathrm{H}|) \leqslant \log|\mathrm{H}|,$$

so that $\kappa \leqslant \log|\mathrm{H}|$. However, this upper bound is often not optimal in bipartite systems.

The asymptotic continuity for the restricted relative entropy was proven for measurements in [LW14]. Let $\mathsf{L} \subseteq \mathsf{M}(\underline{\mathrm{H}})$, and let $\mathcal{K} \subseteq \mathcal{D}(\mathrm{H})$ be star shaped around the maximally mixed state. Let $\varepsilon := \|\rho - \tilde{\rho}\|_{\mathsf{L}} \leqslant 1/e$ for any two states ρ and $\tilde{\rho}$ on H, then we have

$$|D_{\mathsf{L}}(\rho\|\mathcal{K}) - D_{\mathsf{L}}(\tilde{\rho}\|\mathcal{K})| \leqslant 4\varepsilon\log(3|\mathrm{H}|) + 4\eta(\varepsilon).$$

Notice how this asymptotic continuity is qualitatively different from the other ones. On one side the scaling factor depends on the input dimension, but on the other the distinguishability is the one at the output. Almost as if the optimal distinguishability where achieved for output dimensions comparable to the input dimension.

We conjecture that the restricted relative entropy is asymptotically continuous even for general subset of channels, as we will need such asymptotic continuity in Chapter 4.

Conjecture 9. *Let $\mathsf{L} \in \mathsf{CPTP}(\mathrm{H})$ be the set of quantum channels and let $\mathcal{K}$ be a closed convex set of states on H containing the maximally mixed state. Let $\varepsilon := \|\rho - \tilde{\rho}\|_{\mathsf{L}} \leqslant 1/e$ for two ε-close states ρ and $\tilde{\rho}$ on H, then it should hold*

$$|D_{\mathsf{L}}(\rho\|\mathcal{K}) - D_{\mathsf{L}}(\tilde{\rho}\|\mathcal{K})| = O(\varepsilon\log|\mathrm{H}|)$$

We will only need the conjecture for $\mathsf{LOCC}/\mathsf{SEP}$, and thus we can relax the conjecture to only classes of channels that also contain the channels with the corresponding records of the Kraus operators. We display a failed attempt of the proof in Section 4.5 and comment on its likeness.

2.3.2 Bipartite relative entropies

The typical first example example is the relative entropy of entanglement, namely the relative entropy from separable states [Ved+97]. It is common notation to denote the relative entropy of entanglement

$$E_R(\rho) := D(\rho\|\mathcal{S}(\mathrm{A}{:}\mathrm{B}))$$

which we will keep using when there is no ambiguity in the systems. The relative entropy of entanglement is upper bounded by $\min\{\log|\mathrm{A}|, \log|\mathrm{B}|\}$, achieved by the maximally

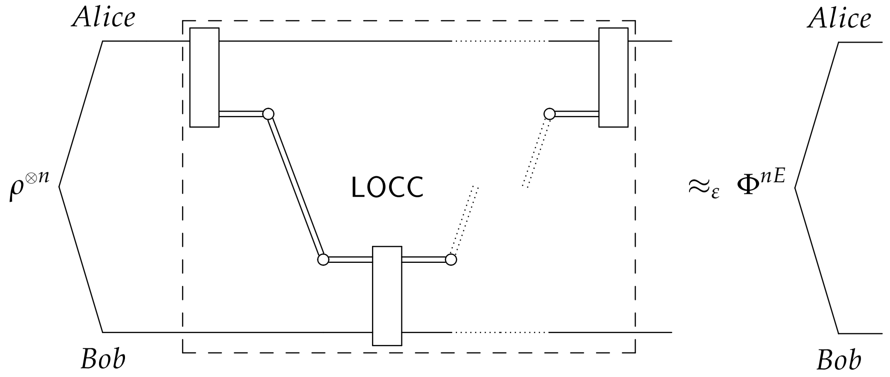


Figure 2.3: Quantum circuit for entanglement distillation. The circuit in the dashed box is the LOCC map Λ, optimized to get ε-close to Φ^{nE} with the highest nE possible. The double lines represent the classical communication going back and forth.

entangled state (for which the infimum is achieved by the maximally correlated state). The asymptotic continuity of the relative entropy of entanglement is thus

$$|E_R(\rho) - E_R(\tilde{\rho})| \leqslant \varepsilon \min\{\log|\mathrm{A}|, \log|\mathrm{B}|\} + g(\varepsilon). \tag{2.12}$$

for ε-close states ρ and $\tilde{\rho}$. The relative entropy with respect to PPT states yields the smaller measure

$$D(\rho\|\mathcal{P}(\mathrm{A{:}B})) \leqslant E_R(\rho)$$

which is monotone under PPT, while E_R is monotone only under SEP.

2.4 Entanglement Distillation

We have seen that entanglement, in particular the maximally entangled state, can be a powerful resource. However, it is also extremely susceptible to noise. Any attempt to produce real maximally entangled states will instead produce some mixed state, which can still be useful, but need to be best made use of in a non trivial way. More specifically, if the produced noisy state ρ is close enough to Φ, there exist protocols that can take two copies $\rho \otimes \rho$ (where each copy is shared between Alice and Bob) and produce a single copy $\tilde{\rho}$ closer to Φ than ρ was. If Alice and Bob want maximally entangled states, there are thus ways, given enough noisy entanglement, to get arbitrarily close to one. This is called entanglement distillation, sometimes called entanglement purification.

The scenario we consider is the one where Alice and Bob have a source that tries to produce the maximally entangled state Φ, but instead produces a mixed state ρ. Quantum information cannot be cloned and thus $\rho^{\otimes 2} \equiv \rho \otimes \rho$ cannot be produced from ρ alone. Here we will assume that Alice and Bob can run the source n times and get out the tensor product of n copies of ρ (an independent identically distributed source), which we denote with $\rho^{\otimes n}$, where each copy is shared between Alice and Bob. All the entanglement Alice and Bob can share is described by the source, which they can use arbitrarily many times as required, but otherwise they can only perform their bipartite operations.

In this section we review the known methods of quantifying and bounding entanglement distillation. Many concepts in this section are not exclusive to entanglement distillation, and thus we will try to highlight the general statements when possible.

2.4.1 Distillable Entanglement

To introduce the concept of distillation, consider for a moment the state

$$\frac{1}{2}\Phi^{+}_{\mathrm{AB}} \otimes |0\rangle\langle 0|_{\mathrm{A}'} + \frac{1}{2}\Phi^{-}_{\mathrm{AB}} \otimes |1\rangle\langle 1|_{\mathrm{A}'}$$

This is a mixed state with a maximally entangled state phase flipped with uniform probability. Usually the mixture would result in a separable state, but in this case Alice has a second register A', which is actually classical, with the record of the phase flip. Alice can thus read the record, use it to correct the phase, and then discard the record, resulting in Φ being perfectly shared between Alice and Bob. Thus maximally entangled states might not be available, but they can be extracted using LOCC (in this specific case even local operations). Recall that we defined the classes of operations as those classes containing local operations and being closed under tensor product and composition[1]. For any class of operations within PPT, this includes LO, $\mathsf{LOCC}_{\rightarrow}$, LOCC and SEP, then the amount of maximally entangled states that can be extracted directly from a state ρ is

$$E^{0}_{D,\mathsf{L}(\mathrm{A:B})}(\rho) := \sup\left\{\log e : \Phi^{\log e} \in \mathsf{L}(\mathbb{C}^{e}{:}\mathbb{C}^{e}\langle\mathrm{A:B}) \circ \rho\right\},$$

usually called zero-error single-copy rate. The base two logarithm results in this quantity being measured in units of qubit maximally entangled states Φ^{+}. The maximally entangled states are the target states of the distillation. Different distillation tasks under the same class of maps, like the key distillation we will see in the next chapter, can generally be described by simply changing the target states.

When noise is unavoidable, which is especially the case for quantum systems, it will be enough to just get close to the target state. If we then allow some error ε, the amount of entanglement that can be extracted becomes the single-copy rate

$$E^{\varepsilon}_{D,\mathsf{L}(\mathrm{A:B})}(\rho) := \sup\left\{\log e : \Phi^{\log e} \approx_{\varepsilon} \mathsf{L}(\mathbb{C}^{e}{:}\mathbb{C}^{e}\langle\mathrm{A:B}) \circ \rho\right\},$$

which now, for example, is non zero for the isotropic states with ε noise. If the protocols can take advantage of many copies, then their finite rate is defined by the amount of extracted target states —maximally entangled states in our case— averaged over the number of copies (Figure 2.3). For entanglement distillation, the best finite L entanglement rate for a given error ε and a given amount of copies n is thus

$$\frac{1}{n}E^{\varepsilon}_{D,\mathsf{L}(\mathrm{A:B})}(\rho^{\otimes n}).$$

A finite rate E is called achievable if for any choice of $\varepsilon > 0$ there exists a sequence of protocols on increasing number of copies, with finite rate limiting to E. Equivalently, a rate E is achievable if for all $\varepsilon > 0$

$$E \leqslant \limsup_{n\to\infty} \frac{1}{n}E^{\varepsilon}_{D,\mathsf{L}(\mathrm{A:B})}(\rho^{\otimes n}).$$

Since $E^{\varepsilon,n}_{D}(\rho)$ is monotonically decreasing in ε, the highest achievable rate is obtained taking the infimum over ε. The distillable entanglement under L is defined as the highest achievable L entanglement rate [Rai99b], namely:

$$E_{D,\mathsf{L}(\mathrm{A:B})}(\rho) := \lim_{\varepsilon\to 0}\limsup_{n\to\infty} \frac{1}{n}E^{\varepsilon}_{D,\mathsf{L}(\mathrm{A:B})}(\rho^{\otimes n}).$$

[1] Being close under composition is in particular important to make the rate monotone under the defining class of operations itself.

When clear from the context we will let the systems in the class of channels be implicit, and simply write $E_{D,\mathsf{L}}$. Because of the hierarchy of bipartite channels we have

$$E_{D,\mathsf{LO}} \leqslant E_{D,\mathsf{LOCC}_\rightarrow} \leqslant E_{D,\mathsf{LOCC}} \leqslant E_{D,\mathsf{SEP}} \leqslant E_{D,\mathsf{PPT}}.$$

If the set of channels is not specified, then it is assumed to be LOCC, so that we can write E_D for $E_{D,\mathsf{LOCC}}$, and $E_D^{\rightarrow}$ for $E_{D,\mathsf{LOCC}_\rightarrow}$.

By construction such rates are always additive on tensor products, namely $E_{D,\mathsf{L}}(\rho^{\otimes n}) = nE_{D,\mathsf{L}}(\rho)$; they are said to be regularized.

2.4.2 Known achievable rates

The distillable entanglement for pure states is known. Furthermore the process is reversible and can be done with one-way LOCC. For pure states $|\psi\rangle\langle\psi| \in \mathcal{D}(\mathrm{A{:}B})$, the distillable entanglement equals the entropy of entanglement [Ben+96a]:

$$E_{D,\mathsf{LOCC}_\rightarrow} = E_{D,\mathsf{PPT}} = H(\operatorname{tr}_{\mathrm{A}} |\psi\rangle\langle\psi|) = H(\operatorname{tr}_{\mathrm{B}} |\psi\rangle\langle\psi|).$$

More generally, it was shown in [DW05] that the coherent information is an achievable rate for one-way distillation. This is known as the hashing bound:

$$E_D^{\rightarrow}(\rho) \geqslant I(\mathrm{A}\rangle\mathrm{B})_\rho.$$

It is important to notice that, while it is valuable to prove lower bounds on the distillable entanglement (and is the only known computable bound), having zero coherent information gives no statement about the distillable entanglement of a state, because the coherent information is not monotone under local operations at Alice. Indeed, even for the maximally entangled state, the bound can be made zero by simply generating local randomness. Finally, the distillable entanglement is also known for states with support only on the maximally correlated subspace, as it is then proved that the hashing bound matches known upper bounds on the distillable entanglement [HH04], and is thus optimal.

2.4.3 Upper bounds

All known easily computable upper bounds on entanglement distillation are actually bounds on $E_{D,\mathsf{PPT}}$. The simplest of them is arguably the logarithmic negativity [Rai99a], defined as [VW02]

$$E_N(\rho) = \log \|\rho^\Gamma\|. \tag{2.13}$$

There are other computable bounds on $E_{D,\mathsf{PPT}}$ [Rai01; WD16], but we will only make use of the logarithmic negativity. In particular, all PPT states have zero distillable entanglement [HHH98]. The argument is simple and it generalizes to any tensor stable positive map [MRW16], which in this case is the transpose map. Namely, because $\rho^{\otimes n}$ is PPT if ρ is PPT, then any PPT protocol will necessarily output a PPT state. Formally

$$\mathsf{L}(\mathbb{C}^e{:}\mathbb{C}^e \langle \mathrm{A}^{\otimes n}{:}\mathrm{B}^{\otimes n}) \circ \rho^{\otimes n} \subseteq \mathcal{P}(\mathbb{C}^e{:}\mathbb{C}^e). \tag{2.14}$$

However, the distance of maximally entangled states from PPT states is large, namely $\left\|\Phi^{\log e} - \mathcal{P}(\mathbb{C}^e{:}\mathbb{C}^e)\right\| \geqslant 1 - \frac{1}{e}$, which is shown by simply twirling the set of PPT states into isotropic states.

In this respect, separable and PPT states behave similarly, in the sense that their distillable entanglement $E_{D,\mathsf{L}}$ is always zero, irrespective of the choice of $\mathsf{L} = \mathsf{LOCC}, \mathsf{SEP}, \mathsf{PPT}$. At the same time we expect E_D and $E_{D,PPT}$ to behave very differently for NPT states.

Indeed, one of the themes of this thesis is that intuitively, the entanglement distillation property of states such as

$$\gamma^{\pm}_{\mathrm{AA'BB'}} = \frac{1}{2}\big(\Phi^{+}_{\mathrm{AB}} \otimes \sigma^{+}_{\mathrm{A'B'}} + \Phi^{-}_{\mathrm{AB}} \otimes \sigma^{-}_{\mathrm{A'B'}}\big).$$

should roughly be determined by the distinguishability properties of the states $\sigma^{\pm}$. More precisely, we expect that if we choose $\sigma^{\pm}$ to be indistinguishable under LOCC, but distinguishable under PPT, then we should obtain states like the above with vanishing E_D but constant $E_{D,\mathsf{PPT}}$. While we are not able to prove this, almost all of our results seem to point in this direction.

Regularization

An important tool in defining upper bounds on distillation rates is the regularization of entanglement measures, which together with asymptotic continuity, is able [DHR02] to deal with the regularization already present in such rates, namely the limit for many copies. Let E be a function on states, like a general entanglement measure, the standard regularization defines $E^{\infty}(\rho) = \lim_{n\to\infty} \frac{1}{n} E(\rho^{\otimes n})$, whenever the limit is well defined. We will use it a bit more generally, and define regularizations also for functions E of two states, like the restricted relative entropies, as $E^{\infty}(\rho,\sigma) = \lim_{n\to\infty} \frac{1}{n} E(\rho^{\otimes n}, \sigma^{\otimes n})$, again whenever the limit is well defined. The usual tools to show that the limit is well defined, are super-additivity and sub-additivity. A sequence E_n is called super-additive if it satisfies for any n and m

$$E_{m+n} \geqslant E_n + E_m, \tag{2.15}$$

while it is called sub-additive if

$$E_{m+n} \leqslant E_n + E_m. \tag{2.16}$$

Fekete's lemmas [Fek23] state that super-additive and sub-additive sequences produce sequences $\frac{1}{n}E_n$ that are either convergent or divergent. More specifically, if E_n is super-additive, $\lim_{n\to\infty} \frac{1}{n}E_n = \sup_{n\in\mathbb{N}} \frac{1}{n}E_n$, while if it is sub-additive, then $\lim_{n\to\infty} \frac{1}{n}E_n = \inf_{n\in\mathbb{N}} \frac{1}{n}E_n$.

These lemmas guarantee that we can regularize $D_{\mathsf{L}}(\cdot\|\mathcal{K})$ as long as the combinations of L and $\mathcal{K}$ make it super- or sub-additive. The unrestricted relative entropy is always sub-additive for classes of states $\mathcal{K}$ closed under tensor product, and thus for such classes the regularized relative entropy from $\mathcal{K}$ is defined as

$$D^{\infty}(\rho\|\mathcal{K}(\mathrm{H}^{\otimes n})) := \lim_{n\to\infty} \frac{1}{n} D(\rho^{\otimes n}\|\mathcal{K}(\mathrm{H}^{\otimes n})) \leqslant D(\rho\|\mathcal{K}(\mathrm{H})).$$

For the relative entropy of entanglement we will keep the notation

$$E_R^{\infty}(\rho) := \lim_{n\to\infty} \frac{1}{n} E_R(\rho^{\otimes n}).$$

Similarly, if L is a class closed under tensor products, like LOCC, SEP and PPT, then the L relative entropy is easily checked to be super-additive. Thus the regularized L relative entropy is defined as

$$D_{\mathsf{L}}^{\infty}(\rho\|\sigma) := \lim_{n\to\infty} \frac{1}{n} D_{\mathsf{L}(\mathrm{H}^{\otimes n})}(\rho^{\otimes n}\|\sigma^{\otimes n}) \geqslant D_{\mathsf{L}(\mathrm{H})}(\rho\|\sigma). \tag{2.17}$$

To distinguish between measurements, partial measurements, different bipartitions, etc, we will use the input systems to denote the parties, and write for example

$$D^{\infty}_{\mathsf{SEP}(\underline{\mathrm{A}}:\underline{\mathrm{B}})} = \lim_{n\to\infty} \frac{1}{n} D_{\mathsf{L}\left(\underline{\mathrm{A}}^{\otimes n}:\underline{\mathrm{B}}^{\otimes n}\right)}(\rho^{\otimes n}\|\sigma^{\otimes n})$$

$$D^{\infty}_{\mathsf{SEP}(\underline{\mathrm{A}}:\mathrm{B})} = \lim_{n\to\infty} \frac{1}{n} D_{\mathsf{L}\left(\underline{\mathrm{A}}^{\otimes n}:\mathrm{B}^{\otimes n}\right)}(\rho^{\otimes n}\|\sigma^{\otimes n})$$

as done for the distillable entanglement.

The L relative entropy from $\mathcal{K}$ can also be regularized if it is super-additive. However, it is only known to be super-additive for measurements closed under tensor product, when the set of states is closed under its measurement operators [Pia09, Theorem 2(d)]. These properties are enough to guarantee that the L relative entropy with respect to $\mathcal{K}$ is also super-additive. While the original statement only considers separable states, PPT states, separable measurements and LOCC measurements, the proof is made in generality using only these properties. Thanks to this theorem, the following regularization is now well defined. Let L be a class of measurements, and let $\mathcal{K}$ be a convex class of states closed under measurement operators in L, then the regularized L relative entropy with respect to $\mathcal{K}$ is defined as:

$$D^{\infty}_{\mathsf{L}}(\rho\|\mathcal{K}) := \lim_{n\to\infty} \frac{1}{n} D_{\mathsf{L}\left(\underline{\mathrm{H}}^{\otimes n}\right)}(\rho^{\otimes n}\|\mathcal{K}(\mathrm{H}^{\otimes n})).$$

Therefore the regularization is well defined for the LOCC, SEP and PPT relative entropy with respect to PPT state, and for the LOCC and SEP relative entropy with respect to separable states.

Relative-entropy upper bounds

The regularized relative entropy of entanglement and from PPT states have been shown to be upper bounds on the PPT distillable entanglement [Hor+09]

$$E_{D,\mathsf{PPT}}(\rho) \leqslant D^{\infty}(\rho\|\mathcal{P}(\mathrm{A}:\mathrm{B})) \leqslant E_R^{\infty}(\rho)$$

thus showing

$$E_{D,\mathsf{PPT}}(\rho) \leqslant D(\rho\|\mathcal{P}(\mathrm{A}:\mathrm{B})) \leqslant E_R(\rho).$$

The regularized restricted relative entropies can also lead to upper bounds on distillation rates, however the normal restricted relative entropies generally are not themselves upper bounds. For L between LO, $\mathsf{LOCC}_{\rightarrow}$, LOCC, SEP or PPT

$$E_{D,\mathsf{L}} \leqslant D^{\infty}_{\mathsf{L}(\underline{\mathrm{A}}:\underline{\mathrm{B}})}(\rho\|\mathcal{S}(\mathrm{A}:\mathrm{B})).$$

Following the proof of asymptotic continuity for the measured relative entropy, the authors in [LW14] also proved the above bound for LOCC; an expanded proof for $\mathsf{LOCC}_{\rightarrow}$ can be found in [1]. While we do not have asymptotic continuity for partial measurements, since these include all measurements, we can prove regularized upper bounds for partial measurements by simple relaxation. Namely for any separable state $\sigma \in \mathcal{S}(\mathrm{A}:\mathrm{B})$ we have:

$$E_{D,\mathsf{L}} \leqslant D^{\infty}_{\mathsf{L}(\underline{\mathrm{A}}:\underline{\mathrm{B}})}(\rho\|\mathcal{S}(\mathrm{A}:\mathrm{B})) \leqslant D^{\infty}_{\mathsf{L}(\underline{\mathrm{A}}:\underline{\mathrm{B}})}(\rho\|\sigma) \leqslant D^{\infty}_{\mathsf{L}(\underline{\mathrm{A}}:\mathrm{B})}(\rho\|\sigma).$$

where we used that separable states are closed under tensor product.

In particular we will use the case of $\mathsf{LOCC}_{\rightarrow}$. By monotonicity of the relative entropy, the supremum in $D_{\mathsf{L}(\underline{\mathrm{A}}:\mathrm{B})}$ and its regularization is always achieved for Bob performing the identity channels in his step of the one-way protocol. By Equation (1.5), we can then remove the classical communication and reduce to just measurement on Alice. We thus have

$$E_{D,\mathsf{LOCC}_{\rightarrow}} \leqslant D^{\infty}_{\mathsf{LOCC}(\underline{\mathrm{A}}:\mathrm{B})}(\rho\|\sigma) = D^{\infty}_{\mathsf{M}(\underline{\mathrm{A}})}(\rho\|\sigma) \tag{2.18}$$

where we left implicit the identity channel in $\mathsf{M}(\underline{\mathrm{A}}) \equiv \mathsf{M}(\underline{\mathrm{A}}) \otimes \mathrm{id}_{\mathrm{B}}$.

2.5 Werner and Isotropic symmetries

We will now discuss two important twirls and the states they produce, namely the isotropic twirl [HH99] and the Werner twirl [Wer89], which have become canonical tools in quantum information. Let $U \in \mathcal{U}(\mathrm{A}) = \mathcal{U}(\mathrm{B})$, then the Werner twirl is defined as the twirl over all the unitaries $U \otimes U \in \mathcal{U}(\mathrm{AB})$, while the isotropic twirl is over all the unitaries $U \otimes \overline{U} \in \mathcal{U}(\mathrm{AB})$. Werner and isotropic states, and more generally Werner and isotropic operators, are defined as the invariants of the Werner and isotropic twirls. While in principle these twirls are defined continuously over the unitary group, they can be implemented [DLT02] by sampling over the Clifford subgroup $\mathcal{C}(\mathrm{A})$ [Got98], which is finite [Cal+98]. Notice that while the Werner twirl is basis independent, the isotropic twirl is basis dependent, as it involves complex conjugation. Therefore, the Werner states will be basis independent, while the isotropic states will be basis dependent (just like the Pauli twirl is basis-dependent, because the Bell states are basis dependent).

Both the Werner and isotropic twirl can be implemented using one-way LOCC. Furthermore, it is easily checked that

$$\left[(U \otimes U)K(U \otimes U)^\dagger\right]^\Gamma = (U \otimes \overline{U})K^\Gamma(U \otimes \overline{U})^\dagger.$$

Therefore the partial transpose maps the invariant subspace of the Werner twirl into the invariant subspace of the isotropic twirl.

Werner

The invariant operators for the Werner twirl are the subspace generated by the identity $\mathbb{1}_{\mathrm{AB}}$ and the swap S_{AB}. This is the same subspace generated by the projectors onto the symmetric and antisymmetric subspace. More precisely, let us define the projectors

$$\begin{aligned} P_s &:= \frac{\mathbb{1} + S}{2} = \sum_{i \geqslant j \in \mathbb{Z}_d} \frac{1}{2}(|ij\rangle + |ji\rangle)(h.c.) \\ P_a &:= \frac{\mathbb{1} - S}{2} = \sum_{i > j \in \mathbb{Z}_d} \frac{1}{2}(|ij\rangle - |ji\rangle)(h.c.), \end{aligned}$$

the dimensions $d_s := \operatorname{tr} P_s = d(d+1)/2$ and $d_a := \operatorname{tr} P_a = d(d-1)/2$, and the uniform mixtures

$$\begin{aligned} \rho_s &:= \frac{P_s}{\operatorname{tr} P_s} = \frac{2P_s}{d(d+1)} \\ \rho_a &:= \frac{P_a}{\operatorname{tr} P_a} = \frac{2P_a}{d(d-1)}. \end{aligned}$$

Any Werner invariant operator is a linear combination of P_s and P_a. The Werner twirl can then be rewritten as

$$\begin{aligned} \mathcal{W}_{\mathrm{AB}}(K) &:= \frac{1}{|\mathcal{C}(\mathrm{A})|} \sum_{U \in \mathcal{C}(\mathrm{A})} (U \otimes U)K(U \otimes U)^\dagger \\ &= \operatorname{tr}(P_s K)\rho_s + \operatorname{tr}(P_a K)\rho_a. \end{aligned}$$

Let $p \in [0,1]$, then a general Werner state has the form

$$\rho_{\mathrm{W}}(p) := p\rho_s + (1-p)\rho_a,$$

therefore ρ_s and ρ_a are also called the extremal Werner states. $\rho_{\mathrm{W}}(1 + \frac{1}{d})$ gives the maximally mixed state. By twirling the uniform mixture on $\mathbb{1} - P_{\hat{\Phi}}$ we obtain $\rho_{\mathrm{W}}(\frac{1}{2}) = \frac{1}{2}\rho_s + \frac{1}{2}\rho_a$

which lies on the boundary of the separable states. This is easily verified by checking that for any $p \geqslant \frac{1}{2}$ the Werner states are NPT. All PPT Werner states are thus separable.

The local norms for the extremal Werner states mentioned in Proposition 6 are easily computed to be

$$\|\rho_a - \rho_a\|_{\mathsf{LO}(\underline{\mathrm{A}}:\underline{\mathrm{B}})} = \|\rho_a - \rho_a\|_{\mathsf{LOCC}(\underline{\mathrm{A}}:\underline{\mathrm{B}})} = \|\rho_a - \rho_a\|_{\mathsf{SEP}(\underline{\mathrm{A}}:\underline{\mathrm{B}})} = \|\rho_a - \rho_a\|_{\mathsf{PPT}(\underline{\mathrm{A}}:\underline{\mathrm{B}})} = \frac{4}{d+1}.$$

For LO the most intuitive measurement to achieve this value is measurement one in the computational basis, while for LOCC and SEP the most straightforward choice is the measurement on the maximally correlated subspace with POVM $\{P_{\hat{\Phi}}, \mathbb{1} - P_{\hat{\Phi}}\}$. Twirling the measurement on the correlated subspace gives the extremal PPT Werner measurement that proves the upper bound, and the twirled POVM is given by $\left\{\frac{d}{d_s}P_s, P_a + \frac{d_a}{d_s}P_s\right\}$.

Isotropic

First we define the projector orthogonal to Φ and its uniform mixture:

$$\begin{aligned}\Phi_{\perp,\mathrm{AB}} &:= \mathbb{1}_{\mathrm{AB}} - \Phi_{\mathrm{AB}} \\ \phi_{\perp,\mathrm{AB}} &:= \frac{\Phi_{\perp,\mathrm{AB}}}{\operatorname{tr}\Phi_{\perp,\mathrm{AB}}}.\end{aligned}$$

The invariant operators for the isotropic twirl are the subspace generated by Φ and $\Phi_\perp$. The action of the isotropic twirl $\mathcal{I}$ on any operator $K \in \mathcal{H}_+(\mathrm{AB})$ is

$$\begin{aligned}\mathcal{I}(K) &:= \frac{1}{|\mathcal{C}(\mathrm{A})|}\sum_{U\in\mathcal{C}(\mathrm{A})}(U\otimes\overline{U})K(U\otimes\overline{U})^\dagger \\ &= (\operatorname{tr}K\Phi_{\mathrm{AB}})\Phi_{\mathrm{AB}} + (\operatorname{tr}K\Phi_{\perp,\mathrm{AB}})\phi_{\perp,\mathrm{AB}}.\end{aligned}$$

where $\operatorname{tr}(K\Phi)$ is known as the fidelity to the maximally entangled state. Let $p \in [0,1]$, then a general isotropic state of fidelity p has the form

$$\rho_{\mathrm{I}}(p) := p\Phi + (1-p)\Phi_\perp,$$

The maximally mixed state is obtained for $\rho_{\mathrm{I}}(\frac{1}{d^2})$, and the separable state on the boundary to the entangled states is $\rho_{\mathrm{I}}(\frac{1}{d})$, which is again verified by twirling the uniform mixture on $P_{\hat{\Phi}}$ and checking that otherwise the isotropic states are NPT for fidelities larger than $\frac{1}{d}$. Under partial transposition the separable Werner and isotropic states map intro each other, with the maximally mixed state as the only invariant state under partial transposition and the (separable) extremes mapping into each other as:

$$\phi_\perp \leftrightarrow \frac{1}{2}\rho_s + \frac{1}{2}\rho_a \qquad \text{and} \qquad \frac{1}{d}\Phi + \left(1 - \frac{1}{d}\right)\phi_\perp \leftrightarrow \rho_s.$$

Distinguishability. With the above parametrization the norm distance and the relative entropy between two isotropic states is quickly computed as

$$\begin{aligned}\|\rho_{\mathrm{I}}(p) - \rho_{\mathrm{I}}(q)\|_1 &= 2|p-q| \\ D(\rho_{\mathrm{I}}(p)\|\rho_{\mathrm{I}}(q)) &= \eta(p\|q) + \eta(1-p\|1-q)\end{aligned}$$

where $\eta(x\|y) := x(\log x - \log y)$. Namely, because the states in the mixture are orthogonal, the distances coincide with the classical norm and relative entropy of the input binary probability distributions given by p and q. If ρ is entangled ($p \geqslant 1/d$) then the minimal distance from the separable states and relative entropy of entanglement are

$$\begin{aligned} \|\rho_{\mathrm{I}}(p) - \mathcal{S}(\mathrm{A{:}B})\| &= 2\left(p - \frac{1}{d}\right) \\ E_R(\rho_{\mathrm{I}}(p)) = D(\rho_{\mathrm{I}}(p)\|\mathcal{S}(\mathrm{A{:}B})) &= \eta\left(p\middle\|\frac{1}{d}\right) + \eta\left(1-p\middle\|1-\frac{1}{d}\right) \end{aligned}$$

which are simply the values achieved for $q = 1/d$, the separable isotropic state on the boundary. In particular $E_R(\Phi) = \log d$. The isotropic twirl can always be used to reduce the output of an entanglement distillation protocol to an isotropic state, without changing the finite rates. More generally, if a state is ε-close to Φ then the isotropic twirl will always produce an isotropic state with fidelity larger than $1-\varepsilon$. Indeed, if σ is ε-close to Φ, then

$$\varepsilon > \frac{1}{2}\|\Phi - \sigma\|_1 \geqslant \frac{1}{2}\|\mathcal{I}(\Phi - \sigma)\|_1 = \frac{1}{2}\|\Phi - \mathcal{I}(\sigma)\|_1 = 1 - \mathrm{tr}(\Phi\sigma).$$

The proof normally used to compute the norm and relative entropy distances of isotropic states, has been generalized in [LW14] to compute the local restricted distances for the maximally entangled state. In the following lemma we generalize the proof further to arbitrary isotropic states. We will use this later in Section 3.2, to lower bound the distinguishability between two states after using the isotropic twirl as part of a distillation protocol.

Lemma 10. *Let* $|\mathrm{A}| = |\mathrm{B}|$. *For any isotropic states* $\rho = \rho_{\mathrm{I}}(p)$ *and* $\sigma = \rho_{\mathrm{I}}(q)$ *on* AB *we have*

$$\begin{aligned} \|\rho - \sigma\|_{\mathsf{LO}(\underline{\mathrm{A}}:\underline{\mathrm{B}})} = \|\rho - \sigma\|_{\mathsf{PPT}(\underline{\mathrm{A}}:\underline{\mathrm{B}})} &= 2\frac{d}{d+1}|p-q| \\ D_{\mathsf{LO}(\underline{\mathrm{A}}:\underline{\mathrm{B}})}(\rho\|\sigma) = D_{\mathsf{PPT}(\underline{\mathrm{A}}:\underline{\mathrm{B}})}(\rho\|\sigma) &= \frac{d}{d+1}\left[\eta\left(p + \frac{1}{d}\middle\|q + \frac{1}{d}\right) + \eta(1-p\|1-q)\right], \end{aligned}$$

and consequently if ρ *is entangled then*

$$\begin{aligned} \|\rho - \mathcal{S}(\mathrm{A{:}B})\|_{\mathsf{LO}(\underline{\mathrm{A}}:\underline{\mathrm{B}})} = \|\rho - \mathcal{S}(\mathrm{A{:}B})\|_{\mathsf{PPT}(\underline{\mathrm{A}}:\underline{\mathrm{B}})} &= 2\frac{d}{d+1}\left(p - \frac{1}{d}\right) \\ D_{\mathsf{LO}(\underline{\mathrm{A}}:\underline{\mathrm{B}})}(\rho\|\mathcal{S}(\mathrm{A{:}B})) = D_{\mathsf{PPT}(\underline{\mathrm{A}}:\underline{\mathrm{B}})}(\rho\|\mathcal{S}(\mathrm{A{:}B})) &= \frac{d}{d+1}\left[\eta\left(p+\frac{1}{d}\middle\|\frac{2}{d}\right) + \eta\left(1-p\middle\|1-\frac{1}{d}\right)\right]. \end{aligned}$$

Proof. The proof follows step by step the proof for the local relative entropy of Φ found in [LW14, Proposition 4]. First we estimate the lower bounds using the (local) measurement in the computational basis. For that purpose, let $\tau = \mathbb{1}/d^2$ be the maximally mixed state and consider the following parametrization of the isotropic states:

$$\begin{aligned} \rho &= p\Phi + (1-p)\Phi_\perp = a\Phi + (1-a)\tau \\ \sigma &= q\Phi + (1-q)\Phi_\perp = b\Phi + (1-b)\tau \end{aligned}$$

which gives

$$\begin{aligned} p &= a\frac{d^2-1}{d^2} + \frac{1}{d^2} \\ q &= b\frac{d^2-1}{d^2} + \frac{1}{d^2} \\ |p-q| &= \frac{d^2-1}{d^2}|a-b|. \end{aligned}$$

For any L norm we then have

$$\|\rho-\sigma\|_{\mathsf{L}} = |p-q|\cdot\|\Phi-\Phi_{\perp}\|_{\mathsf{L}} = |a-b|\cdot\|\Phi-\tau\|_{\mathsf{L}},$$

implying that the optimal measurement is independent of the isotropic states. In particular we obtain

$$\frac{d^2-1}{d^2}\|\Phi-\Phi_{\perp}\|_{\mathsf{L}} = \|\Phi-\tau\|_{\mathsf{L}}.$$

The measurement in the computational basis then gives

$$\|\Phi-\tau\|_{\mathsf{LO}(\underline{\mathrm{A}}:\underline{\mathrm{B}})} \geqslant 2\frac{d-1}{d}. \tag{2.19}$$

For the relative entropy the same measurement yields

$$\begin{aligned}
&D_{\mathsf{LO}(\underline{\mathrm{A}}:\underline{\mathrm{B}})}(\rho\|\sigma)\\
&\qquad\geqslant \sum_{ij}\eta\left(\frac{a}{d}\delta_{ij}+\frac{1-a}{d^2}\middle\|\frac{b}{d}\delta_{ij}+\frac{1-b}{d^2}\right)\\
&\qquad\geqslant \frac{d-1}{d}\Big(\eta\big(a+\tfrac{1}{d-1}\|b+\tfrac{1}{d-1}\big)+\eta(1-a\|1-b)\Big)\\
&\qquad= \frac{d}{d+1}\Big(\eta\big(p+\tfrac{1}{d}\|q+\tfrac{1}{d}\big)+\eta(1-p\|1-q)\Big)
\end{aligned} \tag{2.20}$$

Notice that in both cases, the outcome of the local measurement is the same as the outcome of the binary projective measurement on the maximally correlated subspace and the remaining orthogonal subspace.

To upper bound $\|\Phi-\tau\|_{\mathsf{PPT}}$ (and thus $\|\rho-\sigma\|_{\mathsf{PPT}}$), we use that any measurement acting on ρ and σ can be reduced to an isotropic measurement using $\operatorname{tr} M\rho_{\mathrm{I}}(p) = \operatorname{tr}\mathcal{I}(M)\rho_{\mathrm{I}}(p)$, where M is any positive-semidefinite operator [HH99]. If M is a PPT operator, then $\mathcal{I}(M)$ will be a PPT isotropic operator, all of which can be decomposed into a combination of the two extremal PPT isotropic operators $\Phi_{\perp}$ and $\Phi+\Phi_{\perp}/(d+1)$, see Figure 2.4. We can thus fine grain the measurement into operators proportional to the two extremal ones and then join them into an isotropic binary measurement. The same is true for the relative entropy using first joint convexity to fine grain the measurement into the extremal operators, and then using $\eta(ax\|ay)+\eta(bx,by) = \eta((a+b)x\|(a+b)y)$ to join them into a binary measurement. The result is, that we can restrict to binary measurements without loss of generality, and that the optimal measurement is the binary measurement with the two extremal PPT points as measurement operators.

Let $x,y\in[0,1]$, since a general isotropic measurement operator has the form $I^{x,y} = x\Phi+y\Phi_{\perp}$, see Figure 2.4, then a general isotropic binary measurement $\mathcal{M} = (I^{x,y},\mathbb{1}-I^{x,y})$ has dual operator of the form:

$$K_I^{x,y} := 2I^{x,y}-\mathbb{1} = (2x-1)\Phi+(2y-1)\Phi_{\perp}.$$

We thus find that:

$$\big\|\mathcal{M}(\Phi-\Phi_{\perp})\big\|_1 = \operatorname{tr}\big[K_I^{x,y}(\Phi-\Phi_{\perp})\big] = 2\cdot|x-y|. \tag{2.21}$$

The extremal separable operators are at $x=1$ and $y=\frac{1}{d+1}$, giving

$$\|\Phi-\Phi_{\perp}\|_{\mathsf{PPT}(\mathrm{A}:\mathrm{B})} \leqslant 2\frac{d}{d+1}$$

and matching the lower bound of Equation (2.19). This proves

$$\|\rho-\sigma\|_{\mathsf{LO}(\mathrm{A}:\mathrm{B})} = \|\rho-\sigma\|_{\mathsf{PPT}(\mathrm{A}:\mathrm{B})} = 2\frac{d}{d+1}|p-q|.$$

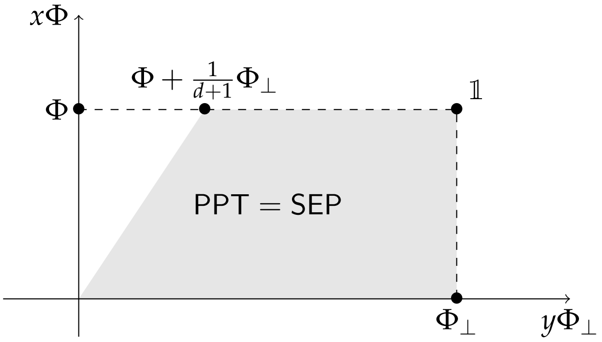


Figure 2.4: The space of isotropic operators. The rectangle is the space of measurement operators, the grey area are separable measurement operators. Notice that any separable operator can be written as positive linear combination of $\Phi_\perp$ and $\Phi + \frac{1}{d+1}\Phi_\perp$.

Similarly for the relative entropy we find that the extremal measurement achieves

$$\begin{aligned} &D_{\mathsf{PPT}(\underline{A}:\underline{B})}(\rho\|\sigma) \\ &\leqslant \eta\left(p + \frac{1-p}{d+1}\middle\| q + \frac{1-q}{d+1}\right) + \eta\left(\frac{(1-p)d}{d+1}\middle\|\frac{(1-q)d}{d+1}\right) \\ &= \frac{1}{d+1}\eta(pd+1\|qd+1) + \frac{d}{d+1}\eta(1-p\|1-q) \\ &= \frac{d}{d+1}\left(\eta\left(p+\frac{1}{d}\middle\| q+\frac{1}{d}\right) + \eta(1-p\|1-q)\right) \end{aligned}$$

matching the lower bound in Equation (2.20) and proving

$$\begin{aligned} D_{\mathsf{LO}(\underline{A}:\underline{B})}(\rho\|\sigma) &= D_{\mathsf{PPT}(\underline{A}:\underline{B})}(\rho\|\sigma) \\ &= \frac{d}{d+1}\left(\eta\left(p+\frac{1}{d}\middle\| q+\frac{1}{d}\right) + \eta(1-p\|1-q)\right). \end{aligned} \tag{2.22}$$

Let us omit (A:B) for the remainder of the proof. For the second part of the claim, we need to use the isotropic twirl to argue that it is enough to look at isotropic separable states to compute $\|\rho - \mathcal{S}\|_{\mathsf{LO}}$ and $D_{\mathsf{LO}}(\rho\|\mathcal{S})$. However, as a channel, the isotropic twirl needs shared randomness/communication and thus is not in LO. We need to use the convexity of the LO norm and the joint convexity of the LO relative entropy to de-randomize it. Namely, let $\varsigma \in \mathcal{S}$ be such that $\|\rho - \mathcal{S}\|_{\mathsf{LO}} = \|\rho - \varsigma\|_{\mathsf{LO}}$, then

$$\|\rho - \mathcal{I}(\varsigma)\|_{\mathsf{LO}} = \|\mathcal{I}(\rho) - \mathcal{I}(\varsigma)\|_{\mathsf{LO}} \leqslant \frac{1}{|\mathcal{C}(\mathrm{A})|}\sum_{U\in\mathcal{C}(\mathrm{A})}\left\|(U\otimes\overline{U})(\rho-\varsigma)(U\otimes\overline{U})^\dagger\right\|_{\mathsf{LO}}.$$

Since $U\otimes\overline{U}$ is a reversible local operation, we use $\|(U\otimes\overline{U})X(U\otimes\overline{U})^\dagger\|_{\mathsf{LO}} = \|X\|_{\mathsf{LO}}$ and get

$$\|\rho - \mathcal{I}(\varsigma)\|_{\mathsf{LO}} \leqslant \|\rho-\varsigma\|_{\mathsf{LO}} = \|\rho-\mathcal{S}\|_{\mathsf{LO}} \leqslant \|\rho-\mathcal{S}\|_{\mathsf{PPT}} \leqslant \|\rho - \rho_{\mathrm{I}}(\tfrac{1}{d})\|_{\mathsf{PPT}}.$$

Similarly, if $\varsigma\in\mathcal{S}$ is such that $D_{\mathsf{LO}}(\varrho\|\mathcal{S}) = D_{\mathsf{LO}}(\varrho\|\varsigma)$, then using joint convexity we get

$$D_{\mathsf{LO}}(\rho\|\mathcal{I}(\varsigma)) \leqslant D_{\mathsf{LO}}(\rho\|\varsigma) = D_{\mathsf{LO}}(\rho\|\mathcal{S}) \leqslant D_{\mathsf{PPT}}(\rho\|\mathcal{S}) \leqslant D_{\mathsf{PPT}}(\rho\|\rho_{\mathrm{I}}(\tfrac{1}{d})).$$

However it is straightforward to check that over the separable isotropic states, Equations (2.21) and (2.22) achieve the minimum for $\sigma = \rho_{\mathrm{I}}(\frac{1}{d})$ and therefore we find

$$\begin{aligned} \|\rho - \rho_{\mathrm{I}}(\tfrac{1}{d})\|_{\mathsf{LO}} &\leqslant \|\rho - \mathcal{I}(\varsigma)\|_{\mathsf{LO}} \\ D_{\mathsf{LO}}(\rho\|\rho_{\mathrm{I}}(\tfrac{1}{d})) &\leqslant D_{\mathsf{LO}}(\rho\|\mathcal{I}(\varsigma)) \end{aligned}$$

concluding the proof. □

Chapter 3

Private states

In this thesis, one of the main topics is the study of perfect keys, uniformly random strings shared between two parties but secret from anybody else, which are the targets of quantum key distribution. A quantum key distribution scheme generally needs to perform key distillation, the task of extracting almost perfect key, from noisy input states. If we fix the input state and ask what is the maximal amount of perfect key that can be extracted, like for entanglement distillation this defines a rate called the distillable key, which is another entanglement measure. Maximally entangled states contain perfect correlations that are pure and in product with the environment, and thus measuring them leads to perfect key. It turns out that there exist mixed states, the private states, that also lead to perfect key just by measuring. An important contribution to the understanding of quantum key distribution has been the discovery of private states from which secret key, but no maximally entangled states, can be extracted. The construction of those states was based on an intuition that the quantum mechanical phenomena of data hiding and privacy might be related. We have mentioned that there exist experimental realizations of these states, these can be found in [Dob+11].

In this chapter, after reviewing the basics of key distillation and private states, we prove an equivalent characterization of private states that reduces to classical the correlations between the source of the perfect key and the impurity of the privates states. This will allow us to derive a formal connection with quantum data hiding, confirming the intuition behind the construction of the states from which key but not pure entanglement can be extracted, that the separation between the distillable key and distillable entanglement is due to quantum data hiding. We will connect to this in the final chapter, where we will argue that this could be a purely bipartite property that does not survive the teleportation process.

3.1 Key Distillation

In key distillation we assume that the adversary is not an attacker (active), but he is passive, in which case we usually refer to as an eavesdropper, which we call Eve. In particular, Eve can only listen to the classical communication between two parties, but not tamper with it. With an attacker, the equivalent requirement is that the communication between Alice and Bob is authenticated. An eavesdropper also cannot actively attack Alice and Bob's local operations, and thus we assume that Alice and Bob can ensure that the environment of their local noisy operations does not leak to the outer environment and thus is not collected by Eve. The quantum part of the eavesdropping is the assumption that Eve can collect the information that leaks from the quantum communication to the environment, and thus can have an extension of the state shared by Alice and Bob. Since any extension can be derived from purifying environments, the worst case scenario is to assume that Eve holds the purification of the quantum communication. Since the purification also holds all the quantum information about the quantum communication that is theoretically possible, the worst case eavesdropping is also the worst case attack on the quantum communication.[1]

3.1.1 Local Operations and Public Communication

We will return to the quantum part of the eavesdropping in the next section. In this section we modify the class of LOCC operations to include the effect of the eavesdropping on the classical communication, modelling the information that is leaked to the environment in the process of distilling the key. Since the eavesdropper Eve can listen to any message, the worst case scenario is her collecting all messages. Equivalently we can think of the communication as holding a public record that anyone can read, including Eve; this is called public communication.

Public instruments and one-way LOPC

We thus define a public instrument as an instrument that outputs two copies of the measurement outcome. Namely, let C be the system for the message/communication and let $\mathrm{M} = \mathrm{C}$ be the system for the public record. Then we define public instruments as any channel $\Lambda \in \mathsf{CPTP}(\mathrm{H}_{\mathrm{in}}\rangle\mathrm{H}_{\mathrm{out}}\mathrm{CM})$ of the form

$$\Lambda = \sum_{i\in\mathbb{Z}_{|\mathrm{C}|}} \Lambda_i(\rho) \otimes |i\rangle\langle i|_{\mathrm{C}} \otimes |i\rangle\langle i|_{\mathrm{M}}$$

where $\Lambda_i \in \mathsf{CP}(\mathrm{H}_{\mathrm{in}}\rangle\mathrm{H}_{\mathrm{out}})$. We denote with $\mathsf{P.IN}(\mathrm{H}_{\mathrm{in}}\rangle\mathrm{H}_{\mathrm{out}}\underline{\mathrm{C}}|\underline{\mathrm{M}})$, or $\mathsf{P.IN}(\mathrm{H}_{\mathrm{out}}\underline{\mathrm{C}}|\underline{\mathrm{M}}\langle\mathrm{H}_{\mathrm{in}})$, the set of public instruments, using the conditioning "$|\mathrm{M}$" to highlight that when we compose further operations, we consider only one of CM as part of the available output, while the other one as an inaccessible third party. Later we will match this conditioning with the one in the conditional mutual information.

As for LOCC, in a one-way protocol all the instruments can be grouped together and the messages can be sent in a single round. One-way local operation with public communication (LOPC) from Alice to Bob are then defined as

$$\begin{aligned}&\mathsf{LOPC}_{\rightarrow}(\mathrm{A}_{\mathrm{out}}{:}\mathrm{B}_{\mathrm{out}}|\underline{\mathrm{M}}\langle\mathrm{A}_{\mathrm{in}}{:}\mathrm{B}_{\mathrm{in}})\\ &\qquad := [\mathrm{id}_{\mathrm{A}_{\mathrm{out}}\mathrm{M}} \otimes \mathsf{CPTP}(\mathrm{B}_{\mathrm{out}}\langle\mathrm{CB}_{\mathrm{in}})] \circ [\mathsf{P.IN}(\mathrm{A}_{\mathrm{out}}\underline{\mathrm{C}}|\underline{\mathrm{M}}\langle\mathrm{A}_{\mathrm{in}}) \otimes \mathrm{id}_{\mathrm{B}_{\mathrm{in}}}],\end{aligned}$$

while from Bob to Alice as

$$\begin{aligned}&\mathsf{LOPC}_{\leftarrow}(\mathrm{A}_{\mathrm{out}}{:}\mathrm{B}_{\mathrm{out}}|\underline{\mathrm{M}}\langle\mathrm{A}_{\mathrm{in}}{:}\mathrm{B}_{\mathrm{in}})\\ &\qquad := [\mathsf{CPTP}(\mathrm{A}_{\mathrm{out}}\langle\mathrm{A}_{\mathrm{in}}\mathrm{C}) \otimes \mathrm{id}_{\mathrm{B}_{\mathrm{out}}\mathrm{M}}] \circ [\mathrm{id}_{\mathrm{A}_{\mathrm{in}}} \otimes \mathsf{P.IN}(\underline{\mathrm{C}}\mathrm{B}_{\mathrm{out}}|\underline{\mathrm{M}}\langle\mathrm{B}_{\mathrm{in}})].\end{aligned}$$

Two-way LOPC

Two-way LOPC is also defined similarly to LOCC. However, because fixing a global system M for the communication in the protocols does not tell us the systems of each communication round, we need to keep track of the system for each round individually.

[1] There can still be a big difference between quantum eavesdroppers and quantum attackers with regard to the local operations. Current devices might be secure against quantum eavesdroppers, but we can safely say that Vadim Makarov is not one.

First we define a single round of two-way LOPC as a function of the input/output systems for Alice and Bob, and a list of two systems for the public communication:

$$\begin{aligned}&\mathsf{LOPC}^1(\mathrm{A}_{\mathrm{out}}{:}\mathrm{B}_{\mathrm{out}}|\,\{\underline{\mathrm{M}}_{\rightarrow},\underline{\mathrm{M}}_{\leftarrow}\}\langle\mathrm{A}_{\mathrm{in}}{:}\mathrm{B}_{\mathrm{in}})\\&\quad:=\bigcup_{\mathrm{A,B}}[\mathsf{LOPC}_{\leftarrow}(\mathrm{A}_{\mathrm{out}}{:}\mathrm{B}_{\mathrm{out}}|\underline{\mathrm{M}}_{\leftarrow}\langle\mathrm{A{:}B})\otimes\mathrm{id}_{\mathrm{M}_{\rightarrow}}]\circ\mathsf{LOPC}_{\rightarrow}(\mathrm{A{:}B}|\underline{\mathrm{M}}_{\rightarrow}\langle\mathrm{A}_{\mathrm{in}}{:}\mathrm{B}_{\mathrm{in}}).\end{aligned}$$

We then group the messages together in a single system as

$$\mathsf{LOPC}^1(\mathrm{A}_{\mathrm{out}}{:}\mathrm{B}_{\mathrm{out}}|\underline{\mathrm{M}}\langle\mathrm{A}_{\mathrm{in}}{:}\mathrm{B}_{\mathrm{in}}):=\bigcup_{\mathrm{M}_{\rightarrow}\mathrm{M}_{\leftarrow}=\mathrm{M}}\mathsf{LOPC}^1(\mathrm{A}_{\mathrm{out}}{:}\mathrm{B}_{\mathrm{out}}|\,\{\underline{\mathrm{M}}_{\rightarrow},\underline{\mathrm{M}}_{\leftarrow}\}\langle\mathrm{A}_{\mathrm{in}}{:}\mathrm{B}_{\mathrm{in}})$$

The first definition contains in principle more information, but this second definition will be enough for our purposes, as $\mathsf{LOPC}^1(\mathrm{A}_{\mathrm{in}}{:}\mathrm{B}_{\mathrm{in}}\rangle\mathrm{A}_{\mathrm{out}}{:}\mathrm{B}_{\mathrm{out}}|\underline{\mathrm{M}})$ is used in a way that is independent of the public communication. This will be enough to define key distillation, as the goal of Alice and Bob is to distill key independently of any eavesdropper, and thus independently of the communication. It should be enough also when upper bounding the power of the eavesdropper, as this will involve optimizing over all of his operations, which include the ones that perfectly match the global public system to the communication in each round.

Since the public system is never touched by Alice and Bob, from now on we will omit the identity maps on the public systems. We now define the n-rounds of two-way LOPC recursively. Similarly we first define the n-rounds explicitly with each message, and then define them with a global public system. Now the list of two systems becomes a list of $2n$-systems

$$\begin{aligned}&\mathsf{LOPC}^n(\mathrm{A}_{\mathrm{out}}{:}\mathrm{B}_{\mathrm{out}}|\,\{\underline{\mathrm{M}}^m_{\rightarrow},\underline{\mathrm{M}}^m_{\leftarrow}\}_{m=1}^n\langle\mathrm{A}_{\mathrm{in}}{:}\mathrm{B}_{\mathrm{in}})\\&\quad:=\bigcup_{\mathrm{A,B}}\mathsf{LOPC}^1(\mathrm{A}_{\mathrm{out}}{:}\mathrm{B}_{\mathrm{out}}|\,\{\underline{\mathrm{M}}^n_{\rightarrow},\underline{\mathrm{M}}^n_{\leftarrow}\}\langle\mathrm{A{:}B})\circ\mathsf{LOPC}^{n-1}(\mathrm{A{:}B}|\,\{\underline{\mathrm{M}}^m_{\rightarrow},\underline{\mathrm{M}}^m_{\leftarrow}\}_{m=1}^{n-1}\langle\mathrm{A}_{\mathrm{in}}{:}\mathrm{B}_{\mathrm{in}})\end{aligned}$$

where as mentioned we omitted the identity map on the public systems of the previous $n-1$ rounds. Then to group all the public communication together we define:

$$\mathsf{LOPC}^n(\mathrm{A}_{\mathrm{in}}{:}\mathrm{B}_{\mathrm{in}}\rangle\mathrm{A}_{\mathrm{out}}{:}\mathrm{B}_{\mathrm{out}}|\underline{\mathrm{M}}):=\bigcup\mathsf{LOPC}^n(\mathrm{A}_{\mathrm{in}}{:}_{\mathrm{in}}\rangle\mathrm{A}_{\mathrm{out}}{:}\mathrm{B}_{\mathrm{out}}|\,\{\underline{\mathrm{M}}^m_{\rightarrow},\underline{\mathrm{M}}^m_{\leftarrow}\}_{m=1}^n)$$

where the union is over systems $\{\mathrm{M}^m_{\rightarrow},\mathrm{M}^m_{\leftarrow}\}_{m=1}^n$ such that $\mathrm{M}^1_{\rightarrow}\mathrm{M}^1_{\leftarrow}\ldots\mathrm{M}^{n-1}_{\rightarrow}\mathrm{M}^{n-1}_{\leftarrow}=\mathrm{M}$.

Finally, we allow arbitrary rounds of communication and define LOPC as

$$\mathsf{LOPC}(\mathrm{A}_{\mathrm{in}}{:}\mathrm{B}_{\mathrm{in}}\rangle\mathrm{A}_{\mathrm{out}}{:}\mathrm{B}_{\mathrm{out}}|\underline{\mathrm{M}}):=\bigcup_{n\in\mathbb{N}}\mathsf{LOPC}^n(\mathrm{A}_{\mathrm{in}}{:}\mathrm{B}_{\mathrm{in}}\rangle\mathrm{A}_{\mathrm{out}}{:}\mathrm{B}_{\mathrm{out}}|\underline{\mathrm{M}}),$$

We then have all the LOPC channels defined by

$$\mathsf{LOPC}(\mathrm{A{:}B}):=\bigcup_{\mathrm{A}_{\mathrm{out}},\mathrm{B}_{\mathrm{out}},\mathrm{M}}\mathsf{LOPC}(\mathrm{A{:}B}\rangle\mathrm{A}_{\mathrm{out}}{:}\mathrm{B}_{\mathrm{out}}|\underline{\mathrm{M}}).$$

LOPC **measurements**

The additional public systems do not interfere with measurements. Measurements and partial measurements are thus simply defined taking the intersection with measurements or instruments in the relevant systems as done in Section 1.4. Namely we have that the LOPC measurements are $\mathsf{LOPC}(\mathrm{A{:}B}\rangle\underline{\mathrm{A}}_{\mathrm{out}}{:}\underline{\mathrm{B}}_{\mathrm{out}}|\underline{\mathrm{M}})$ and $\mathsf{LOPC}(\underline{\mathrm{A}}{:}\underline{\mathrm{B}})$ for two way LOPC, and similarly for one-way LOPC. Analogously, LOPC partial measurements are given by $\mathsf{LOPC}(\mathrm{A{:}B}\rangle\underline{\mathrm{A}}_{\mathrm{out}}{:}\mathrm{B}_{\mathrm{out}}|\underline{\mathrm{M}})$ and $\mathsf{LOPC}(\underline{\mathrm{A}}{:}\mathrm{B})$, and similarly for Bob and for one-way LOPC.

Monotonicity

We have already mentioned that strong sub-additivity implies the monotonicity of the mutual and conditional mutual information under local operations. It is easy to check that these quantities are not monotone under classical communication, exactly because it can create correlations. Making the communication public makes these quantities monotone. More specifically, through the conditioning inequality

$$I(\mathrm{A{:}BC}|\mathrm{E}) \leqslant I(\mathrm{A{:}B}|\mathrm{CE}),$$

strong sub-additivity also implies monotonicity under conditioning on the outcomes of the public instruments, because $I(\mathrm{A{:}B}|\mathrm{CE})$ is exactly the effect of sending the public message, as we will see below. The result is that the conditional mutual information is monotone under LOPC. Note that the same proof applies to the two known multipartite generalizations of the conditional mutual information [Yan+09], namely they will be monotone under the multipartite version of LOPC (however we have no use for these generalizations here). The reason is that the conditioning system is present in every entropy term, thus allowing us to use strong sub-additivity after removing the public message from the parties (which does not change the conditional mutual information).

This monotonicity opens the possibility to define a measured conditional mutual information for states ρ in ABE as

$$I_{\mathsf{LOPC}(\underline{\mathrm{A}}:\underline{\mathrm{B}})}(\rho) := \sup_{\Lambda \in \mathsf{LOPC}(\mathrm{A{:}B}\rangle \mathrm{A_{out}{:}B_{out}}|\underline{\mathrm{M}}) \otimes \mathrm{id_E}} I(\mathrm{A_{out}{:}B_{out}}|\mathrm{ME})_{\Lambda(\rho)}.$$

where the supremum is also over systems $\mathrm{A_{out}}$, $\mathrm{B_{out}}$, M. The same holds for partial measurements $\mathsf{LOPC}(\underline{\mathrm{A}}{:}\mathrm{B})$. We will see later how to use this to construct an entanglement measure in Section 3.1.4, which is of independent interest. The next lemma contains the proof of the claimed monotonicity.

Lemma 11. *The mutual and conditional mutual information are* LOPC *monotones.*
More precisely, let $\rho \in \mathcal{D}(\mathrm{A_{in}B_{in}E})$ *and* $\Lambda \in \mathsf{LOPC}(\mathrm{A_{in}{:}B_{in}}\rangle \mathrm{A_{out}{:}B_{out}}|\underline{\mathrm{M}}) \otimes \mathrm{id_E}$*, then*

$$I(\mathrm{A_{in}{:}B_{in}}|\mathrm{E})_\rho \geqslant I(\mathrm{A_{out}{:}B_{out}}|\mathrm{ME})_{\Lambda(\rho)}.$$

Proof. The result for the mutual information follows from the conditional mutual information one, by making Eve's system trivial. Therefore we only need to show the latter. We already know that the conditional mutual information is monotone under local operations. We thus only need to show monotonicity under the public communication. Let ρ be a state on $\mathrm{A_{in}BE}$ and $\Lambda \in \mathsf{P.IN}(\mathrm{A_{in}}\rangle \mathrm{AC}|\underline{\mathrm{M}}) \otimes \mathrm{id_{BE}}$, so that $\Lambda(\rho)$ is a state on ACMBE. It is easily verified that we can trace out C if M is also present in an entropy, and thus

$$\begin{aligned} H(\mathrm{ACMBE}) &= H(\mathrm{ABME}), \\ H(\mathrm{ACME}) &= H(\mathrm{AME}), \\ H(\mathrm{CBME}) &= H(\mathrm{BME}). \end{aligned}$$

where the conditions hold for any $\rho \in \mathcal{D}(\mathrm{A_{in}BE})$. We thus have by strong sub-additivity

$$I(\mathrm{ACM{:}B}|\mathrm{E}) = I(\mathrm{AM{:}B}|\mathrm{E}) \leqslant I(\mathrm{A{:}B}|\mathrm{ME}) = I(\mathrm{A{:}CB}|\mathrm{ME})$$

Combining with monotonicity under local operations, this gives that the conditional mutual information is monotone under $\mathsf{LOPC}_{\rightarrow}$. A mirroring argument gives monotonicity also under $\mathsf{LOPC}_{\leftarrow}$, and thus LOPC. □

3.1.2 Distillable Key

We have seen how to model the information leaked to the eavesdropper (the passive adversary) in the process of distilling the key. We also mentioned that the quantum part of the eavesdropping is, in the worst case scenario, Eve holding the purification of the quantum communication. We can assume that Alice and Bob have a source that they can run multiple times to produce a tensor product state $\rho^{\otimes n}$ like for entanglement distillation, but it requires some justification.

In key distillation, there are in general no tensor product sources. Even if an underlying pure source could be considered in tensor product (for example the photons generated locally before being sent through a fibre), assuming that the noise is also in tensor product puts assumptions on the power of the eavesdropper. Fortunately, it was proven [CKR09; Ren07] that the security of a fully general noise model, can be reduced to the security of a tensor product source without changing the achievable rates (the finite rates might differ though, as the reduction needs an active operation from Alice and Bob). We will thus assume that Alice and Bob's source produces states $\rho^{\otimes n}$.

The goal of key distillation is to extract key, namely classical bits that are perfectly correlated between Alice and Bob, and independent from any eavesdropper. If E is the system of the eavesdropper, then the states of perfect key have the form

$$\hat{\Phi}_{\mathrm{A^{out}B^{out}}} \otimes \sigma$$

for some state $\sigma \in \mathcal{D}(\mathrm{E})$ and where the maximally correlated state $\hat{\Phi}$ was defined in Section 1.5. This class of states is closed under tensor product and, like for the distillable entanglement, the distillable key will be a measure of key in units of $\hat{\Phi}^{+} \otimes \sigma$. To verify that the distillation procedure achieves this state, we need to keep track of the system of Eve, including the public communication, from the beginning of the protocol, where we assume that Eve holds the purification, to the end. Let $\rho \in \mathcal{D}(\mathrm{AB})$ be the mixed state produced by the source, and denote by $|\rho\rangle\langle\rho| \in \mathcal{D}(\mathrm{ABE})$ a purification. Then the single-shot, zero-error key rate of ρ is given by

$$K_D^0(\rho) := \sup\left\{\log\kappa : \left[\hat{\Phi}^{\log\kappa} \otimes \mathcal{D}(\mathrm{ME})\right] \cap \left[\mathsf{LOPC}(\mathbb{C}^{\kappa}{:}\mathbb{C}^{\kappa}|\underline{\mathrm{M}}\langle\mathrm{A{:}B}) \otimes \mathrm{id}_{\mathrm{E}} \circ |\rho\rangle\langle\rho|\right] \neq \varnothing\right\}$$

where the supremum is also over the public system M. The maximally entangled state Φ^{m} clearly has single-shot zero-error key rate larger than m, as it is already a pure state and thus the purification will be in product. The measurement in the computational basis then achieves the perfect key. More generally, distilling the maximally entangled state implies that we can also distill key with the same size, simply by performing the measurement after the distillation.

We now allow for a small error, as in the case for distillable entanglement, and define the best finite key rate of ρ as

$$K_D^{\varepsilon}(\rho) := \sup\left\{\log\kappa : \hat{\Phi}^{\log\kappa} \otimes \mathcal{D}(\mathrm{ME}) \approx_{\varepsilon} \mathsf{LOPC}(\mathbb{C}^{\kappa}{:}\mathbb{C}^{\kappa}\langle\mathrm{A{:}B}|\underline{\mathrm{M}}) \otimes \mathrm{id}_{\mathrm{E}} \circ |\rho\rangle\langle\rho|\right\}.$$

The distillable key is defined as the best achievable rate for $K_D^{\varepsilon}(\rho)$:

$$K_D(\rho) := \lim_{\varepsilon\to 0} \limsup_{n\to\infty} \frac{1}{n} K_D^{\varepsilon}(\rho^{\otimes n}).$$

We then have $E_D(\rho) \leqslant K_D(\rho)$. The same definition applies for one-way LOPC protocols, defining the one-way distillable key $K_D^{\rightarrow}(\rho)$, also satisfying $E_D^{\rightarrow}(\rho) \leqslant K_D^{\rightarrow}(\rho)$.

The following achievable rate for the one way distillable key was proven together with the hashing bound, and is known as the Devetak-Winter rate [DW05]:

$$\begin{aligned} K_D^{\rightarrow}(\rho) &\geqslant \sup_{\mathcal{M}\in\mathrm{M}(\mathrm{A}\rangle\mathrm{M})} [I(\mathrm{M{:}B}) - I(\mathrm{M{:}E})]_{\mathcal{M}\otimes\mathrm{id}_{\mathrm{BE}}(|\rho\rangle\langle\rho|)} \\ &= \sup_{\mathcal{M}\in\mathrm{M}(\mathrm{A}\rangle\mathrm{M})} [I(\mathrm{M}\rangle\mathrm{B}) - I(\mathrm{M}\rangle\mathrm{E})]_{\mathcal{M}\otimes\mathrm{id}_{\mathrm{BE}}(|\rho\rangle\langle\rho|)}. \end{aligned}$$

3.1.3 Private states

We have already mentioned that the maximally entangled state Φ^+ contains a perfect bit of key, which is extracted simply with a measurement in the computational basis. Private states are those states that generalize this property, namely they are all those states that produce perfect key when measured in the computational basis. More precisely, let Alice and Bob now hold two pairs of systems:

- The key systems $\mathrm{A}_{\text{🔑}}\mathrm{B}_{\text{🔑}}$, of equal dimension $|\mathrm{A}_{\text{🔑}}| = |\mathrm{B}_{\text{🔑}}|$, where we want to perform the measurement in the computational basis and get the perfect key (as indicated by the key symbol 🔑).
- The shield systems $\mathrm{A}_{\text{🛡}}\mathrm{B}_{\text{🛡}}$ of arbitrary dimension for the generalization. The perfect key in $\mathrm{A}_{\text{🔑}}\mathrm{B}_{\text{🔑}}$ might be correlated with $\mathrm{A}_{\text{🛡}}\mathrm{B}_{\text{🛡}}$ but it will be secret from any purifying environment, (as indicated by the shield symbol 🛡).

If the shield systems are trivial ($\mathrm{A}_{\text{🛡}} = \mathrm{B}_{\text{🛡}} = \mathbb{C}$), then the only states that produce perfect key just by measuring are the maximally entangled states. Aside from Φ though, we have already seen an example of a private state in Section 2.4, when we explained that the state

$$\frac{1}{2}\Phi^+_{\mathrm{AB}} \otimes |0\rangle\langle 0|_{\mathrm{A}'} + \frac{1}{2}\Phi^-_{\mathrm{AB}} \otimes |1\rangle\langle 1|_{\mathrm{A}'} \tag{3.1}$$

is perfectly distillable. Of course, we can first distill the entanglement by correcting the phase and then extract the key. However, it is straightforward to verify that this state produces perfect key simply by measuring the maximally entangled state, without applying any correction. This will work in full generality also if, instead of the classical local states $|0\rangle\langle 0|$ and $|1\rangle\langle 1|$, this information is stored in any pair of bipartite orthogonal states σ_0 and σ_1. Recalling the terminology of Section 2.2 for the Holevo information, we would say that measuring directly the maximally entangled states will produce perfect key if the information of the phase is perfectly encoded in any bipartite ensemble. The the perfect encoders $|i\rangle\langle i|$ are the shield of this example, which being local allow for perfect distillation of entanglement. The interesting private states with low distillable entanglement emerge for some perfect encoders that are data hiding states, and thus do not allow for the correction of the phase flip by accessing the information on the shield locally. However, encoding the phase information in data hiding states does not guarantee reduced distillable entanglement. The purpose of this chapter is to make the connection with data hiding formal and provide interesting edge cases, but before that we need to introduce private states in full generality.

Given any state $\rho \in \mathcal{D}(\mathrm{A}_{\text{🔑}}\mathrm{B}_{\text{🔑}}\mathrm{A}_{\text{🛡}}\mathrm{B}_{\text{🛡}})$, the outcome of measuring the computational basis of $\mathrm{A}_{\text{🔑}}\mathrm{B}_{\text{🔑}}$ is called the key-attacked state of ρ, and we will denote it with $\hat{\rho}$. This is compatible with the notation for the maximally correlated state $\hat{\Phi}$, which is indeed the key-attacked state of the maximally entangled state Φ. It was proven in [Hor+05b; Hor+09] that $\gamma \in \mathcal{D}(\mathrm{A}_{\text{🔑}}\mathrm{B}_{\text{🔑}}\mathrm{A}_{\text{🛡}}\mathrm{B}_{\text{🛡}})$ is a private state if and only if:

$$\gamma = \mathcal{T}(\Phi_{\mathrm{A}_{\text{🔑}}\mathrm{B}_{\text{🔑}}} \otimes \sigma) \tag{3.2}$$

for some shield state $\sigma \in \mathcal{D}(\mathrm{A}_{🛡}\mathrm{B}_{🛡})$ and some controlled unitary channel $\mathcal{T}$. Notice that because Φ is the maximally entangled state, only the part of $\mathcal{T}$ that controls on $|ii\rangle\langle ii|$ has an effect of the definition of γ. We will see below where does the rest of the controlled unitary play a role. In the example of Equation (3.1), the shield state is the maximally mixed state of one qubit, and the controlled unitary is a controlled phase, controlled on either of the key qubit at Alice or Bob.

To be more precise, for any control system H and target system T let $\mathsf{CU}(\mathrm{H{:}T}) \subset \mathsf{CPTP}(\mathrm{HT}\rangle\mathrm{HT})$ denote the subset of controlled unitary channels, which is a subgroup of the unitary channels. Then for any pure states $\gamma \in \mathcal{D}(\mathrm{A}_{🔑}\mathrm{B}_{🔑}\mathrm{A}_{🛡}\mathrm{B}_{🛡}\mathrm{E})$ and $\sigma \in \mathcal{D}(\mathrm{A}_{🛡}\mathrm{B}_{🛡}\mathrm{E})$ the result from [Hor+05b] says that $\hat{\gamma}_{\mathrm{A}_{🔑}\mathrm{B}_{🔑}\mathrm{E}}$ is a state of perfect key if and only if $\gamma_{\mathrm{A}_{🔑}\mathrm{B}_{🔑}\mathrm{A}_{🛡}\mathrm{B}_{🛡}}$ is of the form of Equation (3.2), or equivalently

$$\hat{\gamma}_{\mathrm{A}_{🔑}\mathrm{B}_{🔑}\mathrm{E}} = \hat{\Phi}_{\mathrm{A}_{🔑}\mathrm{B}_{🔑}} \otimes \sigma_{\mathrm{E}} \quad \Leftrightarrow \quad \gamma_{\mathrm{A}_{🔑}\mathrm{B}_{🔑}\mathrm{A}_{🛡}\mathrm{B}_{🛡}} \in \mathsf{CU}(\mathrm{A}_{🔑}\mathrm{B}_{🔑}, \mathrm{A}_{🛡}\mathrm{B}_{🛡}) \circ (\Phi_{\mathrm{A}_{🔑}\mathrm{B}_{🔑}} \otimes \sigma_{\mathrm{A}_{🛡}\mathrm{B}_{🛡}})$$

In particular, the controlled unitaries are exactly the invariant operations for $\hat{\gamma}_{\mathrm{A}_{🔑}\mathrm{B}_{🔑}\mathrm{E}}$. Notice

notice that $\hat{\gamma}_{\mathrm{A}_{🔑}\mathrm{B}_{🔑}\mathrm{E}}$ is not the reduced state of the purification of $\hat{\gamma}_{\mathrm{A}_{🔑}\mathrm{B}_{🔑}\mathrm{A}_{🛡}\mathrm{B}_{🛡}}$, which actually equals

$$\hat{\gamma}_{\mathrm{A}_{🔑}\mathrm{B}_{🔑}\mathrm{A}_{🛡}\mathrm{B}_{🛡}} = \mathcal{T}(\hat{\Phi}_{\mathrm{A}_{🔑}\mathrm{B}_{🔑}} \otimes \sigma_{\mathrm{A}_{🛡}\mathrm{B}_{🛡}}).$$

Indeed, measuring the key first and then computing the purification actually gives away the information about the key to Eve, because it gives her the purifying system of the measurement.

Twisting

We will soon define a distillation rate of private states, where we do not care about their shield, but only about their key size, it will then be useful to have a well defined set of private states for a fixed key size, which we do now. The process of tensoring with a state on the shield and conjugating with a controlled unitary, is called twisting, and can be generalized to any pair of control/target system. Let $\rho \in \mathcal{D}(\mathrm{C})$, we can construct the set of all the twistings of ρ on T as

$$\Upsilon(\rho, \mathrm{T}) := \mathsf{CU}(\mathrm{C}, \mathrm{T}) \circ (\rho \otimes \mathcal{D}(\mathrm{T})).$$

where now for general states ρ, all of the controlled unitary will be relevant, and not just the part controlled on $|ii\rangle\langle ii|$, as is the case for the maximally entangled states. Furthermore, just like we have mentioned before for $\hat{\gamma}_{\mathrm{A}_{🔑}\mathrm{B}_{🔑}\mathrm{E}}$, $\mathsf{CU}(\mathrm{C}, \mathrm{T})$ are all the invariant operations for $\hat{\rho}_{\mathrm{CE}}$, where $\rho_{\mathrm{CTE}} \in \mathcal{D}(\mathrm{CTE})$ is any purification of ρ.

The set of all private states on $\mathrm{A}_{🔑}\mathrm{B}_{🔑}\mathrm{A}_{🛡}\mathrm{B}_{🛡}$ is then easily constructed as the set of all the twisting of $\Phi_{\mathrm{A}_{🔑}\mathrm{B}_{🔑}}$ on the shield, namely as

$$\Upsilon(\Phi_{\mathrm{A}_{🔑}\mathrm{B}_{🔑}}, \mathrm{A}_{🛡}\mathrm{B}_{🛡}).$$

As for the maximally entangled state, we will also use the notation γ^m for states $\gamma^m \in \Upsilon(\Phi^m, \mathrm{A}_{🛡}\mathrm{B}_{🛡})$. We will also call private bits the private states with single-qubit key systems, namely the elements of $\Upsilon(\Phi^+, \mathrm{A}_{🛡}\mathrm{B}_{🛡})$, and just like for maximally entangled states we denote with $\gamma^{\pm} := \mathcal{Z}^{\pm}_{\mathrm{B}_{🔑}}(\gamma)$ the encodings for the particular case of private bits (notice that the phase can be applied also on Alice). The class of private states is also closed under tensor product, and private bits are the units in this class.

Private state distillation

All private states are very entangled states, in particular they are all at least as entangled as the maximally entangled state Φ in terms of the relative entropy of entanglement. Namely, using the unitary invariance of the relative entropy, it is straightforward to check that for all private states $\gamma \in \Upsilon(\Phi, \mathrm{A}_{\text{🛡}}\mathrm{B}_{\text{🛡}})$ it holds

$$E_R(\Phi) \leqslant E_R^{\infty}(\gamma) \leqslant E_R(\gamma).$$

In particular, this means that private states are a resource, since they cannot be produced from bipartite systems under separable operations, and we can thus define the following finite rate for the distillation of private states as

$$\tilde{K}_D^{\varepsilon}(\rho) := \sup\left\{\log k : \Upsilon(\Phi^{\log k}, \mathrm{A}_{\text{🛡}}\mathrm{B}_{\text{🛡}}) \approx_{\varepsilon} \mathsf{LOCC}(\mathbb{C}^k\mathrm{A}_{\text{🛡}}{:}\mathbb{C}^k\mathrm{B}_{\text{🛡}}\langle \mathrm{A}{:}\mathrm{B}) \circ \rho\right\},$$

and the corresponding best achievable rate

$$\tilde{K}_D(\rho) := \lim_{\varepsilon\to 0}\limsup_{n\to\infty}\frac{1}{n}\tilde{K}_D^{\varepsilon}(\rho^{\otimes n}).$$

As for the distillable entanglement, changing the maps from LOCC to $\mathsf{LOCC}_{\rightarrow}$ defines the smaller rate $\tilde{K}_D^{\rightarrow}(\rho)$. Similarly, changing to separable maps defines the larger rate $\tilde{K}_{D,\mathsf{SEP}}(\rho)$. These rates are always larger than their distillable entanglement counterparts. Namely, we have $E_{D,\mathsf{L}}(\rho) \leqslant \tilde{K}_{D,\mathsf{L}}(\rho)$, because distilling entanglement, in comparison to distilling private states, is just a restriction on the target states. The relative entropy of entanglement and its regularization, that we mentioned to be upper bounds on the distillable entanglement, are actually upper bounds on private state distillation [Hor+09]. Namely,

$$\tilde{K}_{D,\mathsf{SEP}}(\rho) \leqslant E_R^{\infty}(\rho) \leqslant E_R(\rho). \tag{3.3}$$

Private states might be a bit confusing because there is no restriction on what state can go in the shield. For example, $\Phi^m \otimes \Phi_{\mathrm{A}_{\text{🛡}}\mathrm{B}_{\text{🛡}}}$ is a valid private state, but the shield should clearly be part of the key. Therefore, this state will never be the output of an optimal protocol in private state distillation. The supremum in the rate will be achieved by private states for which the key part cannot be made larger asymptotically. These are called irreducible private states. More precisely, a private state γ^m for states $\gamma^m \in \Upsilon(\Phi^m, \mathrm{A}_{\text{🛡}}\mathrm{B}_{\text{🛡}})$ is said to be irreducible if [Hor+09]

$$\tilde{K}_D(\gamma^m) = m. \tag{3.4}$$

These are the private states that are actually interesting in the definition of distillable key because they are the ones approximated by optimal distillation protocols.

However, the only feasible way to prove that a private state is irreducible is to upper bound the distillable key via some other entanglement measure. In particular, it was also shown in [Hor+09] that:

$$E_R^{\infty}(\gamma^m) \leqslant m + E_R^{\infty}(\hat{\gamma})$$
$$E_R(\gamma^m) \leqslant m + E_R(\hat{\gamma}).$$

We thus have that for any private state $\gamma \in \Upsilon(\Phi^m, \mathrm{A}_{\text{🛡}}\mathrm{B}_{\text{🛡}})$, the relative entropy of entanglement and its regularization are exactly $E_R^{\infty}(\gamma^m) = E_R(\gamma^m) = m$ if and only if the key-attacked state $\hat{\gamma}$ is separable, namely $\hat{\gamma} \in \mathcal{S}(\mathrm{A}_{\text{🔑}}\mathrm{A}_{\text{🛡}}{:}\mathrm{B}_{\text{🔑}}\mathrm{B}_{\text{🛡}})$, in which case γ is also an irreducible private state. Such states are sometimes called strictly irreducible [1; Hor+16]. The example given in Equation (3.1) is strictly irreducible, and in general, if $\tau \in \mathcal{D}(\mathrm{A}_{\text{🛡}}\mathrm{B}_{\text{🛡}})$

is the maximally mixed state, then any private state $\gamma \in \Upsilon(\Phi, \tau)$ will be strictly irreducible, because the key attacked state will simply be the separable state $\hat{\Phi} \otimes \tau$.

There are no known examples of irreducible private states that are not strictly irreducible. As we will see in the next section after we introduce Bell private state, finding such an example would solve a long-standing important open problem in key distillation.

Equivalence with key distillation

Arguably a significant advance in the understanding of key distillation, was the proof in [Hor+05b] that the rate of distillable private states equals the distillable key

$$\begin{aligned}\tilde{K}_D(\rho) &= K_D(\rho)\\ \tilde{K}_D^{\rightarrow}(\rho) &= K_D^{\rightarrow}(\rho),\end{aligned}$$

together with examples of private states with small distillable entanglement, and of PPT states ε-close to private states for arbitrary ε. This result was used to show that a quantum key distribution protocol, that could not be purified into an entanglement distillation protocol, could indeed be purified into a private states distillation protocol [RS07].

There are various consequences of the above. The equality shows through Equation (3.3), that the relative entropy of entanglement and its regularization are upper bounds on the distillable key

$$\tilde{K}_{D,\mathsf{SEP}}(\rho) \leqslant E_R^{\infty}(\rho) \leqslant E_R(\rho). \tag{3.5}$$

The existence of private state with negligible distillable entanglement implies that distillable entanglement and distillable key are different and their difference can be unbounded. The existence of PPT states arbitrarily close to private states, means that PPT operations can generate key even though they cannot generate distillable entanglement, and thus $\tilde{K}_{D,\mathsf{PPT}}$ is not well defined and will be infinite. Finally, the equality of the rates implies that key distillation is actually a bipartite problem. The intuition behind this is the following. We can decide to purify an LOPC protocol at any time, however the system of Eve is already accounted for, so the purifying systems are secure. This comes from the assumption that the local operations at Alice and Bob are not attacked, and thus these systems are actually the auxiliary systems at Alice and Bob's lab that we called shields. Since the total system is pure at all times, we have the equivalent choice of either keeping track of the key systems and Eve's system, or of the key systems and the shields/lab systems. In the second case Eve is determined by the assumption that she holds the purification *at all times*. The public communication is then taken care of by this assumption, because the purification of the classical communication always contains a coherent copy of the message.

Private states explain why some quantum key distribution protocols can have a rate much larger than entanglement based distillation protocols [SS07]. Such protocols can have quite unintuitive processes, whereby adding local noise to key actually allows to achieve larger rates [RGK05]. The noise is indeed part of the shield and allows for some phases to not be corrected [RS07], an interpretastion that will become general and clear in the next section.

Encoding

Each private state γ with a k-dimensional key system, can be thought of as a set of states for encoding a classical message of dimension k as explained in Section 2.2. What is more, it is a straightforward observation that the message can be encoded locally by Alice or Bob acting on the "raw" private state γ, as shown below. Let, without loss of generality, Bob be the encoder of the message. Let γ be a private state on $\mathrm{A}_{\text{🔑}}\mathrm{A}_{\text{🛡}}{:}\mathrm{B}_{\text{🔑}}\mathrm{B}_{\text{🛡}}$, and define for the moment the states $\gamma_j := \mathcal{Z}^j_{\mathrm{B}_{\text{🔑}}}(\gamma)$, where $\mathcal{Z}$ is the phase flip on $\mathrm{B}_{\text{🔑}}$ with the identities on the remaining systems omitted. Since any controlled unitary commutes with the phase flips on the control, it is readily verified that

$$\gamma_j = \mathcal{T}(\phi_{0j} \otimes \sigma), \tag{3.6}$$

and therefore γ_j are orthogonal states (because the orthogonality of the Bell states is preserved under the unitary). Therefore, any private state can encode a classical source $\{p_j\}$, into the ensemble $\{p_j, \gamma_j\}$, and the message can be recovered with a global measurement. Private states are thus both perfect encoders and perfect key distillable.

Because private states are bipartite states, Alice and Bob might thus not have access to global measurements. Private states can thus have low restricted Holevo information, just like they might have low distillable entanglement. This chapter is indeed dedicated to formally connecting these behaviours, and to exploring the limitations of this connection.

3.1.4 Interlude — Measured Squashed Entanglement

The content of this section is unpublished work in collaboration with Karol Horodecki.

This section is not necessary for the understanding of the remaining material, but it is a simple result that might be of independent interest. We have already seen various entanglement measures. Among them we have seen the relative entropy of entanglement as an upper bound on the distillable key, and the relative entropy from PPT states as an upper bound on the distillable entanglement. Another interesting upper bound on the distillable key is the squashed entanglement [CW04; Tuc99; Tuc02]:

$$E_{sq}(\rho) := \frac{1}{2} \inf_{\Lambda \in \mathrm{CPTP}(\mathrm{E}\rangle\mathrm{E}')} I(\mathrm{A}{:}\mathrm{B}|\mathrm{E}')_{\mathrm{id}_{\mathrm{AB}} \otimes \Lambda(|\rho\rangle\langle\rho|)}$$

where $|\rho\rangle\langle\rho| \in \mathcal{D}(\mathrm{ABE})$ is any purification of $\rho \in \mathcal{D}(\mathrm{AB})$. Because different purifications are related by reversible operations on Eve, their choice does not change the squashed entanglement.

The relative entropy of entanglement and the squashed entanglement are incomparable, as the squashed entanglement of the antisymmetric state is arbitrarily smaller than the relative entropy of entanglement [CSW12], and the relative entropy of entanglement of the so-called flower private state (which we will see as an example in Section 3.4) is arbitrarily smaller than the squashed entanglement [Hor+05a; CW05][2]. One important property of squashed entanglement is that it is additive on tensor products, and it is thus its own regularization. The peculiar aspect that is the reason for this section, is that there are no known definitions of squashed entanglement restricted to measurements (or partial measurements). Since defining restricted measures is a recurring theme in this thesis, we dedicate this section to illustrate how to insert measurements and define a form of restricted squashed entanglement, which can be of independent interest. We leave as future work the development of its possible applications.

Measured intrinsic information

Squashed entanglement is inspired by a classical optimized conditional mutual information known as intrinsic information [MW99; RW03], which is an upper bound on

[2] The squashed entanglement is lockable, while the relative entropy of entanglement is not.

the classical distillable key of a tripartite probability distribution. The corresponding definition for the quantum intrinsic information of a tripartite system $\rho \in \mathcal{D}(\mathrm{ABE})$ is

$$I(\mathrm{A{:}B{\downarrow}E})_\rho := \inf_{\xi \in \mathsf{CPTP}(\mathrm{E}\rangle \mathrm{E}^{\mathrm{out}})} I\left(\mathrm{A{:}B|E^{out}}\right)_{\mathrm{id_{AB}} \otimes \xi(\rho)}$$

where the optimization is also over the output system E'. For any state $\rho \in \mathcal{D}(\mathrm{AB})$, squashed entanglement is then simply

$$E_{sq}(\rho) = \frac{1}{2} I(\mathrm{A{:}B{\downarrow}E})_{|\rho\rangle\langle\rho|} \tag{3.7}$$

Notice that, while the squashed entanglement is additive on tensor products, this additivity comes from computing the intrinsic information on a pure state. In general the intrinsic information is simply sub-additive (proven in [CW04] as part of the proof of additivity of squashed entanglement). The intrinsic information can also be proven to be asymptotically continuous, but is not strictly necessary for our purpose.

The infimum over channels at Eve makes the intrinsic information monotonically increasing under local channels at Eve, as these only reduce the set of channels in the optimization. The intrinsic information is also monotone under LOPC. For completeness, we prove these properties at the end of the section. Because of the monotonicity under LOPC, it is now quite natural to give the following definition.

Definition 12. *Let $\rho \in \mathcal{D}(\mathrm{ABE})$ be a tripartite state. We define the measured intrinsic information of ρ as:*

$$I_{\mathsf{LOPC}}(\mathrm{A} : \mathrm{B{\downarrow}E})_\rho := \sup_{\mathcal{M} \in \mathsf{LOPC}\left(\mathrm{A{:}B}\right\rangle \underline{\mathrm{A}}^{\mathrm{out}}{:}\underline{\mathrm{B}}^{\mathrm{out}}|\underline{\mathrm{C}})} I\left(\mathrm{A^{out}{:}B^{out}{\downarrow}CE}\right)_{\mathcal{M}\otimes \mathrm{id_E}(\rho)}$$

where the supremum is also over the output systems $\mathrm{A^{out}}$, $\mathrm{B^{out}}$ *and* C.

Since the optimization over LOPC measurements has no access to Eve's system, it preserves the monotonicity at Eve, and thus the measured intrinsic information is monotonically increasing under channels at Eve. The measured intrinsic information is also an LOPC monotone by definition, because of the supremum over LOPC measurements. We prove these properties below for completeness. Monotonicity of the intrinsic information guarantees that $I_{\mathsf{LOPC}}(\mathrm{A} : \mathrm{B{\downarrow}E})$ is indeed finite, as it immediately follows that

$$I_{\mathsf{LOPC}}(\mathrm{A} : \mathrm{B{\downarrow}E}) \leqslant I(\mathrm{A{:}B{\downarrow}E}).$$

Furthermore, the inequality will in general be strict for entangled states. Indeed while the universal upper bound for the conditional mutual information is

$$I(\mathrm{A{:}B|E}) \leqslant 2 \min \left\{ \log |\mathrm{A}|, \log |\mathrm{B}| \right\},$$

values above $\min \left\{ \log |\mathrm{A}|, \log |\mathrm{B}| \right\}$ can only be achieved by entangled states. For any separable state $\rho \in \mathcal{S}(\mathrm{A{:}B})$, we have instead $I(\mathrm{A{:}B}) \leqslant \min \left\{ \log |\mathrm{A}|, \log |\mathrm{B}| \right\}$[3] which then implies that if $\rho \in \mathcal{D}(\mathrm{ABE})$ is classical on Alice and Bob, as is the case with the outcome of an LOPC measurement, then $I(\mathrm{A{:}B|E})_\rho \leqslant \min \left\{ \log |\mathrm{A}|, \log |\mathrm{B}| \right\}$. We thus have

$$I_{\mathsf{LOPC}}(\mathrm{A} : \mathrm{B{\downarrow}E}) \leqslant \min \left\{ \log |\mathrm{A}|, \log |\mathrm{B}| \right\}, \tag{3.8}$$

even though

$$I(\mathrm{A{:}B{\downarrow}E}) \leqslant 2 \min \left\{ \log |\mathrm{A}|, \log |\mathrm{B}| \right\}.$$

[3] This is shown using the monotonicity under local operations. More precisely the mutual information of $\sum_i p_i \rho_i \otimes \sigma_i$ will be smaller than the mutual information of $\sum_i p_i |i\rangle\langle i| \otimes \sigma_i$, the Holevo information of $\{p_i, \sigma_i\}$, which is upper bounded by $\log |\mathrm{B}|$.

Proofs of the monotonicities.

1. The Intrinsic information is monotonically increasing under channels at Eve. For any channel $\Lambda \in \mathsf{CPTP}\big(\mathrm{E}^{\mathrm{in}}\rangle \mathrm{E}\big)$ and any state $\rho \in \mathcal{D}\big(\mathrm{ABE}^{\mathrm{in}}\big)$:

$$\begin{aligned} I(\mathrm{A{:}B{\downarrow}E})_{\mathrm{id}_{\mathrm{AB}}\otimes\Lambda(\rho)} &= \inf_{\xi\in\mathsf{CPTP}(\mathrm{E}\rangle\mathrm{E}^{\mathrm{out}})} I\big(\mathrm{A{:}B|E}^{\mathrm{out}}\big)_{\mathrm{id}_{\mathrm{AB}}\otimes(\xi\circ\Lambda)(\rho)} \\ &\geqslant \inf_{\xi\circ\Lambda\in\mathsf{CPTP}(\mathrm{E}^{\mathrm{in}}\rangle\mathrm{E}^{\mathrm{out}})} I\big(\mathrm{A{:}B|E}^{\mathrm{out}}\big)_{\mathrm{id}_{\mathrm{AB}}\otimes(\xi\circ\Lambda)(\rho)} \\ &= I\big(\mathrm{A{:}B{\downarrow}E}^{\mathrm{in}}\big)_{\rho}. \end{aligned}$$

2. The intrinsic information is monotonically decreasing under LOPC. Let $\Lambda \in \mathsf{LOPC}\big(\mathrm{A}^{\mathrm{in}}{:}\mathrm{B}^{\mathrm{in}}\rangle \mathrm{A}^{\mathrm{out}}{:}\mathrm{B}^{\mathrm{out}}|\underline{\mathrm{C}}\big)$ be an LOPC operation and let ρ be any input state on $\mathrm{A}^{\mathrm{in}}\mathrm{B}^{\mathrm{in}}\mathrm{E}$. By the monotonicity of the conditional mutual information we have

$$\begin{aligned} I\big(\mathrm{A}^{\mathrm{out}}{:}\mathrm{B}^{\mathrm{out}}{\downarrow}\mathrm{CE}\big)_{\Lambda\otimes\mathrm{id}_{\mathrm{E}}(\rho)} &= \inf_{\xi\in\mathsf{CPTP}(\mathrm{CE}\rangle\mathrm{E}')} I\big(\mathrm{A}^{\mathrm{out}}{:}\mathrm{B}^{\mathrm{out}}|\mathrm{E}'\big)_{(\mathrm{id}_{\mathrm{AB}}\otimes\xi)\circ(\Lambda\otimes\mathrm{id}_{\mathrm{E}})(\rho)} \\ &\leqslant \inf_{\xi\in\mathsf{CPTP}(\mathrm{E}\rangle\mathrm{E}')} I\big(\mathrm{A}^{\mathrm{out}}{:}\mathrm{B}^{\mathrm{out}}|\mathrm{CE}'\big)_{\Lambda\otimes\xi(\rho)} \\ &\leqslant \inf_{\xi\in\mathsf{CPTP}(\mathrm{E}\rangle\mathrm{E}')} I\big(\mathrm{A}^{\mathrm{in}}{:}\mathrm{B}^{\mathrm{in}}|\mathrm{E}'\big)_{\mathrm{id}_{\mathrm{AB}}\otimes\xi(\rho)} \\ &= I\big(\mathrm{A}^{\mathrm{in}}{:}\mathrm{B}^{\mathrm{in}}{\downarrow}\mathrm{E}\big)_{\rho}. \end{aligned}$$

3. The measured intrinsic information is monotonically increasing under channels at Eve. For any channel $\Lambda \in \mathsf{CPTP}\big(\mathrm{E}^{\mathrm{in}}\rangle \mathrm{E}\big)$ and any state $\rho \in \mathcal{D}\big(\mathrm{ABE}^{\mathrm{in}}\big)$:

$$\begin{aligned} I_{\mathsf{LOPC}}(\mathrm{A} : \mathrm{B}{\downarrow}\mathrm{E})_{\mathrm{id}_{\mathrm{AB}}\otimes\Lambda(\rho)} &= \sup_{\mathcal{M}\in\mathsf{LOPC}\left(\mathrm{A{:}B}\rangle \underline{\mathrm{A}}^{\mathrm{out}}{:}\underline{\mathrm{B}}^{\mathrm{out}}|\underline{\mathrm{C}}\right)} I\big(\mathrm{A}^{\mathrm{out}}{:}\mathrm{B}^{\mathrm{out}}{\downarrow}\mathrm{CE}\big)_{\mathcal{M}\otimes\Lambda_{\mathrm{E}}(\rho)} \\ &\geqslant \sup_{\mathcal{M}\in\mathsf{LOPC}\left(\mathrm{A{:}B}\rangle \underline{\mathrm{A}}^{\mathrm{out}}{:}\underline{\mathrm{B}}^{\mathrm{out}}|\underline{\mathrm{C}}\right)} I\big(\mathrm{A}^{\mathrm{out}}{:}\mathrm{B}^{\mathrm{out}}{\downarrow}\mathrm{CE}^{\mathrm{in}}\big)_{\mathcal{M}\otimes\mathrm{id}_{\mathrm{E}}(\rho)} \\ &= I_{\mathsf{LOPC}}(\mathrm{A} : \mathrm{B}{\downarrow}\mathrm{E} \in)_{\rho}. \end{aligned}$$

4. The measured intrinsic information is monotonically decreasing under LOPC. Let $\rho \in \mathcal{D}\big(\mathrm{A}^{\mathrm{in}}\mathrm{B}^{\mathrm{in}}\mathrm{E}\big)$ and let $\Lambda \in \mathsf{LOPC}\big(\mathrm{A}^{\mathrm{in}}{:}\mathrm{B}^{\mathrm{in}}\rangle \mathrm{A{:}B}|\underline{\mathrm{M}}\big)$, then

$$\begin{aligned} I_{\mathsf{LOPC}}(\mathrm{A} : \mathrm{B}{\downarrow}\mathrm{ME})_{\Lambda\otimes\mathrm{id}_{\mathrm{E}}(\rho)} &= \sup_{\mathcal{M}\in\mathsf{LOPC}\left(\mathrm{A{:}B}\rangle \underline{\mathrm{A}}^{\mathrm{out}}{:}\underline{\mathrm{B}}^{\mathrm{out}}|\underline{\mathrm{M}}'\right)} I\big(\mathrm{A}^{\mathrm{out}}{:}\mathrm{B}^{\mathrm{out}}{\downarrow}\mathrm{M}'\mathrm{ME}\big)_{(\mathcal{M}\otimes\mathrm{id}_{\mathrm{CE}})\circ(\Lambda\otimes\mathrm{id}_{\mathrm{E}})(\rho)} \\ &\leqslant \sup_{\mathcal{M}\circ\Lambda\in\mathsf{LOPC}\left(\mathrm{A}^{\mathrm{in}}{:}\mathrm{B}^{\mathrm{in}}\rangle \underline{\mathrm{A}}^{\mathrm{out}}{:}\underline{\mathrm{B}}^{\mathrm{out}}|\underline{\mathrm{M}}'\underline{\mathrm{M}}\right)} I\big(\mathrm{A}^{\mathrm{out}}{:}\mathrm{B}^{\mathrm{out}}{\downarrow}\mathrm{M}'\mathrm{ME}\big)_{(\mathcal{M}\circ\Lambda)\otimes\mathrm{id}_{\mathrm{E}}(\rho)} \\ &\leqslant I_{\mathsf{LOPC}}(\mathrm{A} \in: \mathrm{B} \in {\downarrow}\mathrm{E})_{\rho} \end{aligned}$$

Measured squashed entanglement

Having defined the measured intrinsic information, inserting the measurements in the squashed entanglement is now as simple as generalizing Equation (3.7).

Definition 13. *Let ρ be any state of* AB. *We define the measured squashed entanglement as:*

$$\underline{E}_{sq}(\rho) := I_{\mathsf{LOPC}}(\mathrm{A} : \mathrm{B}{\downarrow}\mathrm{E})_{|\rho\rangle\langle\rho|}$$

where $|\rho\rangle\langle\rho| \in \mathcal{D}(\mathrm{ABE})$ is any purification of ρ.

Notice that there is no factor of 2, to account for Equation (3.8). The measured squashed entanglement will likely not be additive on tensor products any more, and we do not know whether it is sub- or super-additive. Therefore we cannot claim that it admits a regularization. However by monotonicity of the measured intrinsic information we immediately know that

$$\underline{E}_{sq}(\rho) \leqslant 2E_{sq}(\rho),$$

and thus by the additivity of squashed entanglement we have

$$\limsup_{n\to\infty} \frac{1}{n}\underline{E}_{sq}(\rho^{\otimes n}) \leqslant 2E_{sq}(\rho).$$

The measured squashed entanglement is indeed and entanglement measure, as shown by the properties proven below. The proof of the last one is particularly simple given the LOPC monotonicity of the measured intrinsic information, and it implies an equally simple proof that squashed entanglement is an upper bound on the distillable key.

Lemma 14. *On pure states $\rho \in \mathcal{D}(\mathrm{AB})$ we have $\underline{E}_{sq}(\rho) = H(\mathrm{A})_\rho = H(\mathrm{B})_\rho$.*

Lemma 15. *$\underline{E}_{sq}(\rho)$ is an LOCC monotone.*

Lemma 16. *For any bipartite state $\rho \in \mathcal{D}(\mathrm{AB})$*

$$K_D(\rho) \leqslant \limsup_{n\to\infty} \frac{1}{n}\underline{E}_{sq}(\rho^{\otimes n}).$$

Proof of Lemma 14. Let $\rho = |\rho\rangle\langle\rho| \in \mathcal{D}(\mathrm{AB})$ be a pure state. We need to prove that $\underline{E}_{sq}(\rho) = H(\mathrm{A})_\rho$. Notice that on pure states we have $H(\mathrm{A})_\rho = I(\mathrm{A{:}B})_\rho$. We thus need to prove $\underline{E}_{sq}(\rho) = I(\mathrm{A{:}B})_\rho$. First notice that because ρ is pure

$$\underline{E}_{sq}(\rho) := I_{\mathsf{LOPC}}(\mathrm{A} : \mathrm{B}{\downarrow}\mathrm{E})_{|\rho\rangle\langle\rho|} = I_{\mathsf{LOPC}}(\mathrm{A} : \mathrm{B})_\rho.$$

Now, $\underline{E}_{sq}(\rho) \leqslant I(\mathrm{A{:}B})_\rho$ follows immediately from the monotonicity of the conditional mutual information. The lower bound is achieved using local measurement on the Schmidt-decomposition basis of $|\rho\rangle$ (see for example [NC02]). □

Proof of Lemma 15. Let $\rho \in \mathcal{D}(\mathrm{A^{in}{:}B^{in}})$ and $\Lambda \in \mathsf{LOCC}(\mathrm{A^{in}{:}B^{in}}\rangle\mathrm{A^{out}{:}B^{out}})$. We need to show that

$$\underline{E}_{sq}(\Lambda(\rho)) \leqslant \underline{E}_{sq}(\rho). \tag{3.9}$$

To show monotonicity under local operations, we can use the Stinespring dilation as done in the proof of squashed entanglement in [CW04]. Then addition of pure ancillas and local unitaries leave the measured squashed entanglement unchanged, because they are reversible operations in LOPC. For the partial trace, it is enough to notice that $|\rho\rangle_{\tilde{\mathrm{A}}\mathrm{ABE}}$ is both a purification of ρ_{AB} and of $\rho_{\tilde{\mathrm{A}}\mathrm{AB}}$ and thus

$$\begin{aligned}
\underline{E}_{sq}(\rho_{\tilde{\mathrm{A}}\mathrm{AB}}) &= \sup_{\mathcal{M}\in\mathsf{LOPC}(\tilde{\mathrm{A}}\mathrm{A{:}B}\rangle\underline{\mathrm{A}}'{:}\underline{\mathrm{B}}'|\underline{\mathrm{M}})} \inf_{\xi\in\mathsf{CPTP}(\mathrm{ME}\rangle\mathrm{E}')} I(\mathrm{A}'{:}\mathrm{B}'|\mathrm{E}')_{(\mathrm{id}_{\mathrm{A}'\mathrm{B}'}\otimes\xi)\circ(\mathcal{M}\otimes\mathrm{id}_{\mathrm{E}}(|\rho\rangle\langle\rho|)} \\
&\geqslant \sup_{\mathcal{M}\in\mathsf{LOPC}(\mathrm{A{:}B}\rangle\underline{\mathrm{A}}'{:}\underline{\mathrm{B}}'|\underline{\mathrm{M}})} \inf_{\xi\in\mathsf{CPTP}(\mathrm{ME}\rangle\mathrm{E}')} I(\tilde{\mathrm{A}}\mathrm{A}'{:}\mathrm{B}'|\mathrm{E}')_{(\mathrm{id}_{\tilde{\mathrm{A}}\mathrm{A}'\mathrm{B}'}\otimes\xi)\circ(\mathcal{M}\otimes\mathrm{id}_{\tilde{\mathrm{A}}\mathrm{E}}(|\rho\rangle\langle\rho|)} \\
&\geqslant \sup_{\mathcal{M}\in\mathsf{LOPC}(\mathrm{A{:}B}\rangle\underline{\mathrm{A}}'{:}\underline{\mathrm{B}}'|\underline{\mathrm{M}})} \inf_{\xi\in\mathsf{CPTP}(\mathrm{ME}\rangle\mathrm{E}')} I(\mathrm{A}'{:}\mathrm{B}'|\tilde{\mathrm{A}}\mathrm{E}')_{(\mathrm{id}_{\tilde{\mathrm{A}}\mathrm{A}'\mathrm{B}'}\otimes\xi)\circ(\mathcal{M}\otimes\mathrm{id}_{\tilde{\mathrm{A}}\mathrm{E}}(|\rho\rangle\langle\rho|)} \\
&\geqslant \sup_{\mathcal{M}\in\mathsf{LOPC}(\mathrm{A{:}B}\rangle\underline{\mathrm{A}}'{:}\underline{\mathrm{B}}'|\underline{\mathrm{M}})} \inf_{\xi\in\mathsf{CPTP}(\tilde{\mathrm{A}}\mathrm{ME}\rangle\mathrm{E}')} I(\mathrm{A}'{:}\mathrm{B}'|\mathrm{E}')_{(\mathrm{id}_{\mathrm{A}'\mathrm{B}'}\otimes\xi)\circ(\mathcal{M}\otimes\mathrm{id}_{\tilde{\mathrm{A}}\mathrm{E}}(|\rho\rangle\langle\rho|)} \\
&= \underline{E}_{sq}(\rho_{\mathrm{AB}})
\end{aligned}$$

where the first inequality follows because we are taking the supremum over a smaller set, the second follows from the chain rule and strong sub-additivity, and the third one follows because we are taking the infimum over a larger set.

It remains only to prove monotonicity for the classical communication. Let C be the classical system to be communicated, let $\rho \in \mathcal{D}(\mathrm{CAB})$ be a state that is classical on C

$$\rho = \sum_i p_i \, |i\rangle\langle i|_{\mathrm{C}} \otimes \rho_{i,\mathrm{AB}},$$

and let us distinguish with $\rho_{\mathrm{CA:B}}$ and $\rho_{\mathrm{A:CB}}$ the cases where the message system is at Alice or Bob. This state has the following purification $|\rho\rangle\langle\rho| \in \mathcal{D}(\mathrm{ACBE})$:

$$|\rho\rangle = \sum_i \sqrt{p_i} \, |ii\rangle_{\mathrm{CE}'} \otimes |\rho_i\rangle_{\mathrm{ABE}},$$

where with $|\rho_i\rangle \in \mathcal{D}(\mathrm{ABE})$ we denote the purification of $\rho_i \in \mathcal{D}(\mathrm{AB})$.

Then we start with C at Alice and compute:

$$\begin{aligned} \underline{E}_{sq}(\rho_{\mathrm{CA:B}}) &= I_{\mathsf{LOPC}}(\mathrm{CA} : \mathrm{B}{\downarrow}\mathrm{E}'\mathrm{E})_{|\rho\rangle\langle\rho|} \\ &\geqslant I_{\mathsf{LOPC}}(\mathrm{CA} : \mathrm{B}{\downarrow}\mathrm{E}'\mathrm{E})_{\sum |ii\rangle\langle ii|_{\mathrm{CE}'} \otimes |\rho_i\rangle\langle\rho_i|_{\mathrm{ABE}}} \end{aligned}$$

simply by letting Alice perform the local measurement in the computational basis on C. We now send the message using public communication and by monotonicity of the measured intrinsic information we get

$$\underline{E}_{sq}(\rho_{\mathrm{CA:B}}) \geqslant I_{\mathsf{LOPC}}(\mathrm{A} : \mathrm{CB}{\downarrow}\mathrm{ME}'\mathrm{E})_{\sum |iii\rangle\langle iii|_{\mathrm{CME}'} \otimes |\rho_i\rangle\langle\rho_i|_{\mathrm{ABE}}}.$$

The current state can be equivalently achieved by measuring $|\rho\rangle\langle\rho|$ in the computational basis at E', and making a copy at M. We let $\xi \in \mathsf{CPTP}(\mathrm{E}'\rangle\mathrm{E}'\mathrm{M})$ be such map. Then, because the measured intrinsic information is monotonically increasing under local channels at Eve, we have

$$\begin{aligned} \underline{E}_{sq}(\rho_{\mathrm{CA:B}}) &\geqslant I_{\mathsf{LOPC}}(\mathrm{A} : \mathrm{CB}{\downarrow}\mathrm{ME}'\mathrm{E})_{\sum |iii\rangle\langle iii|_{\mathrm{CME}'} \otimes |\rho_i\rangle\langle\rho_i|_{\mathrm{ABE}}} \\ &= I_{\mathsf{LOPC}}(\mathrm{A} : \mathrm{CB}{\downarrow}\mathrm{ME}'\mathrm{E})_{\mathrm{id}_{\mathrm{ABC}} \otimes \xi(|\rho\rangle\langle\rho|)} \\ &\geqslant I_{\mathsf{LOPC}}(\mathrm{A} : \mathrm{CB}{\downarrow}\mathrm{E}'\mathrm{E})_{|\rho\rangle\langle\rho|} \\ &= \underline{E}_{sq}(\rho_{\mathrm{A:CB}}). \end{aligned}$$

□

Proof of Lemma 16. Fix $\varepsilon > 0$ and the number of copies n. Let $|\rho\rangle\langle\rho| \in \mathcal{D}(\mathrm{ABE})$ be a purification. Let $\Lambda \in \mathsf{LOPC}(\mathrm{A}^{\otimes n}{:}\mathrm{B}^{\otimes n}\rangle\mathbb{C}^k{:}\mathbb{C}^k|\underline{\mathrm{M}}) \otimes \mathrm{id}_{\mathrm{E}^{\otimes n}}$ be a protocol that acting on $|\rho\rangle\langle\rho|^{\otimes n}$, gets ε-close to the state with perfect key $\hat{\Phi}^{\log k} \otimes \sigma_{\mathrm{ME}^{\otimes n}}$. Then, by the monotonicity proved in Lemma 15:

$$\begin{aligned} \frac{1}{n}\underline{E}_{sq}(\rho^{\otimes n}) &= \frac{1}{n} I_{\mathsf{LOPC}}(\mathrm{A}^{\otimes n} : \mathrm{B}^{\otimes n}{\downarrow}\mathrm{E}^{\otimes n})_{|\rho\rangle\langle\rho|^{\otimes n}} \\ &\geqslant \frac{1}{n} I_{\mathsf{LOPC}}(\mathbb{C}^k : \mathbb{C}^k{\downarrow}\mathrm{ME}^{\otimes n})_{\Lambda(|\rho\rangle\langle\rho|^{\otimes n})}. \end{aligned}$$

We now let $\mathcal{M} \in \mathsf{LO}(\mathbb{C}^k{:}\mathbb{C}^k)$ be the local measurement in the computational basis. By definition we then have

$$\begin{aligned} \frac{1}{n}\underline{E}_{sq}(\rho^{\otimes n}) &\geqslant I(\mathbb{C}^k{:}\mathbb{C}^k{\downarrow}\mathrm{ME}^{\otimes n})_{(\mathcal{M} \otimes \mathrm{id}_{\mathrm{ME}^{\otimes n}}) \circ \Lambda(|\rho\rangle\langle\rho|)} \\ &= \inf_{\xi \in \mathsf{CPTP}(\mathrm{ME}^{\otimes n}\rangle\mathrm{E}^{\mathrm{out}})} I\big(\mathbb{C}^k{:}\mathbb{C}^k|\mathrm{E}^{\mathrm{out}}\big)_{(\mathcal{M} \otimes \xi) \circ \Lambda(|\rho\rangle\langle\rho|^{\otimes n})}. \end{aligned}$$

By the monotonicity of the trace distance, since $\Lambda(|\rho\rangle\langle\rho|^{\otimes n})$ and $\hat{\Phi}^{\log k} \otimes \sigma_{\mathrm{ME}^{\otimes n}}$ are ε-close, then $(\mathcal{M} \otimes \xi) \circ \Lambda(|\rho\rangle\langle\rho|^{\otimes n})$ and $(\mathcal{M} \otimes \xi) \circ (\hat{\Phi}^{\log k} \otimes \sigma_{\mathrm{ME}^{\otimes n}}) = \hat{\Phi}^{\log k} \otimes \xi(\sigma)_{\mathrm{E}^{\mathrm{out}}}$ will also be ε-close. We now use the continuity of the conditional mutual information [Shi17], which says that for ϵ-close states ϱ and ς, the conditional mutual information satisfies

$$|I(\mathrm{A{:}B|E})_{\varrho} - I(\mathrm{A{:}B|E})\varsigma| \leqslant 2\epsilon \log d + 2g(\varepsilon)$$

where $d = \min\{|\mathrm{A}|, |\mathrm{B}|\}$. Thus we find:

$$\begin{aligned}
\frac{1}{n}\underline{E}_{sq}(\rho^{\otimes n}) &\geqslant \frac{1}{n} \inf_{\xi \in \mathsf{CPTP}(\mathrm{ME}^{\otimes n}\rangle \mathrm{E}^{\mathrm{out}})} \Big(I(\mathbb{C}^k{:}\mathbb{C}^k|\underline{\mathrm{E}}^{\mathrm{out}})_{\hat{\Phi}^{\log k}\otimes\xi(\sigma)} - 2\epsilon \log k + 2g(\varepsilon) \Big) \\
&= \frac{1}{n}(I(\mathbb{C}^k{:}\mathbb{C}^k)_{\hat{\Phi}^{\log k}} - 2\epsilon \log k + 2g(\varepsilon)) \\
&= (1-2\varepsilon)\frac{\log k}{n} + \frac{1}{n}2g(\varepsilon).
\end{aligned}$$

Taking the supremum over the all protocols then gives

$$\frac{1}{n}\underline{E}_{sq}(\rho^{\otimes n}) \geqslant (1-2\varepsilon)\frac{1}{n}K_D^{\varepsilon}(\rho^{\otimes n}) + \frac{1}{n}2g(\varepsilon).$$

Since this is true for all n we then have

$$\begin{aligned}
\limsup_{n\to\infty} \frac{1}{n}\underline{E}_{sq}(\rho^{\otimes n}) &\geqslant \limsup_{n\to\infty} \left[(1-2\varepsilon)\frac{1}{n}K_D^{\varepsilon}(\rho^{\otimes n}) + \frac{1}{n}2g(\varepsilon) \right]. \\
&= (1-2\varepsilon) \limsup_{n\to\infty} \frac{1}{n}K_D^{\varepsilon}(\rho^{\otimes n}).
\end{aligned}$$

Taking the limit $\varepsilon \to 0$ ends the proof. □

We have shown how to make use of the LOPC monotonicity of the conditional mutual information to insert measurements in the squashed entanglement. We have also mentioned that there exist some states, like the flower private state that will be presented in Section 3.4, where the squashed entanglement is very large compared to the relative entropy of entanglement. Forcing Alice and Bob to measure is exactly what could kill Alice and Bob's correlations to the advantage of Eve.

3.2 Bell Private states

The content of this section can be found in [1].

The first example of a private state with low distillable entanglement was constructed as follows [Hor+05b]:

$$\gamma = p^+ \cdot \Phi^+ \otimes \sigma^+ + p^- \cdot \Phi^- \otimes \sigma^- \tag{3.10}$$

$$\begin{aligned}
\hat{\gamma} &= p^+ \cdot \hat{\Phi}^+ \otimes \sigma^+ + p^- \cdot \hat{\Phi}^- \otimes \sigma^- \\
&= \hat{\Phi} \otimes (p^+\sigma^+ + p^-\sigma^-)
\end{aligned} \tag{3.11}$$

with $\sigma^{\pm}$ orthogonal states. If $\sigma^{\pm}$ are chosen to be the extremal Werner states ρ_s, ρ_a, and the probabilities $p^+ = d_s/d^2$ and $p^- = d_a/d^2$, then the logarithmic negativity, and thus the distillable entanglement, can be quickly computed to be $O(1/d)$. The intuition behind the example is the following. Indistinguishable orthogonal states $\sigma^{\pm}$ like the Werner states, should hinder the ability to correct the phase flip locally, and thus they should suppress the distillable entanglement. Nevertheless, because the states are orthogonal, the perfect secret bit is still protected from the environment. Private states like the ones in Equation (3.10) are only an example of what we call Bell private states; we will see different examples of Bell and non-Bell private states in Section 3.4. The purpose of this section is to show that all private states can be converted in a form diagonal in the Bell basis for any dimension. In this way, we can argue that all private states are phase flipped maximally entangled states, with the phase information stored in perfectly encoding states, so that later we can formalize the connection with quantum data hiding for all private states. The conversion to Bell private states acts only on the key systems, preserves the key size, it can be implemented with just local operations and shared randomness (and thus it is in $\mathsf{LOCC}_{\rightarrow}$), and most importantly it is reversible (with just local operations). Because this map is reversible, any entanglement property of a private state will be preserved in the transformation to Bell private states, and therefore the value of any entanglement measure[4] will not change. At the end of the section we will bound the distinguishability of Bell private bits in terms of the distinguishability of their shields. This will be useful later at the end of the chapter, when we will choose the encoding shield states at random.

Key-correlated states

Let now the key systems $\mathrm{A}_{\text{🔑}}$ and $\mathrm{B}_{\text{🔑}}$ have equal dimension, let $m = \log|\mathrm{A}_{\text{🔑}}| = \log|\mathrm{B}_{\text{🔑}}|$, and let all the implicit indexes in this section be summed over $\mathbb{Z}_{2^m}$, unless otherwise stated. We define key-correlated states as those states on $\mathrm{A}_{\text{🔑}}\mathrm{B}_{\text{🔑}}\mathrm{A}_{\text{🛡}}\mathrm{B}_{\text{🛡}}$ supported only on the maximally correlated subspace of $\mathrm{A}_{\text{🔑}}\mathrm{B}_{\text{🔑}}$. They have no bit-flip error and we can write them as:

$$\begin{aligned} \rho &:= \sum_{ij\in\mathbb{Z}_{2^m}} |\phi_{0i}\rangle\langle\phi_{0j}|_{\mathrm{A}_{\text{🔑}}\mathrm{B}_{\text{🔑}}} \otimes P_{ij,\mathrm{A}_{\text{🛡}}\mathrm{B}_{\text{🛡}}} \\ \hat{\rho} &= \frac{1}{2^m} \sum_{i\in\mathbb{Z}_{2^m}} \mathcal{Z}^i_{\mathrm{B}_{\text{🔑}}}(\rho) \end{aligned} \tag{3.12}$$

where $P_{ij} \in \mathcal{L}(\mathrm{A}_{\text{🛡}}\mathrm{B}_{\text{🛡}})$ are the block matrices of the shield, and $\mathcal{Z}^i$ are the phase-flip channels. We introduce the states above because, some of the statements will also be true for this class of states. Notice that key-correlated states include some separable states, as for example the maximally correlated state. We say that a key-correlated state is diagonal in the Bell basis, if it can be written in the form

$$\begin{aligned} \rho &= \sum_i p_i \cdot \phi_{0i} \otimes \sigma_i \\ \hat{\rho} &= \hat{\Phi} \otimes \sum_i p_i \sigma_i \end{aligned} \tag{3.13}$$

for some states $\sigma_i \in \mathcal{D}(\mathrm{A}_{\text{🛡}}\mathrm{B}_{\text{🛡}})$ and probabilities p_i. Notice that because the Bell states ϕ_{0i} are orthogonal and pure, as global states on $\mathrm{A}_{\text{🔑}}\mathrm{B}_{\text{🔑}}$ they encode the classical information just as good as $|i\rangle\langle i|$ and thus we have that the Holevo information of the ensemble of shield states is the same as the mutual information between the key and the shield:

$$\chi(\{p_i, \sigma_i\}) = I(\mathrm{A}_{\text{🔑}}\mathrm{B}_{\text{🔑}}:\mathrm{A}_{\text{🛡}}\mathrm{B}_{\text{🛡}})_\rho.$$

Finally, notice that because the key systems are in a mixture of Bell states, their distillable entanglement, which we recall is equal to the coherent information for states on the maximally correlated subspace, is simply

$$E_D(\rho_{\mathrm{A}_{\text{🔑}}\mathrm{B}_{\text{🔑}}}) = m - H\Big(\textstyle\sum_i p_i\, |i\rangle\langle i|\Big),$$

which for uniform probabilities is zero.

[4] more specifically any function that is $\mathsf{LOCC}_{\rightarrow}$ monotone.

Bell private states

We finally define Bell private states as those key-correlated states that are diagonal in the Bell basis, with shield states σ_i all orthogonal to each other. All Bell private states are private states. This can be proved either by checking that the measurement in $\mathrm{A}_{\text{🗝}}\mathrm{B}_{\text{🗝}}$ gives perfectly secure bits or by showing that they admit an expression as private states, here we show the latter. $P_{\sigma_\perp}$ below plays no active role, it is only needed to complete the twisting, so that the shield state is not required to have full support.

Lemma 17. *γ^m is a Bell private state if and only if it can be written as $\gamma^m = T(\Phi \otimes \sigma)T^\dagger$, where*

$$T = \sum_{ij \in \mathbb{Z}_{2^m}} |ij\rangle\langle ij| \otimes U_\sigma^i \qquad\qquad \sigma = \sum_{i \in \mathbb{Z}_{2^m}} p_i \sigma_i$$

and

$$U_\sigma := \sum_{i \in \mathbb{Z}_{2^m}} \omega^i P_{\sigma_i} + P_{\sigma_\perp} \qquad\qquad P_{\sigma_\perp} := \mathbb{1} - \sum_i P_{\sigma_i}$$

with P_{σ_i} the projectors onto the supports of σ_i.

Proof. We simply provide the sequence of equalities to transform $\sum p_k \phi_{0k} \otimes \sigma_k$ back and forth to the form above.

$$\begin{aligned}
\gamma^m &= \sum_k p_k \phi_{0k} \otimes \sigma_k \\
&= \sum_{ijk} p_k \frac{1}{2^m} \omega^{ik} |ii\rangle\langle jj| \, \omega^{-jk} \otimes P_{\sigma_k} \sigma_k P_{\sigma_k} \\
&= \sum_{ijk} \frac{1}{2^m} |ii\rangle\langle jj| \otimes (\omega^{ik} P_{\sigma_k}) \cdot (p_k \sigma_k) \cdot (\omega^{-jk} P_{\sigma_k}) \\
&= \sum_{ijkab} \frac{1}{2^m} |ii\rangle\langle jj| \otimes (\omega^{ia} P_{\sigma_a}) \cdot (p_k \sigma_k) \cdot (\omega^{-jb} P_{\sigma_b}) \\
&= \sum_{ij} \frac{1}{2^m} |ii\rangle\langle jj| \otimes \Big(\sum_a \omega^a P_{\sigma_a} + P_{\sigma_\perp}\Big)^i \cdot \sigma \cdot \Big(\sum_b \omega^b P_{\sigma_b} + P_{\sigma_\perp}\Big)^{-j} \\
&= \sum_{ij} \frac{1}{2^m} |ii\rangle\langle jj| \otimes U_\sigma^i \sigma U_\sigma^{-j} \\
&= \sum_{ij} (|i\rangle\langle i| \otimes U_\sigma^i) \cdot (\tfrac{1}{2^m} |ii\rangle\langle jj| \otimes \sigma) \cdot (|j\rangle\langle j| \otimes U_\sigma^{j\dagger}) \\
&= T(\Phi^m \otimes \sigma) T^\dagger
\end{aligned}$$

where we used the orthogonality of the σ_k in the identity $\sum_{ka} P_{\sigma_a} \sigma_k = \sum_k P_{\sigma_k} \sigma_k$. □

As a consequence of the above theorem, for Bell private states in any dimension we can provide a block form, which is usually used to display the properties of some private states [PH10]. For $|\mathrm{A}_{\text{🗝}}| = |\mathrm{B}_{\text{🗝}}| = 2$, any private state admits a block form [Hor+09], namely it can be written as:

$$\gamma^1 = \frac{1}{2} \begin{pmatrix} \sqrt{K^\dagger K} & 0 & 0 & K^\dagger \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ K & 0 & 0 & \sqrt{KK^\dagger} \end{pmatrix}$$

where K is a suitable matrix of unit trace norm. This can be easily seen by recalling that any matrix K admits a singular value decomposition and noticing that the decomposition can be used to extract the initial σ and the unitaries U_0 and U_1 in the controlled unitary. However, this does not work in higher dimension because then additional unitaries are needed to specify the controlled unitary. This is not true for Bell private states, indeed, a Bell private state only needs to specify a single unitary U_σ. This allows one to write a block form for all private states by exploiting the fact that U_σ and σ commute.

Corollary 18. *γ^m is a Bell private state if and only if*

$$\gamma^m = \frac{1}{2^m}\sum_{ij} |ii\rangle\langle jj| \otimes |K| \left(\frac{K}{|K|}\right)^{i-j}$$

for some normal operator K such that $\|K\|_1 = 1$ and $K^{2^m} \geqslant 0$, and where K^{-1} is the pseudo inverse.

Proof. Set $K = \sigma U_\sigma$. Then $|K| = \sigma$, $\frac{K}{|K|} = U_\sigma$ and

$$U_\sigma^i \sigma U_\sigma^{-j} = \sigma U_\sigma^{i-j} = |K| \left(\frac{K}{|K|}\right)^{i-j}. \qquad \square$$

With the definitions in place we can now move onto the main result of the section.

The reversible map

To define the reversible LOCC map, consider two copies of the key systems, $\mathrm{A}_{\text{🔑}}\mathrm{B}_{\text{🔑}}$ and $\mathrm{A}'_{\text{🔑}}\mathrm{B}'_{\text{🔑}}$. We use the *BNOT* channel $\mathcal{BNOT}$ with $\mathrm{A}'_{\text{🔑}}\mathrm{B}'_{\text{🔑}}$ as target systems, as defined in Section 1.5.4, and the swap channel $\mathcal{S}$ between $\mathrm{A}_{\text{🔑}}\mathrm{B}_{\text{🔑}}$ and $\mathrm{A}'_{\text{🔑}}\mathrm{B}'_{\text{🔑}}$.

Lemma 19. *Let $\rho \in \mathcal{D}(\mathrm{A}_{\text{🔑}}\mathrm{B}_{\text{🔑}})$ and define $\mathcal{E}^{\mathrm{Bell}}_{\mathrm{A}_{\text{🔑}}\mathrm{B}_{\text{🔑}}} \in \mathsf{CPTP}(\mathrm{A}_{\text{🔑}}\mathrm{B}_{\text{🔑}}\rangle \mathrm{A}'_{\text{🔑}}\mathrm{B}'_{\text{🔑}}\mathrm{A}_{\text{🔑}}\mathrm{B}_{\text{🔑}})$ as*

$$\mathcal{E}^{\mathrm{Bell}}_{\mathrm{A}_{\text{🔑}}\mathrm{B}_{\text{🔑}}}(\rho) := \mathcal{S}_{\mathrm{A}_{\text{🔑}}\mathrm{B}_{\text{🔑}},\mathrm{A}'_{\text{🔑}}\mathrm{B}'_{\text{🔑}}} \circ \mathcal{BNOT}^{-1}_{\mathrm{A}_{\text{🔑}}\mathrm{B}_{\text{🔑}},\mathrm{A}'_{\text{🔑}}\mathrm{B}'_{\text{🔑}}}\left(\rho \otimes \hat{\Phi}_{\mathrm{A}'_{\text{🔑}}\mathrm{B}'_{\text{🔑}}}\right).$$

Then if ρ is a key-correlated state on $\mathrm{A}_{\text{🔑}}\mathrm{B}_{\text{🔑}}\mathrm{A}_{\text{🛡}}\mathrm{B}_{\text{🛡}}$:

$$\begin{aligned} \mathcal{E}^{\mathrm{Bell}}_{\mathrm{A}_{\text{🔑}}\mathrm{B}_{\text{🔑}}}(\rho) &= \frac{1}{2^m}\sum_{i\in\mathbb{Z}_{2^m}} \phi_{i,\mathrm{A}'_{\text{🔑}}\mathrm{B}'_{\text{🔑}}} \otimes \mathcal{Z}^i_{\mathrm{B}_{\text{🔑}}}(\rho) \qquad (3.14) \\ \mathcal{E}^{\mathrm{Bell}}_{\mathrm{A}_{\text{🔑}}\mathrm{B}_{\text{🔑}}}(\hat{\rho}) &= \hat{\Phi}_{\mathrm{A}'_{\text{🔑}}\mathrm{B}'_{\text{🔑}}} \otimes \hat{\rho}. \end{aligned}$$

Proof. From Equation (1.11) we know that the effect of the *CNOT* on phase flips is, for all $i,k \in \mathbb{Z}_{2^m}$

$$(Z^{k+i} \otimes Z^i)CNOT^\dagger = CNOT^\dagger(Z^k \otimes Z^i).$$

Namely, the target adds its phase to the control, and thus we get through Lemma 2, that for all $i,k,l \in \mathbb{Z}_{2^m}$

$$\mathcal{BNOT}^{-1}\left(|\phi_{0k}\rangle\langle\phi_{0l}|_{\mathrm{A}_{\text{🔑}}\mathrm{B}_{\text{🔑}}} \otimes \phi_{0i,\mathrm{A}'_{\text{🔑}}\mathrm{B}'_{\text{🔑}}}\right) = \mathcal{Z}^i_{\mathrm{B}_{\text{🔑}}}\left(|\phi_{0k}\rangle\langle\phi_{0l}|_{\mathrm{A}_{\text{🔑}}\mathrm{B}_{\text{🔑}}}\right) \otimes \phi_{0i,\mathrm{A}'_{\text{🔑}}\mathrm{B}'_{\text{🔑}}}.$$

We now add $\hat{\Phi} = \frac{1}{2^m}\sum_i \phi_{0i}$ as a target, we decompose ρ as in Equation (3.12) (general key-correlated states), and complete the operation with the swap channel to get the result

$$\begin{aligned} \mathcal{E}^{\mathrm{Bell}}_{\mathrm{A}_{\text{🔑}}\mathrm{B}_{\text{🔑}}}(\rho) &= \sum_{kl}\left[\mathcal{S} \circ \mathcal{BNOT}^{-1}(|\phi_{0k}\rangle\langle\phi_{0l}| \otimes \hat{\Phi})\right] \otimes P_{kl} \\ &= \frac{1}{2^m}\sum_{ikl} \mathcal{S}\left(\mathcal{Z}^i_{\mathrm{B}_{\text{🔑}}}(|\phi_{0k}\rangle\langle\phi_{0l}|) \otimes \phi_{0i}\right) \otimes P_{kl} \end{aligned}$$

$$= \frac{1}{2^m} \sum_{ikl} \phi_{0i} \otimes \mathcal{Z}^i_{\mathrm{B}_{🔑}} (|\phi_{0k}\rangle\langle\phi_{0l}|) \otimes P_{kl}$$
$$= \frac{1}{2^m} \sum_{i} \phi_{0i} \otimes \mathcal{Z}^i_{\mathrm{B}_{🔑}} (\rho).$$

The phase flips on the control systems commute with the *BNOT* and thus for the key-attacked state, we get

$$\begin{aligned}
\mathcal{E}^{\mathrm{Bell}}_{\mathrm{A}_{🔑}\mathrm{B}_{🔑}}(\hat\rho) &= \frac{1}{2^m} \sum_i \mathcal{E}^{\mathrm{Bell}}_{\mathrm{A}_{🔑}\mathrm{B}_{🔑}} \circ \mathcal{Z}^i_{\mathrm{B}_{🔑}}(\rho) \\
&= \frac{1}{2^m} \sum_i \mathcal{Z}^i_{\mathrm{B}_{🔑}} \circ \mathcal{E}^{\mathrm{Bell}}_{\mathrm{A}_{🔑}\mathrm{B}_{🔑}}(\rho) \\
&= \frac{1}{2^{2m}} \sum_{ij} \phi_{0j} \otimes \mathcal{Z}^{i+j}_{\mathrm{B}_{🔑}}(\rho) \\
&= \frac{1}{2^m} \sum_j \phi_{0j} \otimes \sum_i \mathcal{Z}^i_{\mathrm{B}_{🔑}}(\rho) \\
&= \hat\Phi \otimes \hat\rho. \qquad \square
\end{aligned}$$

Because $\hat\Phi$ is separable, it can be produced simply with shared randomness or one way LOCC. The *BNOT* and the swap are local as long as the system is partitioned as $\mathrm{A}_{🔑}\mathrm{A}'_{🔑}{:}\mathrm{B}_{🔑}\mathrm{B}'_{🔑}$. The map can be reversed by simply inverting the *BNOT* and tracing out the target, which requires only local operations. Notice that the output key systems are the same as $\mathrm{A}_{🔑}\mathrm{B}_{🔑}$, but the output shield systems are now $\mathrm{A}_{🔑}\mathrm{A}_{🛡}\mathrm{B}_{🔑}\mathrm{B}_{🛡}$. Bell private states now come as a special case. For a private state γ we know from Section 3.1.3 that $\mathcal{Z}^i_{\mathrm{B}_{🔑}}(\gamma)$ are perfect encoding states, namely they are all orthogonal to each other, we thus et a Bell private state at the output. In other words, the private states get mapped into Bell private states that are encoded by the original private states.

Because of the reversible map, we can now always assume without loss of generality that the key distillation protocols distill Bell private states, namely that the distillable key is given by the best achievable rate of distilling Bell private states. The reversible map also proves that after using the key of a private state, the remaining shield can be distilled independently. Namely for any private state $\gamma \in \Upsilon(\Phi, \mathrm{A}_{🛡}\mathrm{B}_{🛡})$, we have

$$K_D(\gamma) \geqslant K(\Phi) + K_D(\hat\gamma)$$

as shown using a bound from [Hor+16] in the following chain of inequalities

$$K_D(\gamma^{\log k}) = K_D(\mathcal{E}^{\mathrm{Bell}}(\gamma)) \geqslant \log k + \frac{1}{k} K_D(\hat\gamma^{\otimes k}) = \log k + K_D(\hat\gamma).$$

This implies that for all irreducible private states we have $K_D(\hat\gamma) = 0$, and thus all optimal distillation protocols can be thought of as putting all the input state in the shield, and then concentrating the key in the key systems until there is no key left in the shield. Whether the remaining shield is separable at the end of the protocol, namely whether all irreducible private states are also strictly irreducible, depends on whether there exist so-called key-bound entangled states [Hor+16]. Key-bound entangled states are entangled states that have zero distillable key, and are a long-standing open question even in the classical setting (see [OSS14] and references there in). The existence of irreducible but not strictly irreducible private states would imply the existence of key-bound entangled states, as shown by the schema below

$$\overbrace{\underbrace{E_R(\hat\gamma) \geqslant E_R(\gamma) - E_R(\Phi)}_{> \text{ if } \gamma \text{ not strictly irred.}} \geqslant K_D(\gamma) - K_D(\Phi)}^{> \text{ if } \hat\gamma \text{ is key-bound entangled}} \underbrace{= K_D(\hat\gamma)}_{\gamma \text{ irred. assump}} = 0.$$

3.2.1 Single-shot Data Hiding

The content of this section comes from [2].

We have seen that we can reduce private states to Bell private states, however the resulting shield states are simply the encoding of the classical information onto the private states itself. It will be a useful tool in the next section, but it will not tell us anything new about the distinguishability of the ensemble of private states $\gamma_j = \mathcal{Z}^j(\gamma)$. If we construct a Bell private state from arbitrary orthogonal states, then there will be a non trivial relation between the distinguishability of the shields and the distinguishability of the Bell private states. In this section we will compute bounds on the local distance of Bell private bits, as a function of the local distance of their shields. Thus relating the data hiding property of the shield to the data hiding properties of the private bits. These will, by construction, be bounds only on the distinguishability of single copies of the private states, but it will be enough for our purposes when we choose the shield states at random in the end of the chapter. In the next section we will see how the distillable entanglement provides a bound on the asymptotic distinguishability of private states.

Recall that for any pair of orthogonal states $\sigma^{\pm}$ on $\mathrm{A}_{\text{🛡}}\mathrm{B}_{\text{🛡}}$, we have the (Bell) private states on $\mathrm{A}_{\text{🔑}}\mathrm{A}_{\text{🛡}}\mathrm{B}_{\text{🔑}}\mathrm{B}_{\text{🛡}}$

$$\gamma^{\pm} = \frac{1}{2}\big(\Phi^{+} \otimes \sigma^{\pm} + \Phi^{-} \otimes \sigma^{\mp}\big).$$

The key-attacked state is indeed $\hat{\gamma} = \frac{1}{2}\gamma^{+} + \frac{1}{2}\gamma^{-}$. With the results of Section 2.1.3, it is now not hard to see that, if $\sigma^{\pm}$ on $\mathrm{A}_{\text{🛡}}\mathrm{B}_{\text{🛡}}$ are data-hiding for PPT or separable measurements on $\mathrm{A}_{\text{🛡}}\mathrm{B}_{\text{🛡}}$, then so are the constructed private states $\gamma^{\pm}$ on $\mathrm{A}_{\text{🔑}}\mathrm{A}_{\text{🛡}}\mathrm{B}_{\text{🔑}}\mathrm{B}_{\text{🛡}}$ for PPT or separable measurements on $\mathrm{A}_{\text{🔑}}\mathrm{A}_{\text{🛡}}\mathrm{B}_{\text{🔑}}\mathrm{B}_{\text{🛡}}$.

Lemma 20. *Let $|\mathrm{A}_{\text{🔑}}| = |\mathrm{B}_{\text{🔑}}| = 2$ and let $\gamma^{\pm} \in \mathcal{D}(\mathrm{A}_{\text{🔑}}\mathrm{A}_{\text{🛡}}\mathrm{B}_{\text{🔑}}\mathrm{B}_{\text{🛡}})$ be Bell private states with shields $\sigma^{\pm} \in \mathcal{D}(\mathrm{A}_{\text{🛡}}\mathrm{B}_{\text{🛡}})$. Then,*

$$\begin{aligned}\frac{1}{2}\|\gamma^{+} - \gamma^{-}\|_{\mathsf{SEP}(\underline{\mathrm{A}}_{\text{🔑}}\underline{\mathrm{A}}_{\text{🛡}}:\underline{\mathrm{B}}_{\text{🔑}}\underline{\mathrm{B}}_{\text{🛡}})} &\leqslant \frac{3}{2}\|\sigma^{+} - \sigma^{-}\|_{\mathsf{SEP}(\underline{\mathrm{A}}_{\text{🛡}}:\underline{\mathrm{B}}_{\text{🛡}})},\\ \frac{1}{2}\|\gamma^{+} - \gamma^{-}\|_{\mathsf{PPT}(\underline{\mathrm{A}}_{\text{🔑}}\underline{\mathrm{A}}_{\text{🛡}}:\underline{\mathrm{B}}_{\text{🔑}}\underline{\mathrm{B}}_{\text{🛡}})} &\leqslant \|\sigma^{+} - \sigma^{-}\|_{\mathsf{PPT}(\underline{\mathrm{A}}_{\text{🛡}}:\underline{\mathrm{B}}_{\text{🛡}})}.\end{aligned}$$

If the key-attacked state is separable, this implies

$$\begin{aligned}\|\gamma^{\pm} - \mathcal{S}(\mathrm{A}_{\text{🔑}}\mathrm{A}_{\text{🛡}}:\mathrm{B}_{\text{🔑}}\mathrm{B}_{\text{🛡}})\|_{\mathsf{SEP}(\underline{\mathrm{A}}_{\text{🔑}}\underline{\mathrm{A}}_{\text{🛡}}:\underline{\mathrm{B}}_{\text{🔑}}\underline{\mathrm{B}}_{\text{🛡}})} &\leqslant \frac{3}{2}\|\sigma^{+} - \sigma^{-}\|_{\mathsf{SEP}(\underline{\mathrm{A}}_{\text{🛡}}:\underline{\mathrm{B}}_{\text{🛡}})},\\ \|\gamma^{\pm} - \mathcal{S}(\mathrm{A}_{\text{🔑}}\mathrm{A}_{\text{🛡}}:\mathrm{B}_{\text{🔑}}\mathrm{B}_{\text{🛡}})\|_{\mathsf{PPT}(\underline{\mathrm{A}}_{\text{🔑}}\underline{\mathrm{A}}_{\text{🛡}}:\underline{\mathrm{B}}_{\text{🔑}}\underline{\mathrm{B}}_{\text{🛡}})} &\leqslant \|\sigma^{+} - \sigma^{-}\|_{\mathsf{PPT}(\underline{\mathrm{A}}_{\text{🛡}}:\underline{\mathrm{B}}_{\text{🛡}})}.\end{aligned}$$

Proof. Let $\Delta = \sigma^{+} - \sigma^{-}$. Note that the states γ and $\hat{\gamma}$ are such that

$$\gamma^{\pm} - \hat{\gamma} = \pm\frac{1}{2}(\gamma^{+} - \gamma^{-}) = \pm\frac{1}{4}\left(\Phi^{+} - \Phi^{-}\right) \otimes \Delta.$$

Therefore the second pair of bounds derives directly from the first. For the first pair of bounds, we have that for any class of channels closed under local operations, the L norm satisfies

$$\frac{1}{2}\|\gamma^{+} - \gamma^{-}\|_{\mathsf{L}} \leqslant \frac{1}{4}\|\Phi^{+} \otimes \Delta\|_{\mathsf{L}} + \frac{1}{4}\|\Phi^{-} \otimes \Delta\|_{\mathsf{L}} = \frac{1}{2}\|\Phi^{+} \otimes \Delta\|_{\mathsf{L}}.$$

We impose closure under local operations so that $\|\Phi^{+} \otimes \Delta\|_{\mathsf{L}} = \|\Phi^{-} \otimes \Delta\|_{\mathsf{L}}$. The two announced inequalities then follow from Proposition 5, which tells us that $\|\Phi^{\pm} \otimes \Delta\|_{\mathsf{SEP}} \leqslant 3\|\Delta\|_{\mathsf{SEP}}$ and $\|\Phi^{\pm} \otimes \Delta\|_{\mathsf{PPT}} \leqslant 2\|\Delta\|_{\mathsf{PPT}}$. □

The result for PPT measurements was provided only for completeness, as later we will actually want to prove that the privates states we construct are not data-hiding for PPT measurements. The following lemma will be useful in this context.

Lemma 21. *Let* $|\mathrm{A}_{🔑}| = |\mathrm{B}_{🔑}| = 2$, *let* $\gamma^{\pm} \in \mathcal{D}(\mathrm{A}_{🔑}\mathrm{A}_{🛡}\mathrm{B}_{🔑}\mathrm{B}_{🛡})$ *be Bell private states with shields* $\sigma^{\pm} \in \mathcal{D}(\mathrm{A}_{🛡}\mathrm{B}_{🛡})$, *and let* L *be either separable or PPT measurements. Then we have*

$$\left\|\gamma^{\pm} - \mathcal{S}(\mathrm{A}_{🔑}\mathrm{A}_{🛡}{:}\mathrm{B}_{🔑}\mathrm{B}_{🛡})\right\|_{\mathsf{L}(\underline{\mathrm{A}}_{🔑}\underline{\mathrm{A}}_{🛡}:\underline{\mathrm{B}}_{🔑}\underline{\mathrm{B}}_{🛡})} \geqslant \frac{1}{3}\left\|\sigma^{+} - \sigma^{-}\right\|_{\mathsf{L}(\underline{\mathrm{A}}_{🛡}:\underline{\mathrm{B}}_{🛡})}.$$

Proof. Let $c := \frac{1}{2}\|\sigma^{+} - \sigma^{-}\|_{\mathsf{L}(\underline{\mathrm{A}}_{🛡}:\underline{\mathrm{B}}_{🛡})}$. By definition, this means that there exists a binary measurement $(M, \mathbb{1} - M) \in \mathsf{L}(\underline{\mathrm{A}}_{🛡}{:}\underline{\mathrm{B}}_{🛡})$ such that without loss of generality we have

$$\operatorname{tr} M\sigma^{+} - \operatorname{tr} M\sigma^{-} = c = \operatorname{tr}(\mathbb{1} - M)\sigma^{-} - \operatorname{tr}(\mathbb{1} - M)\sigma^{+}.$$

We now apply to γ the distillation protocol Λ that first tries to distinguish $\sigma^{\pm}$ using the above measurement, and then corrects the phase flip of the maximally entangled state accordingly. This only needs additional one way communication and local operations, and thus Λ is still in L. Since after the correction the measurement outcome is not needed anymore, it is traced out at the end of the protocol. The resulting state is

$$\begin{aligned} \Lambda(\gamma) &= \frac{1}{2}\operatorname{tr}\bigl(M\sigma^{+} + (\mathbb{1} - M)\sigma^{-}\bigr)\Phi^{+} + \frac{1}{2}\operatorname{tr}\bigl(M\sigma^{-} + (\mathbb{1} - M)\sigma^{+}\bigr)\Phi^{-} \\ &= \frac{1}{2}(1 + c)\Phi^{+} + \frac{1}{2}(1 - c)\Phi^{-}. \end{aligned}$$

We now apply the isotropic twirl to $\mathrm{A}_{🔑}\mathrm{B}_{🔑}$ to produce the isotropic state with fidelity $(1 + c)/2$, and denote by $\tilde{\Lambda}$ the operation Λ followed by a twirl. Since the twirl again only needs one-way LOCC, then $\tilde{\Lambda}$ is also in L. Notice that $\tilde{\Lambda}(\gamma)$ is always entangled, indeed it was proven in [MWW09] that for separable and PPT measurements we have $c \geqslant 1/|\mathrm{A}_{🔑}|$, which is the threshold for separable isotropic states.

Since $\tilde{\Lambda}$ maps PPT states into PPT states, together with the fact that all PPT isotropic states are also separable, we find that the L distance from separable states of γ can only decrease. Namely, for any separable state $\sigma \in \mathcal{S}(\mathrm{A}_{🔑}\mathrm{A}_{🛡}{:}\mathrm{B}_{🔑}\mathrm{B}_{🛡})$ we have $\tilde{\Lambda}(\sigma) \in \mathcal{S}(\mathrm{A}_{🔑}{:}\mathrm{B}_{🔑})$, and thus

$$\begin{aligned} \|\gamma - \sigma\|_{\mathsf{L}(\underline{\mathrm{A}}_{🔑}\underline{\mathrm{A}}_{🛡}:\underline{\mathrm{B}}_{🔑}\underline{\mathrm{B}}_{🛡})} &\geqslant \left\|\tilde{\Lambda}(\gamma) - \tilde{\Lambda}(\sigma)\right\|_{\mathsf{L}(\underline{\mathrm{A}}_{🔑}:\underline{\mathrm{B}}_{🔑})} \\ &\geqslant \left\|\tilde{\Lambda}(\gamma) - \mathcal{S}(\mathrm{A}_{🔑}{:}\mathrm{B}_{🔑})\right\|_{\mathsf{L}(\underline{\mathrm{A}}_{🔑}:\underline{\mathrm{B}}_{🔑})}. \end{aligned}$$

Taking the infimum then gives

$$\|\gamma - \mathcal{S}(\mathrm{A}_{🔑}\mathrm{A}_{🛡}{:}\mathrm{B}_{🔑}\mathrm{B}_{🛡})\|_{\mathsf{L}(\underline{\mathrm{A}}_{🔑}\underline{\mathrm{A}}_{🛡}:\underline{\mathrm{B}}_{🔑}\underline{\mathrm{B}}_{🛡})} \geqslant \left\|\tilde{\Lambda}(\gamma) - \mathcal{S}(\mathrm{A}_{🔑}{:}\mathrm{B}_{🔑})\right\|_{\mathsf{L}(\underline{\mathrm{A}}_{🔑}:\underline{\mathrm{B}}_{🔑})}.$$

Since $\tilde{\Lambda}(\gamma)$ is an isotropic state, Lemma 10 now gives us the desired lower bounds:

$$\left\|\tilde{\Lambda}(\gamma) - \mathcal{S}(\mathrm{A}_{🔑}{:}\mathrm{B}_{🔑})\right\|_{\mathsf{L}(\underline{\mathrm{A}}_{🔑}:\underline{\mathrm{B}}_{🔑})} = \frac{4}{3}\left(\frac{1 + c}{2} - \frac{1}{2}\right) = \frac{2}{3}c. \qquad \square$$

3.3 Entanglement Distillation and Data Hiding

The content of this section is a refined version of [1].

So far we have seen how to convert private states into Bell private states, and given some bounds on their single-copy distinguishability. We now show that Bell private states with low distillable entanglement are states that hide the phase of the maximally entangled states from local detection, even asymptotically. Specifically, we give a lower bound for the distillable entanglement in terms of the Holevo information. This lower bound is the rate achieved by the best protocol that first distinguishes the shield states and then corrects the phase, as done in Lemma 21. In the next chapter we will be able to use these bounds to connect the distillable entanglement with the key rate in the quantum repeater.

Lemma 22. *Let ρ be a Bell-diagonal key-correlated state, namely of the form $\rho = \sum_i p_i \phi_{0i} \otimes \sigma_i$. Let* L *be either* $\mathsf{LOCC}_{\rightarrow}$, LOCC, SEP *or* PPT. *Then:*

$$E_{D,\mathsf{L}}(\rho) \geqslant E_D(\rho_{\mathrm{A}_{🔑}\mathrm{B}_{🔑}}) + \chi_{\mathsf{L}(\underline{\mathrm{A}}_{🛡}:\mathrm{B}_{🛡})}(\{p_i, \sigma_i\})$$

where $\sigma = \sum_i p_i \sigma_i$, and $E_D(\rho_{\mathrm{A}_{🔑}\mathrm{B}_{🔑}}) = I(\mathrm{A}_{🔑}\rangle \mathrm{B}_{🔑})_\rho$ is independent of the choice of L.

Notice that we can use any class of operations L contained in t PPT, that contains classical communication from Alice to Bob. L simply needs to achieve the hashing bound for states in the maximally correlated subspace. Since L being a class and containing one-way communication imply that L contains $\mathsf{LOCC}_{\rightarrow}$, this is enough to ensure that the hashing bound can be achieved.

Proof. Let $\Lambda \in \mathsf{L}(\underline{\mathrm{A}}_{🛡}:\mathrm{B}_{🛡})$. Because of the classical communication, we can assume without loss of generality that $\Lambda \in \mathsf{L}(\mathrm{A}_{🛡}:\mathrm{B}_{🛡})|\mathbb{C} : \mathrm{M})$ for some system M (see Equation (1.6)). We now let Alice and Bob perform their channel Λ and define:

$$\tilde{\rho} := \sum_i p_i \phi_{0i} \otimes \Lambda(\sigma_i). \tag{3.15}$$

We have $E_{D,\mathsf{L}}(\rho) \geqslant E_{D,\mathsf{L}}(\tilde{\rho})$. Alice and Bob now use the hashing protocol to achieve the rate given by the hashing bound [DW05]. We find:

$$\begin{aligned} E_{D,\mathsf{L}}(\tilde{\rho}) &\geqslant I(\mathrm{A}_{🔑}\rangle \mathrm{B}_{🔑}\mathrm{M}) \\ &= H(\mathrm{B}_{🔑}\mathrm{M})_{\tilde{\rho}} - H(\mathrm{A}_{🔑}\mathrm{B}_{🔑}\mathrm{M})_{\tilde{\rho}}. \end{aligned}$$

However, because the key systems are in a mixture of Bell states, tracing out $\mathrm{A}_{🔑}$ will leave $\mathrm{B}_{🔑}$ in product with M. Therefore:

$$\begin{aligned} E_{D,\mathsf{L}}(\tilde{\rho}) &\geqslant H(\mathrm{B}_{🔑})_{\tilde{\rho}} + H(\mathrm{M})_{\tilde{\rho}} - H(\mathrm{A}_{🔑}\mathrm{B}_{🔑}\mathrm{M})_{\tilde{\rho}} \\ &= [H(\mathrm{A}_{🔑}\mathrm{B}_{🔑})_{\tilde{\rho}} + H(\mathrm{M})_{\tilde{\rho}} - H(\mathrm{A}_{🔑}\mathrm{B}_{🔑}\mathrm{M})_{\tilde{\rho}}] \\ &\quad + [H(\mathrm{B}_{🔑})_{\tilde{\rho}} - H(\mathrm{A}_{🔑}\mathrm{B}_{🔑})_{\tilde{\rho}}] \\ &= I(\mathrm{A}_{🔑}\mathrm{B}_{🔑}:\mathrm{M})_{\tilde{\rho}} + E_D(\rho_{\mathrm{A}_{🔑}\mathrm{B}_{🔑}}), \end{aligned} \tag{3.16}$$

where in the last steps we added and removed the entropy of $\mathrm{A}_{🔑}\mathrm{B}_{🔑}$ and then used that $I(\mathrm{A}_{🔑}\rangle \mathrm{B}_{🔑})$ is the distillable entanglement of the key part [Rai01; HH04], see also Section 3.2. Since $I(\mathrm{A}_{🔑}\mathrm{B}_{🔑}:\mathrm{M})$ is monotone under local operations on $\mathrm{A}_{🔑}\mathrm{B}_{🔑}$, and in particular invariant under unitary operations on $\mathrm{A}_{🔑}\mathrm{B}_{🔑}$ and under tracing out pure ancillas, we have

$$\begin{aligned} I(\mathrm{A}_{🔑}\mathrm{B}_{🔑}:\mathrm{M})_{\tilde{\rho}} &= I(\mathrm{A}_{🔑}\mathrm{B}_{🔑}:\mathrm{M})_{\sum_i p_i |0i\rangle\langle 0i| \otimes \Lambda(\sigma_i)} \\ &= I(\mathrm{B}_{🔑}:\mathrm{M})_{\sum_i p_i |i\rangle\langle i| \otimes \Lambda(\sigma_i)} = \chi(\{p_i, \Lambda(\sigma_i)\}). \end{aligned}$$

Taking the supremum over Λ proves the claim. □

The Holevo information quantifies the distinguishability between states, measuring how well they encode classical data; the restricted Holevo information then quantifies how much of this distinguishability is left when Alice and Bob can only act locally. In the particular case of private states, the σ_i states of Equation (3.13) are orthogonal and thus encode perfectly, and i can be recovered with a global measurement. However, Lemma 22 implies that if the distillable entanglement is low, then the local distinguishability of σ_i is low and i cannot be determined accurately locally: the σ_i are data hiding. We will now see that because of the reversible map, we can think of the private state itself as the data-hiding state, where i is encoded using the local phase flip. This is important, as we will see examples where the L Holevo information of the private states is strictly larger than the one for the shields.

Corollary 23. *Let* $m = \log|\mathrm{A}_{🔑}| = \log|\mathrm{B}_{🔑}|$. *Let ρ be any key-correlated state of* $\mathrm{A}_{🔑}\mathrm{B}_{🔑}\mathrm{A}_{🛡}\mathrm{B}_{🛡}$ *and let* $\rho_i := \mathcal{Z}^i_{\mathrm{B}_{🔑}}(\rho)$. *For any* $\mathsf{L} = \mathsf{LOCC}_{\rightarrow}, \mathsf{LOCC}, \mathsf{SEP}, \mathsf{PPT}$, *it holds:*

$$E_{D,\mathsf{L}}(\rho) \geqslant \chi_{\mathsf{L}(\underline{\mathrm{A}}_{🔑}\underline{\mathrm{A}}_{🛡}:\mathrm{B}_{🔑}\mathrm{B}_{🛡})}\left(\left\{\tfrac{1}{2^m}, \rho_i\right\}\right)$$

Proof. We can use Lemma 22 after using Lemma 19, since the reversible map is contained in any class of operations with communication from Alice to Bob:

$$\begin{aligned} E_{D,\mathsf{L}}(\rho) = E_{D,\mathsf{L}}(\mathcal{E}^{\mathrm{Bell}}(\rho)) &= E_{D,\mathsf{L}}\left(\textstyle\sum_i \tfrac{1}{2^m}\phi_{0i} \otimes \mathcal{Z}^i_{\mathrm{B}_{🔑}}(\rho)\right) \\ &\geqslant \chi_{\mathsf{L}(\underline{\mathrm{A}}_{🔑}\underline{\mathrm{A}}_{🛡}:\mathrm{B}_{🔑}\mathrm{B}_{🛡})}\left(\left\{\tfrac{1}{2^m}, \rho_i\right\}\right). \end{aligned}$$

where the distillable entanglement of the key part is now zero, because of the uniform probability distribution. □

Finally, we remark that these bounds can be regularized. Notice the conceptual discrepancy between the left and the right side of Corollary 23. The distillable entanglement is allowed to perform operations on many copies of the shield, but the Holevo information is single copy. Because the distillable entanglement is already regularized we have for any n

$$E_{D,\mathsf{L}}(\rho) = \frac{1}{n}E_{D,\mathsf{L}}(\rho^{\otimes n}) \geqslant \frac{1}{n}\chi_{\mathsf{L}(\underline{\mathrm{A}}_{🔑}^{\otimes n}\underline{\mathrm{A}}_{🛡}^{\otimes n}:\mathrm{B}_{🔑}^{\otimes n}\mathrm{B}_{🛡}^{\otimes n})}\left(\left\{\tfrac{1}{2^m}, \rho_i\right\}^{\otimes n}\right).$$

If L is closed under tensor product, then the L Holevo information is super-additive, its regularization is well defined and we have

$$E_{D,\mathsf{L}}(\rho) \geqslant \chi^{\infty}_{\mathsf{L}(\underline{\mathrm{A}}_{🔑}\underline{\mathrm{A}}_{🛡}:\mathrm{B}_{🔑}\mathrm{B}_{🛡})}\left(\left\{\tfrac{1}{2^m}, \rho_i\right\}\right).$$

Therefore proving that the distillable entanglement is low, also proves “asymptotic” indistinguishability.

3.4 Interesting classes of private states

Bell private states connect explicitly data hiding and entanglement distillation. Consider a private bit

$$\gamma^{\pm} = p^{+} \cdot \Phi^{+} \otimes \sigma^{\pm} + p^{-} \cdot \Phi^{-} \otimes \sigma^{\mp}.$$

If we consider them encoding states, then we have shown that even if we give Alice and Bob the information of the phase flip, the power of the maximally entangled state is limited, and thus if the shields are disproportionally large, the distinguishability properties of the private bit will not be very different from the ones of the shields. On the other side if we consider private bits for entanglement distillation, then trying to use the distinguishability of the shield can only provide so much entanglement, and if the distillable entanglement is actually low, then the shield state must be data hiding, even asymptotically.

We might be tempted to think, because $\phi^{\pm}$ is not actually available as resource, that distinguishing the shields asymptotically has to be the best protocol, and that touching $\phi^{\pm}$

could only add noise. Indeed, if Alice and Bob simply perform the teleportation protocol using $\Phi^{\pm}$ without the knowledge of the phase, they will introduce a phase error. However, this is not true, the private states can be strictly more distinguishable than their shield.

In this section we present various examples. The first two examples show some private states that are strictly more distinguishable than their shields, showing that the reversible protocol is actually necessary, in that the bound in Corollary 23 can be strictly better than the bound in Lemma 22. In the remaining examples we present private states that have provably small distillable entanglement and thus they are provably data hiding, even for arbitrarily many copies.

3.4.1 The *BNOT* private states

The Bell states, as a basis, are actually perfect encoders (indeed they are a special instance of private states). While the Bell states ϕ_{0i} are perfectly distinguishable even locally, any additional Bell states will not be recoverable by local observers [Bad+03, Equation 8]. In this example we show that using the Bell states to encode the phase flips leads to perfectly distillable private states. Giving an unbounded separation between the restricted Holevo information of the shields and the restricted Holevo information of the private state.

We let $|\mathrm{A}_{\text{🔑}}| = |\mathrm{B}_{\text{🔑}}| = |\mathrm{A}'_{\text{🔑}}| = |\mathrm{B}'_{\text{🔑}}| = |\mathrm{A}_{\text{🛡}}| = |\mathrm{B}_{\text{🛡}}|$ and $m = \log |\mathrm{A}_{\text{🔑}}|$, namely we let the key systems be two copies of $\mathrm{A}_{\text{🔑}}\mathrm{B}_{\text{🔑}}$. We define the *BNOT* private states [1] as:

$$\gamma^{2m} \sum_{ij\in\mathbb{Z}_{2m}} \frac{1}{2^{2m}} \phi_{0i,\mathrm{A}_{\text{🔑}}\mathrm{B}_{\text{🔑}}} \otimes \phi_{0j,\mathrm{A}'_{\text{🔑}}\mathrm{B}'_{\text{🔑}}} \otimes \phi_{ij,\mathrm{A}_{\text{🛡}}\mathrm{B}_{\text{🛡}}}$$

$$\hat{\gamma}^{2m} = \hat{\Phi}^{m}_{\mathrm{A}_{\text{🔑}}\mathrm{B}_{\text{🔑}}} \otimes \hat{\Phi}^{m}_{\mathrm{A}'_{\text{🔑}}\mathrm{B}'_{\text{🔑}}} \otimes \tau^{m}_{\mathrm{A}_{\text{🛡}}\mathrm{B}_{\text{🛡}}}$$

where τ^m is the maximally mixed state.

We now show that using the reversible map is necessary, namely for these states the bound of Lemma 22 is strictly suboptimal, while the bound of Corollary 23 achieves equality. The bound of Lemma 22 computes to

$$\chi_{\text{LOCC}}\left(\left\{\tfrac{1}{2^{2m}}, \phi_{ij}\right\}\right) = m,$$

which is proven using the measurement on the computational basis and the upper bound of [Bad+03, Equation 8].

However, this state is distillable into $2m$ maximally entangled states with just a sequence of unitaries. We recall from Equation (1.10) in Section 1.5, that

$$[\overline{\mathcal{F}} \otimes \mathcal{F}](\phi_{ij}) = \phi_{ji},$$

where $\mathcal{F}$ is the Fourier transform channel. We also recall from Lemma 2 that the *BNOT* channel acts as:

$$\mathcal{BNOT}(\phi_{0j} \otimes \phi_{kl}) = \phi_{0,j-l} \otimes \phi_{kl}.$$

The unitary that distills Φ^{2m} is then obtained by the following sequence (we omit the identity channels):

$$\mathcal{BNOT}_{\mathrm{A}_{\text{🔑}}\mathrm{B}_{\text{🔑}},\mathrm{A}_{\text{🛡}}\mathrm{B}_{\text{🛡}}} \circ [\overline{\mathcal{F}}_{\mathrm{A}_{\text{🛡}}} \otimes \mathcal{F}_{\mathrm{B}_{\text{🛡}}}] \circ \mathcal{BNOT}_{\mathrm{A}'_{\text{🔑}}\mathrm{B}'_{\text{🔑}},\mathrm{A}_{\text{🛡}}\mathrm{B}_{\text{🛡}}}$$

which results in

$$\begin{aligned}
\mathcal{BNOT}_{\mathrm{A}_{\text{🔑}}\mathrm{B}_{\text{🔑}},\mathrm{A}_{\text{🛡}}\mathrm{B}_{\text{🛡}}} \circ [\overline{\mathcal{F}}_{\mathrm{A}_{\text{🛡}}} \otimes \mathcal{F}_{\mathrm{B}_{\text{🛡}}}] \circ \mathcal{BNOT}_{\mathrm{A}'_{\text{🔑}}\mathrm{B}'_{\text{🔑}},\mathrm{A}_{\text{🛡}}\mathrm{B}_{\text{🛡}}} &(\phi_{0,i+k} \otimes \phi_{0,j+l} \otimes \phi_{ij}) \\
&= \mathcal{BNOT}_{\mathrm{A}_{\text{🔑}}\mathrm{B}_{\text{🔑}},\mathrm{A}_{\text{🛡}}\mathrm{B}_{\text{🛡}}} \circ [\overline{\mathcal{F}}_{\mathrm{A}_{\text{🛡}}} \otimes \mathcal{F}_{\mathrm{B}_{\text{🛡}}}](\phi_{0,i+k} \otimes \phi_{0l} \otimes \phi_{ij}) \\
&= \mathcal{BNOT}_{\mathrm{A}_{\text{🔑}}\mathrm{B}_{\text{🔑}},\mathrm{A}_{\text{🛡}}\mathrm{B}_{\text{🛡}}}(\phi_{0,i+k} \otimes \phi_{0l} \otimes \phi_{ji}) \\
&= \phi_{0k} \otimes \phi_{0l} \otimes \phi_{ji},
\end{aligned}$$

and thus there exist a local unitary achieving the transformation

$$\mathcal{BNOT} \circ [\overline{\mathcal{F}} \otimes \mathcal{F}] \circ \mathcal{BNOT}(\mathcal{Z}^k_{\mathrm{B}_{\text{🔑}}} \mathcal{Z}^l_{\mathrm{B}'_{\text{🔑}}}(\gamma^{2m})) = \phi_{0k} \otimes \phi_{0l} \otimes \tau^m$$

thus proving that $E_{D,\mathsf{LO}}(\gamma^{2m}) = 2m$.

We have computed the distillation procedure on $\mathcal{Z}^k_{\mathrm{A}_{\text{🔑}}} \mathcal{Z}^l_{\mathrm{B}'_{\text{🔑}}}(\gamma^{2m})$, because the same unitary now shows that the LO Holevo information of the private states of Corollary 23 achieves the distillable entanglement. Namely, by using the distillation protocol and tracing out the shield, we get by monotonicity

$$\chi_{\mathsf{LO}}\left(\left\{\tfrac{1}{2^{2m}}, \mathcal{Z}^k_{\mathrm{A}_{\text{🔑}}} \mathcal{Z}^l_{\mathrm{B}'_{\text{🔑}}}(\gamma^{2m})\right\}\right) \geqslant \chi_{\mathsf{LO}}\left(\left\{\tfrac{1}{2^{2m}}, \phi_{0k} \otimes \phi_{0l}\right\}\right) = \chi_{\mathsf{LO}}\left(\left\{\tfrac{1}{2^{2m}}, \phi_{k0} \otimes \phi_{l0}\right\}\right) \geqslant 2m,$$

where the value is achieved measuring the computational basis. This bound is now optimal and performs strictly better than Lemma 22.

3.4.2 The Locking private states

The content of this section is unpublished work in collaboration with Māris Ozols and Matthias Christandl.

We have seen with the *BNOT* private states, that the Holevo information can increase when constructing the private state. In this section we show that this can happen even when the shield states are actually separable. However, in this case the separation holds only for one-way LOCC, and we do not have a family of states with asymptotically unbounded separation, but only a specific example. More precisely, we will take the locking basis for two qubits, which is not perfectly distinguishable with $\mathsf{LOCC}_{\rightarrow}$, namely it has $\chi_{\mathsf{LOCC}_{\rightarrow}} < \chi$, and show that the resulting private state is distillable with $\mathsf{LOCC}_{\rightarrow}$. Recall that the locking states from Section 2.2 are defined for $\mathbb{C}^d$:$\mathbb{C}^2$ as

$$\psi_{ij} := \mathcal{F}^j(|i\rangle\langle i|) \otimes |j\rangle\langle j|$$

where $i \in \mathbb{Z}_d$ and $j \in \mathbb{Z}_2$, and they satisfy

$$\chi_{\mathsf{LOCC}_{\rightarrow}(\underline{\mathbb{C}}^d:\mathbb{C}^2)}\left(\left\{\tfrac{1}{2d}, \psi_{ij}\right\}\right) = 1 + \frac{1}{2}\log d.$$

Let now $\mathrm{A}_{\text{🔑}} = \mathrm{B}_{\text{🔑}} = \mathrm{A}_{\text{🛡}} = \mathrm{B}_{\text{🛡}} = \mathbb{C}^2$, and let ψ_{ij} be the locking states in $\mathrm{A}_{\text{🛡}}$:$\mathrm{B}_{\text{🛡}}$. To construct the private state we use the qubit in $\mathrm{A}_{\text{🛡}}$, which is locked, to decide the phase of the maximally entangled state, expecting that it will not be possible to correct it without a message from Bob. The resulting private state is

$$\gamma := \frac{1}{4} \sum_{ij \in \mathbb{Z}_2} \phi_{0i} \otimes \psi_{ij}.$$

Notice that because the qubit at Bob is classical and readily accessible, and the phase/bit flips on the maximally entangled states can be done locally, the above private state is locally equivalent to

$$\frac{1}{4} \sum_{ij \in \mathbb{Z}_2} \phi_{ji} \otimes \psi_{ij}.$$

For a similar reason, the alternative choice

$$\gamma' := \frac{1}{4} \sum_{ij \in \mathbb{Z}_2} \phi_{0i} \otimes \phi_{0j} \otimes \psi_{ij}$$

is locally equivalent to the original choice with an additional distilled maximally entangled state

$$\frac{1}{4}\sum_{ij\in\mathbb{Z}_2} \phi_{0i}\otimes\phi_{00}\otimes\psi_{ij},$$

and thus our initial choice is without loss of generality.

Lemma 24. *The zero-error single-shot distillable entanglement of the Locking private state satisfies:*

$$E^0_{D,\mathsf{LOCC}_\rightarrow}(\gamma)\geqslant 1.$$

Proof. For the purpose we define the "x" and "y" bases of $\mathbb{C}^2$:

$$\begin{aligned}|j_{\mathrm{x}}\rangle :=& \frac{1}{\sqrt{2}}(|0\rangle + (-1)^j\,|1\rangle) = F\,|j\rangle \\ |j_{\mathrm{y}}\rangle :=& \frac{1}{\sqrt{2}}(|0\rangle + i(-1)^j\,|1\rangle),\end{aligned}$$

where we will for now stop using i as and index, so that the i's outside the bra-ket's are the imaginary unit. For these states we have $\langle j|k_{\mathrm{y}}\rangle = \frac{1}{\sqrt{2}}(i)^j(-1)^{jk}$ and $\langle j_{\mathrm{x}}|k_{\mathrm{y}}\rangle = \frac{1}{2}(1+i(-1)^{j+k})$, and thus

$$\begin{aligned}\langle k_{\mathrm{y}}|a\rangle\langle a|l_{\mathrm{y}}\rangle =& \frac{1}{2}(-1)^{ka}(-1)^{la} \\ \langle k_{\mathrm{y}}|a_{\mathrm{x}}\rangle\langle a_{\mathrm{x}}|l_{\mathrm{y}}\rangle =& \frac{1}{4}([1+(-1)^{k+l}] - i(-1)^a[(-1)^k-(-1)^l]) \\ =& \frac{1}{2}(\delta_{kl} - i(-1)^{a+k}[1-\delta_{kl}]).\end{aligned}$$

where the Kronecker deltas have the effect of selecting 1 if k and l are equal, or selecting $-i(-1)^{a+k}$ if they are different. We use these expressions, in the next step, where we define the following orthogonal projectors on $\mathbb{C}^2\otimes\mathbb{C}^2$ at Alice

$$P^j := \sum_{x\in\mathbb{Z}_2} |x\rangle\langle x| \otimes |(x+j)_{\mathrm{y}}\rangle\langle (x+j)_{\mathrm{y}}|\,.$$

and by plugging in the formulas above, we can now compute

$$\begin{aligned}P^j(|k\rangle\langle l|\otimes|a\rangle\langle a|)P^j =& |k\rangle\langle l|\otimes|k_{\mathrm{y}}+j_{\mathrm{y}}\rangle\langle k_{\mathrm{y}}+j_{\mathrm{y}}|a\rangle\langle a|l_{\mathrm{y}}+j_{\mathrm{y}}\rangle\langle l_{\mathrm{y}}+j_{\mathrm{y}}| \\ =& |k\rangle\langle l|\otimes|k_{\mathrm{y}}+j_{\mathrm{y}}\rangle\langle l_{\mathrm{y}}+j_{\mathrm{y}}|\,\frac{1}{2}(-1)^{(k+j)a}(-1)^{(l+j)a} \\ =& \frac{1}{2}(-1)^{ka}\,|k\rangle\langle l|\,(-1)^{la}\otimes|k_{\mathrm{y}}+j_{\mathrm{y}}\rangle\langle l_{\mathrm{y}}+j_{\mathrm{y}}| \\ P^j(|k\rangle\langle l|\otimes|a_{\mathrm{x}}\rangle\langle a_{\mathrm{x}}|)P^j =& |k\rangle\langle l|\otimes|k_{\mathrm{y}}+j_{\mathrm{y}}\rangle\langle k_{\mathrm{y}}+j_{\mathrm{y}}|a_{\mathrm{x}}\rangle\langle a_{\mathrm{x}}|l_{\mathrm{y}}+j_{\mathrm{y}}\rangle\langle l_{\mathrm{y}}+j_{\mathrm{y}}| \\ =& |k\rangle\langle l|\otimes|k_{\mathrm{y}}+j_{\mathrm{y}}\rangle\langle l_{\mathrm{y}}+j_{\mathrm{y}}|\,\frac{1}{2}(\delta_{kl}-i(-1)^{a+k+j}[1-\delta_{kl}]).\end{aligned}$$

This is a piece of the action of P^j on the Locking private state. Indeed, by expanding the Bell states, the Locking private state equals:

$$\begin{aligned}\gamma :=& \frac{1}{8}\sum_{akl\in\mathbb{Z}_2}\Big((-1)^{ka}\,|kk\rangle\langle ll|_{\mathrm{A}_{🔑}\mathrm{B}_{🔑}}\,(-1)^{la}\otimes|a\rangle\langle a|_{\mathrm{A}_{🛡}}\otimes|0\rangle\langle 0|_{\mathrm{B}_{🛡}} + \\ &+(-1)^{ka}\,|kk\rangle\langle ll|_{\mathrm{A}_{🔑}\mathrm{B}_{🔑}}\,(-1)^{la}\otimes|a_{\mathrm{x}}\rangle\langle a_{\mathrm{x}}|_{\mathrm{A}_{🛡}}\otimes|1\rangle\langle 1|_{\mathrm{B}_{🛡}}\Big)\end{aligned}$$

Again we plug in the last expression above and, making heavy use of the effect of the Kronecker deltas, we compute

$$
\begin{aligned}
(P^j \otimes \mathbb{1}_{\mathrm{B}_{\text{🔑}}\mathrm{B}_{\text{🛡}}}) &\left((-1)^{ka} \left|kk\right\rangle\!\left\langle ll\right| (-1)^{la} \otimes |a\rangle\!\langle a| \otimes |0\rangle\!\langle 0|\right) (P^j \otimes \mathbb{1}_{\mathrm{B}_{\text{🔑}}\mathrm{B}_{\text{🛡}}}) \\
&= \frac{1}{2} \left|kk, k_{\mathrm{y}} + j_{\mathrm{y}}\right\rangle\!\left\langle ll, l_{\mathrm{y}} + j_{\mathrm{y}}\right| \otimes |0\rangle\!\langle 0| \\
(P^j \otimes \mathbb{1}_{\mathrm{B}_{\text{🔑}}\mathrm{B}_{\text{🛡}}}) &\left((-1)^{ka} \left|kk\right\rangle\!\left\langle ll\right| (-1)^{la} \otimes |a_{\mathrm{x}}\rangle\!\langle a_{\mathrm{x}}| \otimes |1\rangle\!\langle 1|\right) (P^j \otimes \mathbb{1}_{\mathrm{B}_{\text{🔑}}\mathrm{B}_{\text{🛡}}}) \\
&= \frac{1}{2} \left|kk, k_{\mathrm{y}} + j_{\mathrm{y}}\right\rangle\!\left\langle ll, l_{\mathrm{y}} + j_{\mathrm{y}}\right| \otimes |1\rangle\!\langle 1| \cdot \\
&\quad \cdot \left((-1)^{a(k+l)}\delta_{kl} - i(-1)^{a+k+j}(-1)^{a(k+l)}[1-\delta_{kl}]\right) \\
&= \frac{1}{2} \left|kk, k_{\mathrm{y}} + j_{\mathrm{y}}\right\rangle\!\left\langle ll, l_{\mathrm{y}} + j_{\mathrm{y}}\right| \otimes |1\rangle\!\langle 1| \cdot \\
&\quad \cdot \left(\delta_{kl} - i(-1)^{k+j}[1-\delta_{kl}]\right).
\end{aligned}
$$

We are almost there. The qubit $\left|k_{\mathrm{y}} + j_{\mathrm{y}}\right\rangle\!\left\langle l_{\mathrm{y}} + j_{\mathrm{y}}\right|$ is at Alice, and has the same expression independently of the measurement outcome j. Therefore, Alice can change from the y basis to the computational basis, and then remove k and l controlling a $CNOT$ on her key qubit. This amounts to the following operation

$$
\left|k, k_{\mathrm{y}} + j_{\mathrm{y}}\right\rangle\!\left\langle l, l_{\mathrm{y}} + j_{\mathrm{y}}\right|_{\mathrm{A}_{\text{🔑}}\mathrm{A}_{\text{🛡}}} \to |k, j\rangle\!\langle l, j|_{\mathrm{A}_{\text{🔑}}\mathrm{A}_{\text{🛡}}}
$$

which is local, unitary and independent of the measurement outcome j. We even get $|j\rangle\!\langle j|$ for free at $\mathrm{A}_{\text{🛡}}$, without having recorded the measurement outcome explicitly. We thus have that Alice, with only a local operation, can transform into the following state

$$
\begin{aligned}
\tilde{\gamma} := \frac{1}{8} \sum_{klj \in \mathbb{Z}_2} &\Big(|kk\rangle\!\langle ll|_{\mathrm{A}_{\text{🔑}}\mathrm{B}_{\text{🔑}}} \otimes |j\rangle\!\langle j|_{\mathrm{A}_{\text{🛡}}} \otimes |0\rangle\!\langle 0|_{\mathrm{B}_{\text{🛡}}} + \\
&+ |kk\rangle\!\langle ll|_{\mathrm{A}_{\text{🔑}}\mathrm{B}_{\text{🔑}}} \otimes |j\rangle\!\langle j|_{\mathrm{A}_{\text{🛡}}} \otimes |1\rangle\!\langle 1|_{\mathrm{B}_{\text{🛡}}} \left(\delta_{kl} - i(-1)^{k+j}[1-\delta_{kl}]\right)\Big)
\end{aligned}
$$

(notice that the original locked bit of information a has been summed over, absorbing a factor of $\frac{1}{2}$). Finally, the only remaining thing is a phase gate, which depends on Bob's bit of information. Bob can remove $(-1)^k$ with a controlled phase gate conditioned on his bit $\mathrm{B}_{\text{🛡}}$; this can be done without knowing the measurement outcome j. The last phase depends jointly on Bobs qubit and the measurement outcome, Alice can now send j to Bob, who can then correct the phase. A way to write this precisely is to again, exploit the properties of the Kronecker product and compute

$$
\begin{aligned}
(\delta_{kl} - i(-1)^{k+j}[1-\delta_{kl}]) &= (\delta_{kl} + (-i)^{k+l}(-1)^{k+j}[1-\delta_{kl}]) \\
&= (\delta_{kl} + (i)^k(-1)^k(-i)^l(-1)^{k+j}[1-\delta_{kl}]) \\
&= ((i)^k(-i)^l\delta_{kl} + (i)^k(-i)^l(-1)^j[1-\delta_{kl}]) \\
&= (i)^k(\delta_{kl} + (-1)^j[1-\delta_{kl}])(-i)^l
\end{aligned}
$$

which extract's j-independent the phase flip. To extract the j dependent phase flip we compute

$$
\begin{aligned}
(\delta_{kl} + (-1)^j[1-\delta_{kl}]) &= (((-1)^j)^{k+l}\delta_{kl} + ((-1)^j)^{k+l}[1-\delta_{kl}]) \\
&= ((-1)^j)^{k+l} = (-1)^{jk}(-1)^{jl}.
\end{aligned}
$$

Therefore, we can rewrite $\tilde{\gamma}$ as

$$\begin{aligned}\tilde{\gamma} :=& \frac{1}{8}\sum_{klj\in\mathbb{Z}_2} \Big(|kk\rangle\langle ll| \otimes |j\rangle\langle j| \otimes |0\rangle\langle 0| + \\ & \qquad + (-1)^{jk}(i)^k\, |kk\rangle\langle ll|\, (-i)^l (-1)^{jl} \otimes |j\rangle\langle j| \otimes |1\rangle\langle 1| \Big) \\ =& \frac{1}{8}\sum_{kljb\in\mathbb{Z}_2} (-1)^{jkb}(i)^{kb}\, |kk\rangle\langle ll|\, (-i)^{lb}(-1)^{jlb} \otimes |j\rangle\langle j| \otimes |b\rangle\langle b| \\ =& \frac{1}{8}\sum_{jb\in\mathbb{Z}_2} \Pi^b_{\mathrm{B}_{\text{🔑}}} \phi_{0,jb} \Pi^{-b}_{\mathrm{B}_{\text{🔑}}} \otimes |j\rangle\langle j| \otimes |b\rangle\langle b| \end{aligned}$$

where $\Pi = |0\rangle\langle 0| + i\,|1\rangle\langle 1|$ is the phase gate. Controlling the phase gate and the phase flips using the classical information j and b, Bob can distill

$$\Phi \otimes \sum_{jb\in\mathbb{Z}_2} |j\rangle\langle j| \otimes |b\rangle\langle b|\,.$$

The shields are now independent of the maximally entangled states, and can be traced out, completing the protocol. □

This leads to the following corollary, which shows a separation between the distinguishability of the private state and the distinguishability of the shield states, even when the shield states are separable.

Corollary 25.

$$\chi_{\mathsf{LOCC}_{\rightarrow}(\mathrm{A}_{\text{🔑}}\mathrm{A}_{\text{🛡}}:\mathrm{B}_{\text{🔑}}\mathrm{B}_{\text{🛡}})}\left(\left\{\tfrac{1}{2}, \gamma^{\pm}\right\}\right) = 1$$

Proof. This is a consequence of the distillation protocol only needing to apply phase gates on Bob's key. $\mathcal{Z}^{\pm}_{\mathrm{B}_{\text{🔑}}}$ thus commutes with the distillation and we have

$$\chi_{\mathsf{LOCC}_{\rightarrow}(\mathrm{A}_{\text{🔑}}\mathrm{A}_{\text{🛡}}:\mathrm{B}_{\text{🔑}}\mathrm{B}_{\text{🛡}})}\left(\left\{\tfrac{1}{2}, \gamma^{\pm}\right\}\right) \geqslant \chi_{\mathsf{LOCC}_{\rightarrow}(\mathrm{A}_{\text{🔑}}\mathrm{A}_{\text{🛡}}:\mathrm{B}_{\text{🔑}}\mathrm{B}_{\text{🛡}})}\left(\left\{\tfrac{1}{2}, \Phi^{\pm}\right\}\right) = 1. \qquad (3.17)$$

The Locking private state is irreducible, therefore distillable key and thus distillable entanglement are at most one, giving the upper bound. □

It may be worth remarking that all the known results about the imperfect distinguishability of domino and locking state are single-copy, and we do not know if sampling the source multiple times, and performing operations on multiple copies, achieves asymptotically perfect distinguishability for the locking states. In particular it is not known whether the regularized restricted Holevo information achieves the Holevo information for these ensembles.

Regarding the dimension of the locking states, we do not know if the distillation protocol is generalizable to $\mathbb{C}^d : \mathbb{C}^2$. The protocol seems to use only the interplay between mutually unbiased bases. Since there always exist at least three of them [Sch60], two of them being the computational and the conjugate bases, it is possible that the measurement in the third is enough to achieve the perfect distillation with one-way $\mathsf{LOCC}_{\rightarrow}$ of the locking states in higher dimension.

Finally, we would like to remark that while Bob cannot communicate to Alice, he can still perform operations that simplify the Locking private states locally. This is not the case if we define the Domino private state

$$\gamma_{\mathrm{domino}} := \sum_{ijk\in\mathbb{Z}_2} \phi_{0i} \otimes \phi_{0j} \otimes \phi_{0k} \otimes \psi_{ijk}$$

where ψ_{ijk} are the outer Domino states presented in Section 2.2.5. We leave as an open question whether there exist an LOCC protocol that can perfectly distill the Domino private state. However, the existence of such a protocol would also imply the existence of a one-way LOCC first step that changes the state without destroying the distillable entanglement, and thus it seems that the imperfect discrimination of the Domino states should be more robust than the imperfect discrimination of the locking states.

3.4.3 The Swap private states

We have just seen two cases were the distinguishability is made perfect by constructing private states, thus showing that using the maximally entangled state can be useful and necessary in the distillation of entanglement. Notice that in either case the classical information about the phase was never recovered. Now we will see opposite examples, where the distillable entanglement is provably low, thus showing that the private states themselves are indistinguishable even when acting on many copies. The first such examples are he swap private states γ_S [Hor+05b], defined for each dimension $d = |\mathrm{A}_{\text{🛡}}| = |\mathrm{B}_{\text{🛡}}|$ as the following Bell private bits:

$$\gamma_S = \frac{1}{2}\left(1 + \frac{1}{d}\right)\Phi_+ \otimes \rho_s + \frac{1}{2}\left(1 - \frac{1}{d}\right)\Phi_- \otimes \rho_a \tag{3.18}$$

for each dimension $d > 1$, where ρ_s and ρ_a are the symmetric and anti-symmetric states in $\mathbb{C}^d \otimes \mathbb{C}^d$. In private state form, they are defined by:

$$\sigma = \frac{\mathbb{1}}{d^2} \qquad\qquad T = \mathbb{1}_2 \otimes (|0\rangle\langle 0| \otimes \mathbb{1} + |1\rangle\langle 1| \otimes S)$$

where S is the swap operator. Because the swap is hermitian, it gives the following block form:

$$\gamma_S = \frac{1}{2}\frac{1}{d^2}\begin{bmatrix} \mathbb{1} & 0 & 0 & S \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ S & 0 & 0 & \mathbb{1} \end{bmatrix}.$$

This is a strictly irreducible private state so $K_D(\gamma_S) = 1$. By the logarithmic negativity, we have the following upper bound on the distillable entanglement which vanishes for large enough d, and that we can use in Corollary 23. Thus the Swap private states are data hiding.

Corollary 26.

$$\chi^{\infty}_{\mathsf{L}(\underline{\mathrm{A}}:\mathrm{B})}(\{\tfrac{1}{2}, \gamma_S^{\pm}\}) \leqslant E_{D,\mathsf{L}}(\gamma_S) \leqslant E_N(\gamma_S) = \log\left(1 + \frac{1}{d}\right)$$

with L *a choice between* LOCC$_{\rightarrow}$, LOCC, SEP *and* PPT.

The swap private states display the worst case scenario of indistinguishability, as it was shown that orthogonal states always have a minimal distinguishability even under local operations, that indeed scales inversely with the dimension of the local systems [MWW09].

3.4.4 The Fourier and Flower private states

In general, any unitary matrix can be used to define a private state [Hor+08]. Here we focus only on the following special case:

$$U = \textstyle\sum_{ij} \frac{1}{\sqrt{d}} u_{ij}\, |i\rangle\langle j| \tag{3.19}$$

such that $|u_{ij}| = 1$, an example being the discrete Fourier transform. For each such U we then define the following operators, to be used in the examples to follow:

$$
\begin{aligned}
\mathbb{U} &:= \textstyle\sum_{ij} u_{ij} \, |ii\rangle\langle jj| \\
\mathbb{U}^{\Gamma} &:= \textstyle\sum_{ij} u_{ij} \, |ij\rangle\langle ji|
\end{aligned}
$$

where $(\cdot)^{\Gamma}$ denotes the partial transpose. Notice that $\mathbb{U}^{\Gamma}$ is a unitary and $\frac{\mathbb{U}}{\sqrt{d}}$ is unitary in the maximally correlated subspace, namely

$$
\begin{aligned}
\frac{\mathbb{U}}{\sqrt{d}} \frac{\mathbb{U}^{\dagger}}{\sqrt{d}} &= \frac{\mathbb{U}^{\dagger}}{\sqrt{d}} \frac{\mathbb{U}}{\sqrt{d}} = \mathbb{1}_{\hat{\Phi}} \\
\mathbb{U}^{\Gamma\dagger} \mathbb{U}^{\Gamma} &= \mathbb{U}^{\Gamma} \mathbb{U}^{\Gamma\dagger} = \mathbb{1} \, .
\end{aligned}
$$

The privates bits stemming from these operators are building blocks for the construction of PPT examples.

Fourier

The Fourier private bits $\gamma_{\mathbb{U}^{\Gamma}}$ [Hor+08] are defined for each $d = |\mathrm{A}_{\text{🛡}}| = |\mathrm{B}_{\text{🛡}}|$ and for each unitary U from Equation (3.19) as

$$
\sigma = \frac{\mathbb{1}}{d^2} \qquad\qquad T = \mathbb{1}_2 \otimes (|0\rangle\langle 0| \otimes \mathbb{1} + |1\rangle\langle 1| \otimes \mathbb{U}^{\Gamma})
$$

or in block form:

$$
\gamma_{\mathbb{U}^{\Gamma}} = \frac{1}{2}\frac{1}{d^2} \begin{bmatrix} \mathbb{1} & 0 & 0 & \mathbb{U}^{\Gamma} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \mathbb{U}^{\Gamma\dagger} & 0 & 0 & \mathbb{1} \end{bmatrix} . \tag{3.20}
$$

Notice that in general these are not Bell private states because $\mathbb{U}^{\Gamma}$ is in general not hermitian. U is usually taken to be the discrete Fourier transform, thus the name.

These are also strictly irreducible private states, thus $K_D(\gamma_{\mathbb{U}^{\Gamma}}) = 1$, and again we have an upper bound on distillable entanglement via the logarithmic negativity, showing with Corollary 23 that they are also data hiding.

Corollary 27.

$$
\chi^{\infty}_{\mathsf{L}(\underline{\mathrm{A}}:\mathrm{B})}\left(\left\{\tfrac{1}{2}, \gamma^{\pm}_{\mathbb{U}^{\Gamma}}\right\}\right) \leqslant E_{D,\mathsf{L}}(\gamma_{\mathbb{U}^{\Gamma}}) \leqslant E_N(\gamma_{\mathbb{U}^{\Gamma}}) = \log\left(1 + \frac{1}{\sqrt{d}}\right).
$$

with L *a choice between* $\mathsf{LOCC}_{\rightarrow}$, LOCC, SEP *and* PPT.

We will see in the next section that when choosing the orthogonal shield states at random, the above scaling is indeed the resulting one for separable measurements with high probability. However, the Fourier private bits are indistinguishable under PPT measurements, while picking the shield states at random will result in private bits distinguishable under PPT measurements (but we will only be able to prove this in the single-copy regime).

Flower

The Flower private bits $\gamma_{\mathbb{U}}$ [Hor+08] similarly are defined for each $d = |\mathrm{A}_{\text{🛡}}| = |\mathrm{B}_{\text{🛡}}|$ and for each unitary U as:

$$
\sigma = \frac{\mathbb{1}_{\hat{\Phi}}}{d} \qquad\qquad T = \mathbb{1}_2 \otimes (|0\rangle\langle 0| \otimes \mathbb{1} + |1\rangle\langle 1| \otimes \mathbb{U})
$$

or in block form:

$$\gamma_{\mathbb{U}} = \frac{1}{2}\frac{1}{d}\begin{bmatrix} \mathbb{1}_{\hat{\Phi}} & 0 & 0 & \frac{\mathbb{U}}{\sqrt{d}} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \frac{\mathbb{U}^{\dagger}}{\sqrt{d}} & 0 & 0 & \mathbb{1}_{\hat{\Phi}} \end{bmatrix}. \tag{3.21}$$

Again, these are not Bell private states in general and they are strictly irreducible, because the key-attacked state is separable. However, for the same reason that makes the logarithmic negativity of the Fourier private states small, the logarithmic negativity of the Flower private states becomes large and thus it cannot be used to find a meaningful bound on the distillable entanglement:

$$E_N(\gamma_{\mathbb{U}}) = \log\left(1+\sqrt{d}\right).$$

On the other hand, the distillable entanglement is indeed high. We can actually compute it explicitly via the hashing bound, because $\gamma_{\mathbb{U}}$ has support only on the maximally correlated subspace of $\mathrm{A}_{\text{🔑}}\mathrm{B}_{\text{🔑}}\mathrm{A}_{\text{🛡}}\mathrm{B}_{\text{🛡}}$ [HH04]. Since $\frac{\mathbb{U}}{\sqrt{d}}$ is a unitary in the maximally correlated subspace, it can be diagonalized in this subspace with phases eigenvalues. In short, we find

$$I(\mathrm{A}_{\text{🔑}}\mathrm{A}_{\text{🛡}}\rangle\mathrm{B}_{\text{🔑}}\mathrm{B}_{\text{🛡}}) = H(\mathrm{B}_{\text{🔑}}\mathrm{B}_{\text{🛡}})_{\gamma_{\mathbb{U}}} - H(\mathrm{A}_{\text{🔑}}\mathrm{B}_{\text{🔑}}\mathrm{A}_{\text{🛡}}\mathrm{B}_{\text{🛡}})_{\gamma_{\mathbb{U}}} = (1+\log d) - \log d.$$

Therefore, for the class of Flower private states:

$$E_D(\gamma_{\mathbb{U}}) = K_D(\gamma_{\mathbb{U}}) = 1. \tag{3.22}$$

While they might not seem interesting, they are used in connection with the Fourier private states to create states close to private states that are also PPT, as we will see below.

3.4.5 The PPT Fourier private states

The PPT Fourier private states $\tilde{\gamma}_{\mathbb{U}^{\Gamma}}$ [Bäu+15] are not exact private states, they are approximate private bits that can be made arbitrarily close to the Fourier private states while still being PPT. The class of PPT Fourier private states defines (for each $d = |\mathrm{A}_{\text{🛡}}| = |\mathrm{B}_{\text{🛡}}|$ and for each U from Equation (3.19)):

$$\tilde{\gamma}_{\mathbb{U}^{\Gamma}} = \frac{1}{1+\frac{1}{\sqrt{d}}}\left(\gamma_{\mathbb{U}^{\Gamma}} + \frac{1}{\sqrt{d}} X_{\mathrm{A}_{\text{🔑}}}\hat{\gamma}_{\mathbb{U}} X_{\mathrm{A}_{\text{🔑}}}\right) \tag{3.23}$$

where $X_{\mathrm{A}_{\text{🔑}}}$ is the bit flip on $\mathrm{A}_{\text{🔑}}$, and its function is to move the key-attacked state $\hat{\gamma}_{\mathbb{U}}$ to the orthogonal subspace. Namely, in block form:

$$\tilde{\gamma}_{\mathbb{U}^{\Gamma}} = \frac{1}{2}\frac{1}{1+\frac{1}{\sqrt{d}}}\begin{bmatrix} \frac{\mathbb{1}}{d^2} & 0 & 0 & \frac{\mathbb{U}^{\Gamma}}{d^2} \\ 0 & \frac{1}{\sqrt{d}}\frac{\mathbb{1}_{\hat{\Phi}}}{d} & 0 & 0 \\ 0 & 0 & \frac{1}{\sqrt{d}}\frac{\mathbb{1}_{\hat{\Phi}}}{d} & 0 \\ \frac{\mathbb{U}^{\Gamma\dagger}}{d^2} & 0 & 0 & \frac{\mathbb{1}}{d^2} \end{bmatrix}.$$

One can check that this $O(1/\sqrt{d})$ noise is just enough to make the states PPT, and remarkably, the amount of noise needed in the mixture goes to zero for large d. The PPT Fourier private states are engineered to become close to the set of separable states after partial transposition. Indeed, since $\mathbb{1}$ and $\mathbb{1}_{\hat{\Phi}}$ are PPT invariant, we find

$$\tilde{\gamma}^{\Gamma}_{\mathbb{U}^{\Gamma}} = \frac{1}{2}\frac{1}{1+\frac{1}{\sqrt{d}}}\begin{bmatrix} \frac{\mathbb{1}}{d^2} & 0 & 0 & 0 \\ 0 & \frac{1}{\sqrt{d}}\frac{\mathbb{1}_{\hat{\Phi}}}{d} & \frac{1}{\sqrt{d}}\frac{\mathbb{U}}{d\sqrt{d}} & 0 \\ 0 & \frac{1}{\sqrt{d}}\frac{\mathbb{U}^{\dagger}}{d\sqrt{d}} & \frac{1}{\sqrt{d}}\frac{\mathbb{1}_{\hat{\Phi}}}{d} & 0 \\ 0 & 0 & 0 & \frac{\mathbb{1}}{d^2} \end{bmatrix}$$

and thus:

$$\tilde{\gamma}^{\Gamma}_{\mathbb{U}^{\Gamma}} = \frac{1}{1+\frac{1}{\sqrt{d}}}\left(\hat{\gamma}_{\mathbb{U}^{\Gamma}} + \frac{1}{\sqrt{d}} X_{\mathrm{A}_{\text{🔑}}} \gamma_{\mathbb{U}} X_{\mathrm{A}_{\text{🔑}}}\right)$$

which is suddenly mostly a separable key-attacked state with a vanishing mixture of a Flower private state.

Since the PPT Fourier private states are PPT, their distillable entanglement is exactly zero, but we cannot use this because the bounds on the Holevo information do not apply. However, $\tilde{\gamma}_{\mathbb{U}^{\Gamma}}$ is obtained by adding separable noise to $\gamma_{\mathbb{U}^{\Gamma}}$, and thus by monotonicity, their Holevo information must be small. Namely

Corollary 28.

$$\chi^{\infty}_{\mathsf{L}(\underline{\mathrm{A}}:\mathrm{B})}\left(\left\{\tfrac{1}{2}, \tilde{\gamma}^{\pm}_{\mathbb{U}^{\Gamma}}\right\}\right) \leqslant \chi^{\infty}_{\mathsf{L}(\underline{\mathrm{A}}:\mathrm{B})}\left(\left\{\tfrac{1}{2}, \gamma^{\pm}_{\mathbb{U}^{\Gamma}}\right\}\right) \leqslant E_{D,\mathsf{L}}(\gamma_{\mathbb{U}^{\Gamma}}) \leqslant \log\left(1+\frac{1}{\sqrt{d}}\right).$$

with L *a choice between* $\mathsf{LOCC}_{\to}$, LOCC, SEP *and* PPT.

3.4.6 The PPT invariant private states

The PPT invariant private states $\tilde{\gamma}_{\Gamma}$ [Hor+08] are another class of approximate private bits that are PPT invariant, namely they satisfy $\tilde{\gamma}_{\Gamma}^{\Gamma} = \tilde{\gamma}_{\Gamma}$ By substituting in Equation (3.23) the key-attacked state with the Flower private state, the expression becomes PPT invariant. Namely, the class of PPT invariant private states defines

$$\tilde{\gamma}_{\Gamma} = \frac{1}{1+\frac{1}{\sqrt{d}}}\left(\gamma_{\mathbb{U}^{\Gamma}} + \frac{1}{\sqrt{d}} X_{\mathrm{A}_{\text{🔑}}} \gamma_{\mathbb{U}} X^{\dagger}_{\mathrm{A}_{\text{🔑}}}\right) = (\tilde{\gamma}_{\Gamma})^{\Gamma}\,. \tag{3.24}$$

with block form

$$\tilde{\gamma}_{\Gamma} = \frac{1}{2}\frac{1}{1+\frac{1}{\sqrt{d}}}\begin{bmatrix} \frac{\mathbb{1}}{d^2} & 0 & 0 & \frac{\mathbb{U}^{\Gamma}}{d^2} \\ 0 & \frac{1}{\sqrt{d}}\frac{\mathbb{1}_{\hat{\Phi}}}{d} & \frac{1}{\sqrt{d}}\frac{\mathbb{U}}{d\sqrt{d}} & 0 \\ 0 & \frac{1}{\sqrt{d}}\frac{\mathbb{U}^{\dagger}}{d\sqrt{d}} & \frac{1}{\sqrt{d}}\frac{\mathbb{1}_{\hat{\Phi}}}{d} & 0 \\ \frac{{\mathbb{U}^{\Gamma}}^{\dagger}}{d^2} & 0 & 0 & \frac{\mathbb{1}}{d^2} \end{bmatrix}$$

which is manifestly PPT invariant.

Similarly to the Fourier private state, we cannot use Corollary 23 directly, but a bound can still be computed combining monotonicity and the bound in terms of the relative entropy from PPT states. To do this we introduce a qubit system H and define on system $\mathrm{A}_{\text{🔑}}\mathrm{A}_{\text{🛡}}{:}\mathrm{B}_{\text{🔑}}\mathrm{B}_{\text{🛡}}\mathrm{H}$ the following private bits:

$$\gamma_{\Gamma} = \frac{1}{1+\frac{1}{\sqrt{d}}}\left(\gamma_{\mathbb{U}^{\Gamma}} \otimes |0\rangle\langle 0| + \frac{1}{\sqrt{d}}\gamma_{\mathbb{U}} \otimes |1\rangle\langle 1|\right),$$

and the following PPT state:

$$\tilde{\gamma}_{\Gamma} = \frac{1}{1+\frac{1}{\sqrt{d}}}\left(\tilde{\gamma}_{\mathbb{U}^{\Gamma}} \otimes |0\rangle\langle 0| + \frac{1}{\sqrt{d}}\hat{\gamma}_{\mathbb{U}} \otimes |1\rangle\langle 1|\right).$$

Who holds the additional qubit is irrelevant, but by making it part of the shield it is straightforward to show that γ_{Γ} is actually a strictly irreducible private state. One can obtain $\tilde{\gamma}_{\Gamma}$ from γ_{Γ} with local operations: use the additional qubit to locally bit flip the key of $\gamma_{\mathbb{U}} \otimes |1\rangle\langle 1|$ but not of $\gamma_{\mathbb{U}^{\Gamma}} \otimes |0\rangle\langle 0|$, then trace out H. Furthermore, $\tilde{\gamma}_{\Gamma}$ is clearly PPT, since it is mixture of the PPT states $\tilde{\gamma}_{\mathbb{U}^{\Gamma}}$ and $\hat{\gamma}_{\mathbb{U}}$.

Lemma 29.

$$\chi^{\infty}_{\mathsf{L}(\underline{\mathrm{A}}:\mathrm{B})}\left(\left\{\tfrac{1}{2},\tilde{\gamma}_{\Gamma}^{\pm}\right\}\right) \leqslant \chi^{\infty}_{\mathsf{L}(\underline{\mathrm{A}}:\mathrm{B})}\left(\left\{\tfrac{1}{2},\gamma_{\Gamma}^{\pm}\right\}\right) \leqslant E_{D,\mathsf{L}}(\gamma_{\Gamma}) \leqslant \frac{1+\log e}{1+\sqrt{d}}.$$

with L *a choice between* $\mathsf{LOCC}_{\rightarrow}$, LOCC, SEP *and* PPT.

Proof. We find that γ_{Γ} is close to $\tilde{\gamma}_{\Gamma}$:

$$\begin{aligned} D(\gamma_{\Gamma}\|\tilde{\gamma}_{\Gamma}) &= \frac{1}{1+\frac{1}{\sqrt{d}}}\left(D(\gamma_{\mathbb{U}^{\Gamma}}\|\tilde{\gamma}_{\mathbb{U}^{\Gamma}}) + \frac{1}{\sqrt{d}}D(\gamma_{\mathbb{U}}\|\hat{\gamma}_{\mathbb{U}})\right) \\ &= \frac{1}{1+\frac{1}{\sqrt{d}}}\left(\log\left(1+\frac{1}{\sqrt{d}}\right)+\frac{1}{\sqrt{d}}\right) \\ &\leqslant \frac{\sqrt{d}}{1+\sqrt{d}}\left(\frac{\log e}{\sqrt{d}}+\frac{1}{\sqrt{d}}\right) \\ &= \frac{1+\log e}{1+\sqrt{d}}. \end{aligned} \tag{3.25}$$

Thus by monotonicity under local operations we have

$$D(\tilde{\gamma}_{\Gamma}\|\mathcal{P}(\mathrm{A}_{\text{🔑}}\mathrm{A}_{\text{🛡}}{:}\mathrm{B}_{\text{🔑}}\mathrm{B}_{\text{🛡}})) \leqslant D(\gamma_{\Gamma}\|\mathcal{P}(\mathrm{A}_{\text{🔑}}\mathrm{A}_{\text{🛡}}{:}\mathrm{B}_{\text{🔑}}\mathrm{B}_{\text{🛡}}\mathrm{H})) \leqslant D(\gamma_{\Gamma}\|\tilde{\gamma}_{\Gamma}) \leqslant \frac{1+\log e}{1+\sqrt{d}}$$

It was shown in [Aud+02] that this is an upper bound on the PPT distillable entanglement, namely $E_{D,\mathsf{PPT}}(\rho) \leqslant D(\rho\|\mathcal{P})$. Together with Corollary 23, this ends the proof. □

3.5 Random shield states

The content of this section comes from [2].

The study of random states with probabilistic tools and high dimensional analysis has in recent years significantly advanced our understanding of entanglement [HLW06; ASY14; AL14]. In this section, we use such techniques in order to construct bipartite quantum states that exhibit a large gap between their key and their distinguishability. In order to do so, we follow the prescription of [AL14] to choose the shield states at random and in high dimension, which with high probability produces states that are indistinguishable under separable measurements but distinguishable under PPT measurements. This property is preserved when constructing the private states. The result are private states that are distillable under PPT operations, but that should behave like undistillable states under separable operations. Because they are distillable under PPT operations all the computable upper bounds on the distillable entanglement will be useless, because, as we mentioned, they are actually upper bounds on the distillable entanglement under PPT operations. Nonetheless, proving their indistinguishability under separable measurements will find applications in the next chapter.

3.5.1 Random states

Construction 30 (Random orthogonal states [AL14, Section 6.1]). *Let* $|\mathrm{A}_{\text{🛡}}| = |\mathrm{B}_{\text{🛡}}| = d$ *and without loss of generality assume that* d *is even, and let* P *be an orthogonal projector on some fixed* $d^2/2$*-dimensional subspace of* $\mathrm{A}_{\text{🛡}}\mathrm{B}_{\text{🛡}}$*. Define first the two following orthogonal states:*

$$\bar{\sigma}^{+} := \frac{P}{\operatorname{tr} P} \qquad\qquad \bar{\sigma}^{-} := \frac{P^{\perp}}{\operatorname{tr} P^{\perp}}.$$

Then, let U be a Haar-distributed random unitary on $\mathrm{A}_{\text{🛡}}\mathrm{B}_{\text{🛡}}$, and define the two following random orthogonal states:

$$\sigma^{\pm} := U\bar{\sigma}^{\pm}U^{\dagger}.$$

It was proved in [AL14, Section 6.1] that such random orthogonal states $\sigma^{\pm}$ have interesting data-hiding properties. More precisely, with high probability separable (and thus LOCC) measurements almost do not distinguish them, while PPT ones do. So our goal is now to construct random private states out of them, and show that they exhibit some similar features. However random states are still highly entangled states. We recall the following bound from Proposition 5. For any Hermitian operator K on $\mathrm{A}'\mathrm{B}'$ and any state ρ on AB, we have

$$\|\rho \otimes K\|_{\mathsf{SEP}(\underline{\mathrm{A}}\underline{\mathrm{A}}':\underline{\mathrm{B}}\underline{\mathrm{B}}')} \leqslant (2\mathcal{R}_{\mathrm{A:B}}(\rho)+1)\|K\|_{\mathsf{SEP}(\underline{\mathrm{A}}':\underline{\mathrm{B}}')},$$

To bound the increase in distinguishability provided by using random states as a resource for distinguishability, we estimate their robustness in the next lemma. Again this will need to wait until the next chapter to find its application.

Lemma 31. *Let σ be a random state on $\mathrm{A}_{\text{🛡}}\mathrm{B}_{\text{🛡}}$ as in Construction 30 and let $\tau = \mathbb{1}/d^2$ be the maximally mixed state on $\mathrm{A}_{\text{🛡}}\mathrm{B}_{\text{🛡}}$. Namely, let $\sigma = UPU^{\dagger}/(d^2/2)$ where P is the projector on a $d^2/2$-dimensional subspace of $\mathrm{A}_{\text{🛡}}\mathrm{B}_{\text{🛡}}$ and U is a Haar-distributed random unitary on $\mathrm{A}_{\text{🛡}}\mathrm{B}_{\text{🛡}}$. Then*

$$\mathbf{P}\Big(\mathcal{R}_{\mathrm{A}_{\text{🛡}}:\mathrm{B}_{\text{🛡}}}(\sigma\|\tau) \leqslant C\sqrt{d}\log d\Big) \geqslant 1-e^{-c_0 d^3\log^2 d}$$

and consequently

$$\mathbf{P}\Big(\mathcal{R}_{\mathrm{A}_{\text{🛡}}:\mathrm{B}_{\text{🛡}}}(\sigma) \leqslant C\sqrt{d}\log d\Big) \geqslant 1-e^{-c_0 d^3\log^2 d}.$$

where $C, c_0 > 0$ are universal constants (independent of d).

Proof. The second claim follows from the first by the definition in Equation (2.7), so we only need to prove the upper bound estimate on $\mathcal{R}_{\mathrm{A}_{\text{🛡}}:\mathrm{B}_{\text{🛡}}}(\sigma\|\tau)$. Let $\mathcal{S} \equiv \mathcal{S}(\mathrm{A}_{\text{🛡}}:\mathrm{B}_{\text{🛡}})$ and $\mathcal{R} \equiv \mathcal{R}_{\mathrm{A}_{\text{🛡}}:\mathrm{B}_{\text{🛡}}}$.

We use the same notation as in [ASY14] and define $\mathcal{S}_0 = \mathcal{S} - \tau$ as the set of separable states translated to the subspace of traceless Hermitian operators with the maximally mixed state at the origin. From the definition we have that for any state ϱ

$$\begin{aligned}\mathcal{R}(\varrho\|\tau) &= \inf\left\{s : \frac{1}{1+s}(\varrho + s\,\tau) \in \mathcal{S}\right\} \\ &= \inf\left\{s : \frac{1}{1+s}(\varrho - \tau) \in \mathcal{S}_0\right\}. \end{aligned} \tag{3.26}$$

Let us, with some abuse of notation, denote by $\|\cdot\|_{\mathcal{S}_0}$ the gauge of $\mathcal{S}_0$ (it is not homogeneous because $\mathcal{S}_0$ is not symmetric, and hence is not actually a norm). From Equation (3.26), we thus have that if ϱ is entangled

$$\|\varrho - \tau\|_{\mathcal{S}_0} = \mathcal{R}(\varrho\|\tau) + 1 \tag{3.27}$$

(if ϱ is separable this does not hold, as is the case for $\varrho = \tau$).

Let $\bar{\sigma} = P/\operatorname{tr} P = 2P/d^2$, and let us introduce the notation $\varrho_0 \equiv \varrho - \tau$ for any state ϱ. Notice that $\sigma_0 = U\bar{\sigma}_0U^{\dagger}$. We have reduced the problem of estimating $\mathcal{R}(\sigma\|\tau)$ to the problem of estimating $\|\sigma_0\|_{\mathcal{S}_0}$ and the statement to prove is thus

$$\mathbf{P}\left(\left\|U\bar{\sigma}_0U^{\dagger}\right\|_{\mathcal{S}_0} \leqslant C\sqrt{d}\log d\right) \geqslant 1-e^{-c_0 d^3\log^2 d}\,.$$

For this purpose, we first compute the expectation value $\mathbf{E}\left\|U\bar{\sigma}_0U^\dagger\right\|_{\mathcal{S}_0}$ over the random variable U. Then we estimate the Lipschitz constant of $\left\|U\bar{\sigma}_0U^\dagger\right\|_{\mathcal{S}_0}$ as a function of U and use it to argue that being close to the expected value happens with high probability. Notice that $\|X\|_{\mathcal{S}_0}$ is not unitary invariant, however the function $\mathbf{E}\left\|UXU^\dagger\right\|_{\mathcal{S}_0}$ is unitary invariant on X, and is still convex.

Let us compute the expectation value $\mathbf{E}\left\|U\bar{\sigma}_0U^\dagger\right\|_{\mathcal{S}_0}$. With minor modifications, we know from [AL14, Lemma 6] that for any unitary-invariant convex function g of any traceless Hermitian operators X and Y on $\mathbb{C}^d \otimes \mathbb{C}^d$, we have[5]

$$\frac{g(X)}{\|X\|_\infty} \leqslant 2d^2\frac{g(Y)}{\|Y\|_1}, \tag{3.28}$$

which applied twice leads to

$$\frac{1}{2d^2}\frac{\|X\|_1}{\|Y\|_\infty} \leqslant \frac{g(X)}{g(Y)} \leqslant 2d^2\frac{\|X\|_\infty}{\|Y\|_1}. \tag{3.29}$$

Now, we let $Y = \bar{\sigma}_0$ for which $\|\bar{\sigma}_0\|_1 = 1$ and $\|\bar{\sigma}\|_\infty = 1/d^2$:

$$\frac{1}{2}\|X\|_1 \leqslant \frac{g(X)}{g(\bar{\sigma}_0)} \leqslant 2d^2\|X\|_\infty.$$

Then we let $g(X) = \mathbf{E}\left\|UXU^\dagger\right\|_{\mathcal{S}_0}$, which is unitary invariant by construction and convex by the convexity of $\|X\|_{\mathcal{S}_0}$:

$$\frac{1}{2}\|X\|_1 \leqslant \frac{\mathbf{E}\left\|UXU^\dagger\right\|_{\mathcal{S}_0}}{\mathbf{E}\left\|U\bar{\sigma}_0U^\dagger\right\|_{\mathcal{S}_0}} \leqslant 2d^2\|X\|_\infty.$$

We now let X be a Gaussian vector G on the traceless Hermitian operators (gaussian unitary ensemble) on $\mathbb{C}^d \otimes \mathbb{C}^d$. This makes $\mathbf{E}\left\|UGU^\dagger\right\|_{\mathcal{S}_0} = \mathbf{E}\|G\|_{\mathcal{S}_0}$. We then take the expectation values over the remaining random variable G on each side of the inequalities and get

$$\frac{1}{4}\,\mathbf{E}\|G\|_1 \leqslant \frac{\mathbf{E}\|G\|_{\mathcal{S}_0}}{\mathbf{E}\left\|U\bar{\sigma}_0U^\dagger\right\|_{\mathcal{S}_0}} \leqslant 2d^2\,\mathbf{E}\|G\|_\infty.$$

For such a Gaussian random matrix, it is well know that $\mathbf{E}\|G\|_1 \sim d^3$ and $\mathbf{E}\|G\|_\infty \sim d$, where with "$\sim$" we denote having the same order. This proves

$$\mathbf{E}\left\|U\bar{\sigma}_0U^\dagger\right\|_{\mathcal{S}_0} \sim \mathbf{E}\|G\|_{\mathcal{S}_0}/d^3. \tag{3.30}$$

In particular, we know from [ASY14, Section 4] that $\mathbf{E}\|G\|_{\mathcal{S}_0}$ is at most of order $d^{7/2}\log d$ and therefore there exists a universal constant $C > 0$ such that

$$\mathbf{E}\|\sigma_0\|_{\mathcal{S}_0} = \mathbf{E}\left\|U\bar{\sigma}_0U^\dagger\right\|_{\mathcal{S}_0} \leqslant C\sqrt{d}\log d. \tag{3.31}$$

Now we have to show that this average behaviour is generic for large d, because $f(U) := \left\|U\bar{\sigma}_0U^\dagger\right\|_{\mathcal{S}_0}$ is regular enough. We claim that f is 4-Lipschitz (in the Euclidean

[5] The original [AL14, Lemma 6] is stated only for permutation invariant norms. However, what is proven in the proof is the more general statement that if x and y are zero-sum vectors in $\mathbb{R}^d$, then $x/\|x\|_\infty \prec 2d\,y/\|y\|_1$ where $\prec$ denotes majorisation. A direct application of [Bha13, Theorem II.3.3] then proves [AL14, Lemma 6] and more generally $g(x)/\|x\|_\infty \leqslant 2d\,g(y)/\|y\|_1$ for all permutation-invariant convex functions g. Applying the latter to the spectrum of traceless Hermitian operators for a unitary invariant function gives us Equation (3.28).

norm). Indeed, for any unitaries U, V on $\mathbb{C}^d \otimes \mathbb{C}^d$, we have by the triangle inequality for $\|\cdot\|_{\mathcal{S}_0}$

$$\begin{aligned}
|f(U) - f(V)| &= \left| \left\| U\bar{\sigma}_0 U^\dagger \right\|_{\mathcal{S}_0} - \left\| V\bar{\sigma}_0 V^\dagger \right\|_{\mathcal{S}_0} \right| \\
&\leqslant \left\| U\bar{\sigma}_0 U^\dagger - V\bar{\sigma}_0 V^\dagger \right\|_{\mathcal{S}_0} \\
&= \left\| U\bar{\sigma} U^\dagger - V\bar{\sigma} V^\dagger \right\|_{\mathcal{S}_0} \\
&= \frac{2}{d^2} \left\| UPU^\dagger - VPV^\dagger \right\|_{\mathcal{S}_0}.
\end{aligned}$$

It was proven in [GB02] that

$$\frac{1}{d\sqrt{d^2-1}} B_2 \subseteq \mathcal{S}_0,$$

which implies

$$\|\cdot\|_{\mathcal{S}_0} \leqslant d\sqrt{d^2-1} \|\cdot\|_2 \leqslant d^2 \|\cdot\|_2 .$$

Therefore we get

$$\begin{aligned}
|f(U) - f(V)| &\leqslant 2 \left\| UPU^\dagger - VPV^\dagger \right\|_2 \\
&\leqslant 2 \left\| (U-V)PU^\dagger \right\|_2 + 2 \left\| VP(U-V)^\dagger \right\|_2 \\
&= 4 \|(U-V)P\|_2 \\
&\leqslant 4 \|U-V\|_2 \|P\|_\infty \\
&= 4 \|U-V\|_2 .
\end{aligned}$$

where the first inequality is by the triangle inequality and the last inequality is by Hölder inequality $\|XY\|_2 \leqslant \|X\|_2 \|Y\|_\infty$.

Now, we know from [MM+13, Corollary 17] that any L-Lipschitz function g on the unitaries on $\mathbb{C}^D$ (equipped with the Euclidean metric) satisfies the concentration estimate: if U is a Haar-distributed unitary on $\mathbb{C}^D$, then for all $\varepsilon > 0$, $\mathbf{P}(g(U) > \mathbf{E}\, g + \varepsilon) \leqslant e^{-c_0 D \varepsilon^2 / L^2}$, where $c_0 > 0$ is a universal constant. Combining the above estimate on the Lipschitz constant of f with the estimate on its expected value from Equation (3.31), we thus get for all $\varepsilon > 0$

$$\begin{aligned}
\mathbf{P}\Big(f(U) > C\sqrt{d} \log d + \varepsilon \Big) &\leqslant \mathbf{P}(f(U) > \mathbf{E}\, f + \varepsilon) \\
&\leqslant e^{-c_0' d^2 \varepsilon^2}.
\end{aligned}$$

The advertised result follows from choosing $\varepsilon = C\sqrt{d} \log d$ in the above deviation probability, and combining it with Equation (3.27) . □

As we would expect this holds for both random states simultaneously.

Corollary 32. *Let $\sigma^\pm$ be random orthogonal states on $\mathrm{A}_{\text{🛡}}\mathrm{B}_{\text{🛡}}$ as in Construction 30. Then*

$$\mathbf{P}\Big(\max \left\{ \mathcal{R}_{\mathrm{A}_{\text{🛡}}:\mathrm{B}_{\text{🛡}}}(\sigma^+), \mathcal{R}_{\mathrm{A}_{\text{🛡}}:\mathrm{B}_{\text{🛡}}}(\sigma^-) \right\} \leqslant C\sqrt{d} \log d \Big) \geqslant 1 - e^{-c_0 d^3 \log^2 d}.$$

where $C, c_0 > 0$ are universal constants.

Proof. This is a direct consequence of Lemma 31. By the union bound, we have

$$\begin{aligned}
&\mathbf{P}\Big(\mathcal{R}(\sigma^+) \geqslant C\sqrt{d} \log d \text{ or } \mathcal{R}(\sigma^-) \geqslant C\sqrt{d} \log d \Big) \\
&\leqslant \mathbf{P}\Big(\mathcal{R}(\sigma^+) \geqslant C\sqrt{d} \log d \Big) + \mathbf{P}\Big(\mathcal{R}(\sigma^-) \geqslant C\sqrt{d} \log d \Big) \\
&\leqslant 2e^{-c_0 d^3 \log^2 d}
\end{aligned}$$

□

3.5.2 Random private states

We will now use the random orthogonal states just introduced to construct a Bell privat bit. Together with the single-shot results, in particular Lemmas 20 and 21, we show that with the shield states picked at random, the resulting private state will be data hiding under separable measurements, but distinguishable under PPT measurements. In particular, the distinguishability under PPT measurements of the shield implies that the resulting random private state is distillable with PPT operations (by just performing the measurement and correcting the phase). Since the PPT distillable entanglement is upper bounded by the distance from the PPT set, the constructed private states will be far from the set of PPT states. For these random private states we thus have no useful upper bound on the LOCC distillable entanglement, as all the upper bound we know are upper bounds on the PPT one. It is interesting that this is a generic behaviour, and it will be particularly interesting when we will still be able to show a repeater bound under the asymptotic continuity conjecture.

Construction 33. *Let $|\mathrm{A}| = |\mathrm{B}| = 2$ and let $\sigma^{\pm}$ be two random orthogonal states on $\mathrm{A}_{\text{🛡}}\mathrm{B}_{\text{🛡}}$ as in Construction 30. We define a random private state γ on $\mathrm{A}\mathrm{A}_{\text{🛡}}\mathrm{B}\mathrm{B}_{\text{🛡}}$ to be a private state as in Section 3.2.1 with random states $\sigma^{\pm}$ as shield states, namely*

$$\gamma_{\mathrm{A}\mathrm{A}_{\text{🛡}}\mathrm{B}\mathrm{B}_{\text{🛡}}} = \frac{1}{2}\left(\Phi^{+}_{\mathrm{A}_{\text{🔑}}\mathrm{B}_{\text{🔑}}} \otimes \sigma^{+}_{\mathrm{A}_{\text{🛡}}\mathrm{B}_{\text{🛡}}} + \Phi^{-}_{\mathrm{A}_{\text{🔑}}\mathrm{B}_{\text{🔑}}} \otimes \sigma^{-}_{\mathrm{A}_{\text{🛡}}\mathrm{B}_{\text{🛡}}}\right).$$

Observe that, for such a construction the key-attacked state is always separable. Indeed, since $P + P^{\perp} = \mathbb{1}$, we simply have $(\sigma^{+} + \sigma^{-})/2 = \mathbb{1}/d^2$ and therefore $\hat{\gamma} = (\Phi^{+} + \Phi^{-})/2 \otimes \mathbb{1}/d^2$. These are thus strictly irreducible private states, and the distillable key is thus exactly one bit [Hor+09]. Finally, note that $\sigma^{\pm}$ are orthogonal by construction but, as we already know from [AL14, Section 6.1], this orthogonality is completely hidden to local observers. The results from Section 3.2 now directly imply that this is true also for the constructed private states in the limit for large shield sizes.

Theorem 34. *Let γ be a random private state on $\mathrm{A}\mathrm{A}_{\text{🛡}}\mathrm{B}\mathrm{B}_{\text{🛡}}$ as defined by Construction 33. Then,*

$$\mathbf{P}\left(\|\gamma - \mathcal{S}(\mathrm{A}\mathrm{A}_{\text{🛡}}{:}\mathrm{B}\mathrm{B}_{\text{🛡}})\|_{\mathsf{SEP}(\underline{\mathrm{A}}_{\text{🔑}}\underline{\mathrm{A}}_{\text{🛡}}:\underline{\mathrm{B}}_{\text{🔑}}\underline{\mathrm{B}}_{\text{🛡}})} \leqslant \frac{C}{\sqrt{d}}\right) \geqslant 1 - e^{-c_0 d^3}$$
$$\mathbf{P}\left(\|\gamma - \mathcal{S}(\mathrm{A}\mathrm{A}_{\text{🛡}}{:}\mathrm{B}\mathrm{B}_{\text{🛡}})\|_{\mathsf{PPT}(\underline{\mathrm{A}}_{\text{🔑}}\underline{\mathrm{A}}_{\text{🛡}}:\underline{\mathrm{B}}_{\text{🔑}}\underline{\mathrm{B}}_{\text{🛡}})} \geqslant c\right) \geqslant 1 - e^{-c_0 d^4}$$

where $c_0, c, C > 0$ are universal constants.

Proof. For the first claim, we know from [AL14, Section 6.1] that there exist universal constants $c_0, C > 0$ such that,

$$\mathbf{P}\big(\|\sigma^{+} - \sigma^{-}\|_{\mathsf{SEP}} \leqslant C/\sqrt{d}\big) \geqslant 1 - e^{-c_0 d^3}.$$

Hence by Lemma 20 above, we have (just relabelling $3C/2$ as C), that

$$\mathbf{P}\big(\|\gamma - \mathcal{S}\|_{\mathsf{SEP}} \leqslant C/\sqrt{d}\big) \geqslant 1 - e^{-c_0 d^3}.$$

For the second claim, we know from [AL14, Theorem 5] that there exist universal constants $c_0, c > 0$ such that

$$\mathbf{P}\big(\|\sigma^{+} - \sigma^{-}\|_{\mathsf{PPT}} \geqslant c\big) \geqslant 1 - e^{-c_0 d^4}.$$

Hence by Lemma 21 above, we have (just relabelling $c/3$ as c),

$$\mathbf{P}(\|\gamma - \mathcal{S}\|_{\mathsf{PPT}} \geqslant c) \geqslant 1 - e^{-c_0 d^4}.$$

This concludes the proof of Theorem 34. □

In words, Theorem 34 tells us the following: the considered random private state γ is, with probability going to 1 as the dimension d grows, at a SEP distance of at most $C/\sqrt{d}$ and at a PPT distance of at least c from the set of separable states. So in conclusion, what we learn from it is that there exist private states which are barely distinguishable from being separable for observers which can only perform separable (and even more so LOCC) measurements on them. However, this data hiding property is not maintained when relaxing to PPT measurements, since these private states keep a constant distinguishability from separable states under PPT measurements. We now derive the analogue of Theorem 34 when distinguishability is measured in local relative entropy.

Theorem 35. *Let γ be a random private state on $\mathrm{A}_{\text{🔑}}\mathrm{A}_{\text{🛡}}\mathrm{B}_{\text{🔑}}\mathrm{B}_{\text{🛡}}$ as defined by Construction 33. Then*

$$\mathbf{P}\left(D_{\mathsf{SEP}(\underline{\mathrm{A}}_{\text{🔑}}\underline{\mathrm{A}}_{\text{🛡}}:\underline{\mathrm{B}}_{\text{🔑}}\underline{\mathrm{B}}_{\text{🛡}})}(\gamma\|\mathcal{S}(\mathrm{AA}':\mathrm{BB}')) \leqslant C\frac{\log d}{\sqrt{d}}\right) \geqslant 1-e^{-c_0 d}$$
$$\mathbf{P}\left(D_{\mathsf{PPT}(\underline{\mathrm{A}}_{\text{🔑}}\underline{\mathrm{A}}_{\text{🛡}}:\underline{\mathrm{B}}_{\text{🔑}}\underline{\mathrm{B}}_{\text{🛡}})}(\gamma\|\mathcal{S}(\mathrm{AA}':\mathrm{BB}')) \geqslant c\right) \geqslant 1-e^{-c_0 d^4}$$

where $c_0, c, C > 0$ are universal constants.

The upper bound on $D_{\mathsf{SEP}}(\gamma\|\mathcal{S})$ in Theorem 35 above is not tight: the $\log d$ factor can actually be removed. The derivation of this improved upper bound is relegated to Section 3.5.3 as it is much more involved, and requires developing several additional tools (which might be of independent interest).

Proof. For the first probability estimate, we know from Theorem 34 that, with probability greater than $1-e^{-c_0 d}$, $\|\gamma-\hat{\gamma}\|_{\mathsf{SEP}} \leqslant C/\sqrt{d}$. Hence by Equation (2.12), we get that, with probability greater than $1-e^{-c_0 d}$,

$$\begin{aligned} D_{\mathsf{SEP}}(\gamma\|\mathcal{S}) &= |D_{\mathsf{SEP}}(\gamma\|\mathcal{S}) - D_{\mathsf{SEP}}(\hat{\gamma}\|\mathcal{S})| \\ &\leqslant \frac{C}{\sqrt{d}}\log(2d) + g\left(\frac{C}{\sqrt{d}}\right) \\ &\leqslant \frac{C'\log d}{\sqrt{d}}, \end{aligned}$$

where the equality is due to $\hat{\gamma} \in \mathcal{S}$, so that $D_{\mathsf{SEP}}(\hat{\gamma}\|\mathcal{S}) = 0$.

For the second probability estimate, we know from Theorem 34 that, with probability greater than $1-e^{-c_0 d^4}$, $\|\gamma-\mathcal{S}\|_{\mathsf{PPT}} \geqslant c$. By Pinsker's inequality[Pin60; Ver14], this implies that, with probability greater than $1-e^{-c_0 d^4}$,

$$D_{\mathsf{PPT}}(\gamma\|\mathcal{S}) \geqslant \frac{1}{2\ln 2}\|\gamma-\mathcal{S}\|_{\mathsf{PPT}}^2 \geqslant \frac{c^2}{2\ln 2} = c'. \qquad \square$$

Remark 36. Note that the proof of Theorem 35 in fact establishes something slightly stronger, namely that $C\log d/\sqrt{d}$ is also an upper bound on $D_{\mathsf{SEP}(\underline{\mathrm{A}}_{\text{🔑}}\underline{\mathrm{A}}_{\text{🛡}}:\mathrm{B}_{\text{🔑}}\mathrm{B}_{\text{🛡}})}(\gamma\|\mathcal{S})$. Indeed, since $\mathsf{SEP}(\mathrm{A}_{\text{🔑}}\mathrm{A}_{\text{🛡}}:\mathrm{B}_{\text{🔑}}\mathrm{B}_{\text{🛡}})$ contains classical communication from $\mathrm{A}_{\text{🔑}}\mathrm{A}_{\text{🛡}}$ to $\mathrm{B}_{\text{🔑}}\mathrm{B}_{\text{🛡}}$, we know by Equation (2.6) that we actually have

$$\|\gamma-\hat{\gamma}\|_{\mathsf{SEP}(\underline{\mathrm{A}}_{\text{🔑}}\underline{\mathrm{A}}_{\text{🛡}}:\mathrm{B}_{\text{🔑}}\mathrm{B}_{\text{🛡}})} = \|\gamma-\hat{\gamma}\|_{\mathsf{SEP}(\underline{\mathrm{A}}_{\text{🔑}}\underline{\mathrm{A}}_{\text{🛡}}:\underline{\mathrm{B}}_{\text{🔑}}\underline{\mathrm{B}}_{\text{🛡}})}.$$

And thus Theorem 35 holds for $D_{\mathsf{SEP}(\underline{\mathrm{A}}_{\text{🔑}}\underline{\mathrm{A}}_{\text{🛡}}:\mathrm{B}_{\text{🔑}}\mathrm{B}_{\text{🛡}})}$ exactly as it holds for $D_{\mathsf{SEP}(\underline{\mathrm{A}}_{\text{🔑}}\underline{\mathrm{A}}_{\text{🛡}}:\underline{\mathrm{B}}_{\text{🔑}}\underline{\mathrm{B}}_{\text{🛡}})}$.

Theorem 35 teaches us that the same qualitative conclusion as that of Theorem 34 holds when measuring distance from the set of separable states in relative entropy rather than trace norm: with probability going to one as the dimension d grows, our random private state γ has a very small relative entropy of entanglement when restricted to separable (and even more so LOCC) measurements, but a high one when only restricted to PPT measurements.

Finally, let us emphasize that because our random private states are distinguishable under PPT operations, then they will also be distillable under PPT operations, and thus all the upper bounds on the distillable entanglement will be bounded away by the distillable entanglement itself. The usual computable measures based on the PPT criterion will be useless. For instance, the logarithmic negativity of our random private state γ, i.e. $E_N(\gamma) := \log \|\gamma^\Gamma\|_1$, is with high probability high. Indeed, in matrix notation we have

$$\gamma^\Gamma = \frac{1}{2}\begin{pmatrix} \mathbb{1}/d^2 & 0 & 0 & 0 \\ 0 & 0 & (\sigma^+ - \sigma^-)^\Gamma/2 & 0 \\ 0 & (\sigma^+ - \sigma^-)^\Gamma/2 & 0 & 0 \\ 0 & 0 & 0 & \mathbb{1}/d^2 \end{pmatrix},$$

and we thus easily see that $\|\gamma^\Gamma\|_1 = 1 + \left\|(\sigma^+ - \sigma^-)^\Gamma\right\|_1$. Now, the spectrum of the random matrix $(\sigma^+ - \sigma^-)^\Gamma$ can be precisely studied (see e.g. [Mon13, Section 3]), but for our purposes it is in fact enough to simply know that there exists a universal constant $c > 0$ such that $\left\|(\sigma^+ - \sigma^-)^\Gamma\right\|_1 \geqslant c$ with high probability. And therefore, $E_N(\gamma) \geqslant \log(1+c)$ with high probability.

3.5.3 Improvement

In this section we give a lengthier proof of Theorem 35, that does away with the $\log d$ factor in the scaling of the SEP relative entropy.

Restricted operator ordering

Contrary to the definitions of L norm and L relative entropies (see Section 2.1), the definition below, as far as we are aware of, has not been introduced in the literature before.

Definition 37 (L (partial) ordering). *For any Hermitian operators X, Y on H, we define the notion of ordering in restriction to a set of measurements L by:*

$$X \leqslant_{\mathsf{L}} Y \text{ if } \forall\, \mathcal{M} \in \mathsf{L},\ \mathcal{M}(X) \leqslant \mathcal{M}(Y)\,.$$

Note that the condition in Definition 37 above can be rewritten as point-wise ordering, namely if $\{M_i\}_{i\in I}$, are the measurement operators of $\mathcal{M}$

$$\mathcal{M}(X) \leqslant \mathcal{M}(Y) \text{ if } \forall\, i \in I,\ \mathrm{tr}(T_i X) \leqslant \mathrm{tr}(T_i Y).$$

We now explore how this notion of measurement ordering connects to that of measurement distance. We begin with the following easy observations in Lemmas 38 and 39 below. Lemma 38 will be used later in the section, while Lemma 39 is just stated here as an independent comment.

Lemma 38. *Let ρ, σ be states and L be a set of measurements on H. If $\rho \leqslant_{\mathsf{L}} (1+\epsilon)\sigma$ for some $\epsilon > 0$, then $D_{\mathsf{L}}(\rho \,\|\, \sigma) \leqslant \log(1+\epsilon)$.*

Proof. If p, q are probability distributions satisfying $p \leqslant (1+\epsilon)q$ for some $\epsilon > 0$, then clearly

$$D(p \,\|\, q) = \sum_i p_i \log\left(\frac{p_i}{q_i}\right) \leqslant \sum_i p_i \log(1+\epsilon) = \log(1+\epsilon).$$

Now, for any $\mathcal{M} \in \mathsf{L}$, $\mathcal{M}(\rho)$ and $\mathcal{M}(\sigma)$ are classical probability distributions. So what we have shown is that, if $\mathcal{M}(\rho) \leqslant (1+\epsilon)\mathcal{M}(\sigma)$ for all $\mathcal{M} \in \mathsf{L}$, then $D\left(\mathcal{M}(\rho) \| \mathcal{M}(\sigma)\right) \leqslant \log(1+\epsilon)$ for all $\mathcal{M} \in \mathsf{L}$. And this is exactly the statement in Lemma 38. □

Lemma 39. *Let ρ, σ be states and L be a set of measurements on H. If $\rho \leqslant_{\mathsf{L}} (1+\epsilon)\sigma$ and $\sigma \leqslant_{\mathsf{L}} (1+\epsilon)\rho$ for some $0 < \epsilon < 1$, then $\|\rho - \sigma\|_{\mathsf{L}} \leqslant \epsilon/(1-\epsilon/2) \leqslant 2\epsilon$.*

Proof. If p, q are probability distributions satisfying $p \leqslant (1+\epsilon)q$ and $q \leqslant (1+\epsilon)p$ for some $0 < \epsilon < 1$, then for all i

$$\begin{aligned} p_i - q_i &\leqslant \epsilon q_i \\ q_i - p_i &\leqslant \epsilon p_i. \end{aligned}$$

Hence as a consequence,

$$\begin{aligned} \sum_i |p_i - q_i| &\leqslant \epsilon \sum_i \max(p_i, q_i) \\ &= \epsilon \sum_i \frac{p_i + q_i + |p_i - q_i|}{2} \\ &= \epsilon \left(1 + \frac{1}{2} \sum_i |p_i - q_i| \right). \end{aligned}$$

Now, for any $\mathcal{M} \in \mathsf{L}$, $\mathcal{M}(\rho)$ and $\mathcal{M}(\sigma)$ are classical probability distributions. So what we have shown is that, if $\mathcal{M}(\rho) \leqslant (1+\epsilon)\mathcal{M}(\sigma)$ and $\mathcal{M}(\sigma) \leqslant (1+\epsilon)\mathcal{M}(\rho)$ for all $\mathcal{M} \in \mathsf{L}$, then $\|\mathcal{M}(\rho) - \mathcal{M}(\sigma)\|_1 \leqslant \epsilon/(1-\epsilon/2)$ for all $\mathcal{M} \in \mathsf{L}$. And this is exactly the statement in Lemma 39. □

SEP ordering

We start with establishing a technical result about the maximum overlap with separable states for the difference of two random orthogonal states. It has some similarities with the SEP data hiding result of [AL14, Theorem 5], but cannot be directly derived from it, which is why we re-do the whole argument.

Proposition 40. *Let $\sigma^{\pm}$ be random orthogonal states on $\mathrm{A}_{\text{🛡}}\mathrm{B}_{\text{🛡}}$ as defined by Construction 30. Then, there exist universal constants $c_0, C > 0$ such that*

$$\mathbf{P}\left(\sup_{\sigma \in \mathcal{S}(\mathrm{A}_{\text{🛡}}:\mathrm{B}_{\text{🛡}})} \left| \mathrm{tr}\left(\sigma[\sigma^+ - \sigma^-]\right) \right| \leqslant \frac{C}{d^{5/2}} \right) \geqslant 1 - e^{-c_0 d}.$$

Proof. With some abuse of notation, let us define the following function on the traceless Hermitian operators

$$\|X\|_{\mathcal{S}^\circ} := \sup_{\sigma \in \mathcal{S}(\mathrm{A}_{\text{🛡}}:\mathrm{B}_{\text{🛡}})} \left| \mathrm{tr}\left(\sigma X\right) \right|$$

and let us remark that it satisfies the triangle inequality. To be more precise, $\|X\|_{\mathcal{S}^\circ}$ is the support function of the symmetrization $\Sigma = \mathrm{conv}\left\{-\mathcal{S} \cup \mathcal{S}\right\}$ of the separable states around the origin, and thus $\|X\|_{\mathcal{S}^\circ}$ is actually $\|X\|_{\Sigma^\circ}$, namely it is the gauge of the polar of Σ, see [AS06] for more details. Notice that $\|X\|_{\mathcal{S}^\circ}$ is not to be confused with $\|X\|_{\mathcal{S}_0}$ introduced in Lemma 31. In particular the former is always smaller than the ∞-norm, while the latter is always larger than the trace norm. Still, we will follow the same proof structure of Lemma 31.

Recall that the random states are defined by $\sigma^{\pm} = U\bar{\sigma}^{\pm}U^{\dagger}$, where U is a Haar-distributed unitary on $\mathbb{C}^d \otimes \mathbb{C}^d$ and $\bar{\sigma}^{\pm}$ some fixed orthogonal maximally mixed states on $d^2/2$-dimensional subspaces of $\mathbb{C}^d \otimes \mathbb{C}^d$. Let

$$\begin{aligned} \bar{\Delta} &= \bar{\sigma}^+ - \bar{\sigma}^- \\ \Delta &= \sigma^+ - \sigma^- = U\bar{\Delta}U^{\dagger}. \end{aligned}$$

The statement to prove therefore is

$$\mathbf{P}\left(\|\Delta\|_{\mathcal{S}^\circ} \leqslant \frac{C}{d^{5/2}} \right) \geqslant 1 - e^{-c_0 d}.$$

To prove the statement, we compute the expectation value $\mathbf{E}\,\|\Delta\|_{\mathcal{S}^\circ} = \mathbf{E}\,\left\|U\bar{\Delta}U^\dagger\right\|_{\mathcal{S}^\circ}$ over the random variable U, then we estimate the Lipschitz constant of $\left\|U\bar{\Delta}U^\dagger\right\|_{\mathcal{S}^\circ}$ as a function of U and use this to argue that being close to the expected value happens with high probability. $\|X\|_{\mathcal{S}^\circ}$ is not unitary invariant, however the function $\mathbf{E}\,\left\|UXU^\dagger\right\|_{\mathcal{S}^\circ}$ is unitary invariant on X, while still being convex.

As explained in Lemma 31 leading to Equation (3.29), we know from [AL14, Lemma 6] that for any unitary-invariant convex function g of any traceless Hermitian operators X and Y on $\mathbb{C}^d \otimes \mathbb{C}^d$, we have

$$\frac{1}{2d^2}\frac{\|X\|_1}{\|Y\|_\infty} \leqslant \frac{g(X)}{g(Y)} \leqslant 2d^2\frac{\|X\|_\infty}{\|Y\|_1}\,.$$

Now, we let $Y = \bar{\Delta}$ for which $\left\|\bar{\Delta}\right\|_1 = 2$ and $\left\|\bar{\Delta}\right\|_\infty = 2/d^2$:

$$\frac{1}{4}\|X\|_1 \leqslant \frac{g(X)}{g(\bar{\Delta})} \leqslant d^2\|X\|_\infty\,.$$

Then we let $g(X) = \mathbf{E}\,\left\|UXU^\dagger\right\|_{\mathcal{S}^\circ}$, which is unitary invariant by construction and convex by the convexity of $\|X\|_{\mathcal{S}^\circ}$:

$$\frac{1}{4}\|X\|_1 \leqslant \frac{\mathbf{E}\,\left\|UXU^\dagger\right\|_{\mathcal{S}^\circ}}{\mathbf{E}\,\left\|U\bar{\Delta}U^\dagger\right\|_{\mathcal{S}^\circ}} \leqslant d^2\|X\|_\infty\,.$$

We now let X be again a Gaussian vector G on the traceless Hermitian operators (Gaussian unitary ensemble) on $\mathbb{C}^d \otimes \mathbb{C}^d$. This makes $\mathbf{E}\,\left\|UGU^\dagger\right\|_{\mathcal{S}^\circ} = \mathbf{E}\,\|G\|_{\mathcal{S}^\circ}$. As in Lemma 31, taking expectation values over the inequalities gives

$$\frac{1}{4}\,\mathbf{E}\,\|G\|_1 \leqslant \frac{\mathbf{E}\,\|G\|_{\mathcal{S}^\circ}}{\mathbf{E}\,\|\Delta\|_{\mathcal{S}^\circ}} \leqslant d^2\,\mathbf{E}\,\|G\|_\infty\,,$$

and using that $\mathbf{E}\,\|G\|_1 \sim d^3$ and $\mathbf{E}\,\|G\|_\infty \sim d$ further gives:

$$\mathbf{E}\,\|\Delta\|_{\mathcal{S}^\circ} \sim \frac{1}{d^3}\,\mathbf{E}\,\|G\|_{\mathcal{S}^\circ}\,.$$

We now know from [AS06, Equation 7] [6] that $\mathbf{E}\,\|G\|_{\mathcal{S}^\circ}$ is at most of order $\sqrt{d}$ and, therefore there exists a universal constant $C > 0$ such that

$$\mathbf{E}\,\|\Delta\|_{\mathcal{S}^\circ} \leqslant \frac{C}{d^{5/2}}. \tag{3.32}$$

Now we have to show that this average behaviour is generic for large d, because the function $f(U) := \left\|U\bar{\Delta}U^\dagger\right\|_{\mathcal{S}^\circ}$ is regular enough in the Euclidean norm: we claim that it is a

[6] As remarked in [AS06, Section 2.1], for a convex set K of $\mathbb{R}^n$ we have $\mathbf{E}\,\|G\|_{K^\circ} = \gamma\, w(K) \geqslant \gamma\,\mathrm{vrad}(K)$, where $\gamma = \mathbf{E}\,\|G\|_2$, w is the mean width and vrad is the volume radius. Furthermore, [AS06] shows that in particular $w(\Sigma) \sim \mathrm{vrad}(\Sigma)$. What [AS06, Equation 7] and the remarks below show, is that we have $\mathrm{vrad}(\Sigma) \sim \sqrt{d}/\gamma$, which gives us $\mathbf{E}\,\|G\|_{\mathcal{S}^\circ} \sim \sqrt{d}$.

$8/d^2$-Lipschitz function. Indeed by the triangle inequality for $\|\cdot\|_{\mathcal{S}^\circ}$, for any unitaries U and V on $\mathbb{C}^d \otimes \mathbb{C}^d$ we have

$$\begin{aligned}
&|f(U) - f(V)| \\
&\quad = \left| \left\| U\bar{\Delta}U^\dagger \right\|_{\mathcal{S}^\circ} - \left\| V\bar{\Delta}V^\dagger \right\|_{\mathcal{S}^\circ} \right| \\
&\quad \leqslant \left\| U\bar{\Delta}U^\dagger - V\bar{\Delta}V^\dagger \right\|_{\mathcal{S}^\circ} \\
&\quad \leqslant \left\| U\bar{\sigma}^+U^\dagger - V\bar{\sigma}^+V^\dagger \right\|_{\mathcal{S}^\circ} + \left\| U\bar{\sigma}^-U^\dagger - V\bar{\sigma}^-V^\dagger \right\|_{\mathcal{S}^\circ}.
\end{aligned}$$

We can then use that $\mathcal{S} \subset B_1$, together with duality of the 1-norm and the ∞-norm, to get

$$\begin{aligned}
&|f(U) - f(V)| \\
&\quad \leqslant \left\| U\bar{\sigma}^+U^\dagger - V\bar{\sigma}^+V^\dagger \right\|_{B_1^\circ} + \left\| U\bar{\sigma}^-U^\dagger - V\bar{\sigma}^-V^\dagger \right\|_{B_1^\circ} \\
&\quad = \left\| U\bar{\sigma}^+U^\dagger - V\bar{\sigma}^+V^\dagger \right\|_\infty + \left\| U\bar{\sigma}^-U^\dagger - V\bar{\sigma}^-V^\dagger \right\|_\infty \\
&\quad \leqslant \left\| U\bar{\sigma}^+(U^\dagger - V^\dagger) \right\|_\infty + \left\| (U-V)\bar{\sigma}^+V^\dagger \right\|_\infty \\
&\qquad + \left\| U\bar{\sigma}^-(U^\dagger - V^\dagger) \right\|_\infty + \left\| (U-V)\bar{\sigma}^-V^\dagger \right\|_\infty \\
&\quad = 2\left\| (U-V)\bar{\sigma}^+ \right\|_\infty + 2\left\| (U-V)\bar{\sigma}^- \right\|_\infty \\
&\quad \leqslant 2\left\| \bar{\sigma}^+ \right\|_\infty \|U-V\|_\infty + 2\left\| \bar{\sigma}^- \right\|_\infty \|U-V\|_\infty \\
&\quad = \frac{8}{d^2}\|U-V\|_\infty \leqslant \frac{8}{d^2}\|U-V\|_2.
\end{aligned}$$

which shows that f is $8/d^2$-Lipschitz.

Now, we know from [MM+13, Corollary 17] that any L-Lipschitz function g on the unitaries on $\mathbb{C}^D$ (equipped with the Euclidean metric) satisfies the concentration estimate: if U is a Haar-distributed unitary on $\mathbb{C}^D$, then for all $\epsilon > 0$, $\mathbf{P}(g(U) > \mathbf{E}\, g + \epsilon) \leqslant e^{-c_0 D\epsilon^2/L^2}$, where $c_0 > 0$ is a universal constant. Combining the above estimate on the Lipschitz constant of $\left\|U\bar{\Delta}U^\dagger\right\|_{\mathcal{S}^\circ}$ with the estimate on its expected value from Equation (3.32), we thus get for all $\epsilon > 0$

$$\begin{aligned}
\mathbf{P}\left(\|\Delta\|_{\mathcal{S}^\circ} > \frac{C}{d^{5/2}} + \epsilon \right) &\leqslant \mathbf{P}\left(\|\Delta\|_{\mathcal{S}^\circ} > \mathbf{E}\,\|\Delta\|_{\mathcal{S}^\circ} + \epsilon \right) \\
&\leqslant e^{-c_0 d^6 \epsilon^2/64}.
\end{aligned}$$

The advertised result follows from choosing $\epsilon = C/d^{5/2}$ (and suitably relabelling the constants). □

Thanks to Proposition 40, we can now show that our random private states and their key-attacked versions are with high probability SEP ordered with a constant close to 1.

Proposition 41. *Let γ and $\hat{\gamma}$ on* $\mathrm{A}_{\text{🔑}}\mathrm{A}_{\text{🛡}}\mathrm{B}_{\text{🔑}}\mathrm{B}_{\text{🛡}}$ *be a random private state and its key-attacked state as defined by Construction 33. Then, there exist universal constants $c_0, C > 0$ such that*

$$\mathbf{P}\left(\gamma \leqslant_{\mathsf{SEP}(\underline{\mathrm{A}}_{\text{🔑}}\underline{\mathrm{A}}_{\text{🛡}}:\underline{\mathrm{B}}_{\text{🔑}}\underline{\mathrm{B}}_{\text{🛡}})} \left(1 + \frac{C}{\sqrt{d}}\right) \hat{\gamma} \right) \geqslant 1 - e^{-c_0 d}.$$

Proof. By Definition 37, to prove Proposition 41 it suffices to show that, with probability greater than $1 - e^{-c_0 d}$, forall positive-semidefinite operators $0 \leqslant M, N \leqslant \mathbb{1}$

$$\operatorname{tr}(M \otimes N\gamma) \leqslant \left(1 + \frac{C}{\sqrt{d}}\right) \operatorname{tr}(M \otimes N\hat{\gamma}). \tag{3.33}$$

Now, let $\Delta = \sigma^+ - \sigma^-$ and observe that, for any $\epsilon > 0$, we have

$$(1+\epsilon)\hat{\gamma} - \gamma = \frac{1}{2}\begin{pmatrix} \epsilon\,\dfrac{\mathbb{1}}{d^2} & 0 & 0 & -\dfrac{\Delta}{2} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -\dfrac{\Delta}{2} & 0 & 0 & \epsilon\,\dfrac{\mathbb{1}}{d^2} \end{pmatrix}.$$

Given $0 \leqslant M \leqslant \mathbb{1}$ on A🔑A🛡 and $0 \leqslant N \leqslant \mathbb{1}$ on BB′, we write them in the block-form

$$M = \begin{pmatrix} M_1 & \tilde{M} \\ \tilde{M}^\dagger & M_2 \end{pmatrix} \text{ and } N = \begin{pmatrix} N_1 & \tilde{N} \\ \tilde{N}^\dagger & N_2 \end{pmatrix},$$

where $M_1, M_2, \tilde{M}$ are operators on A🛡 and $N_1, N_2, \tilde{N}$ are operators on B′, with $0 \leqslant M_1, M_2 \leqslant \mathbb{1}$ and $0 \leqslant N_1, N_2 \leqslant \mathbb{1}$. A straightforward calculation shows that

$$\operatorname{tr}\left(M \otimes N\left[(1+\epsilon)\,\hat{\gamma} - \gamma\right]\right) = \epsilon\frac{\operatorname{tr} K}{d^2} - \frac{1}{2}\operatorname{tr}\left(\tilde{K}\Delta\right) \tag{3.34}$$

where $K = \left(M_1 \otimes N_1 + M_2 \otimes N_2\right)/2 \geqslant 0$ and $\tilde{K} = \tilde{K}^\dagger = \left(\tilde{M} \otimes \tilde{N} + \tilde{M}^\dagger \otimes \tilde{N}^\dagger\right)/2$, satisfying $\pm\tilde{K} \leqslant K$.

We now expand $\tilde{K}$ into the difference of the positive-semidefinite components of $\tilde{M}$ and $\tilde{N}$. Any operator $K \in \mathcal{L}(\mathrm{H})$ can be decomposed into the sum of an Hermitian and anti-Hermitian operator as $K = K_{\mathrm{Re}} + iK_{\mathrm{Im}}$ where $K_{\mathrm{Re}} = (K + K^\dagger)/2$ and $K_{\mathrm{Im}} = (K - K^\dagger)/2i$ are Hermitian. Also recall that if K is Hermitian then we denote with $K_\pm$ the positive- and negative-semidefinite part. Then we have

$$\operatorname{tr}\left(\Delta'(\rho - \sigma)\right) \leqslant \sum_{x,y\in\{+,-\},\, z\in\{\mathrm{Re},\mathrm{Im}\}} \left|\operatorname{tr}\left(\tilde{M}_{z,x} \otimes \tilde{N}_{z,y}\Delta\right)\right|,$$

where for each term we have $\tilde{M}_{z,x}, \tilde{N}_{z,y} \geqslant 0$, $\operatorname{tr}\tilde{M}_{z,x} \leqslant (\operatorname{tr} M_1 \operatorname{tr} M_2)^{1/2}$ and $\operatorname{tr}\tilde{N}_{z,y} \leqslant (\operatorname{tr} N_1 \operatorname{tr} N_2)^{1/2}$. So for all $x, y \in \{+,-\}$, $z \in \{\mathrm{Re},\mathrm{Im}\}$,

$$\begin{aligned} \left|\operatorname{tr}\left(\tilde{M}_{z,x} \otimes \tilde{N}_{z,y}\Delta\right)\right| &\leqslant \left(\operatorname{tr}\tilde{M}_{z,x} \operatorname{tr}\tilde{N}_{z,y}\right) \sup_{\sigma\in\mathcal{S}} \left|\operatorname{tr}(\sigma\Delta)\right| \\ &\leqslant (\operatorname{tr} M_1 \operatorname{tr} M_2)^{1/2} (\operatorname{tr} N_1 \operatorname{tr} N_2)^{1/2} \sup_{\sigma\in\mathcal{S}} \left|\operatorname{tr}(\sigma\Delta)\right| \\ &\leqslant \frac{\operatorname{tr} M_1 \operatorname{tr} N_1 + \operatorname{tr} M_2 \operatorname{tr} N_2}{2} \sup_{\sigma\in\mathcal{S}} \left|\operatorname{tr}(\sigma\Delta)\right| \\ &= \operatorname{tr}(K) \sup_{\sigma\in\mathcal{S}} \left|\operatorname{tr}(\sigma\Delta)\right|. \end{aligned}$$

Then

$$\operatorname{tr}\left(\tilde{K}\Delta\right) \leqslant 8 \operatorname{tr} K \sup_{\sigma\in\mathcal{S}} \left|\operatorname{tr}(\sigma\Delta)\right|.$$

Now we apply Proposition 40 and get that there exist constants $C, c_0 > 0$ such that with probability greater than $1 - e^{-c_0 d}$

$$\operatorname{tr}\left(\tilde{K}\Delta\right) \leqslant 8 \operatorname{tr}(K) \frac{C}{\sqrt{d}} \frac{1}{d^2}.$$

Inserting the above in Equation (3.34) we obtain that with probability greater than $1 - e^{-c_0 d}$ we have for all $0 \leqslant M, N \leqslant \mathbb{1}$

$$\operatorname{tr}\left(M \otimes N\left[(1+\epsilon)\,\hat{\gamma} - \gamma\right]\right) \geqslant \left(\epsilon - 4\frac{C}{\sqrt{d}}\right)\frac{\operatorname{tr} K}{d^2}.$$

The right-hand-side is positive as soon as $\epsilon \geqslant 4C/\sqrt{d}$, in which case Equation (3.33) indeed holds completing the proof. □

Improved bound

Using Proposition 41, we are now able to prove an upper bound which is better than the one appearing in Theorem 35.

Theorem 42. *Let γ on $\mathrm{A}_{\text{🔑}}\mathrm{A}_{\text{🛡}}\mathrm{B}_{\text{🔑}}\mathrm{B}_{\text{🛡}}$ be a random private state as defined by Construction 33. Then, there exist universal constants $c_0, C > 0$ such that*

$$\mathbf{P}\left(D_{\mathsf{SEP}(\underline{\mathrm{A}}_{\text{🔑}}\underline{\mathrm{A}}_{\text{🛡}}:\underline{\mathrm{B}}_{\text{🔑}}\underline{\mathrm{B}}_{\text{🛡}})}\left(\gamma \| \mathcal{S}(\mathrm{A}_{\text{🔑}}\mathrm{A}_{\text{🛡}}:\mathrm{B}_{\text{🔑}}\mathrm{B}_{\text{🛡}})\right) \leqslant \frac{C}{\sqrt{d}} \right) \geqslant 1 - e^{-c_0 d}\,.$$

Proof. We know from Proposition 41 that, with probability greater than $1 - e^{-c_0 d}$, $\gamma \leqslant_{\mathsf{SEP}} (1 + C/\sqrt{d})\hat{\gamma}$. Because $\hat{\gamma}$ is a separable state, we have by Lemma 38 that, with probability greater than $1 - e^{-c_0 d}$,

$$D_{\mathsf{SEP}}(\gamma \parallel \mathcal{S}) \leqslant D_{\mathsf{SEP}}\left(\gamma \parallel \hat{\gamma}\right) \leqslant \log\left(1 + \frac{C}{\sqrt{d}}\right) \leqslant \frac{C}{\sqrt{d}}.$$

This concludes the proof. □

For the sake of clarity, we focused on one particular way of constructing random private states. However, the properties that we described would hold true for many other random private state models. For instance, one could think of picking as states $\sigma^{\pm}$, two independent uniformly distributed mixed states on $\mathrm{A}_{\text{🛡}}\mathrm{B}_{\text{🛡}}$, or mixtures of order d^2 independent uniformly distributed pure states on $\mathrm{A}_{\text{🛡}}\mathrm{B}_{\text{🛡}}$. These would be with high probability approximately orthogonal, so that the random state γ on $\mathrm{A}_{\text{🔑}}\mathrm{A}_{\text{🛡}}\mathrm{B}_{\text{🔑}}\mathrm{B}_{\text{🛡}}$ formed out of them would be with high probability an approximate private state. Moreover, it would have with high probability all the previously observed features. It thus appears as a generic aspect of private states that their amount of distillable entanglement and their amount of data-hiding have to obey some trade-off. One important open question at this point would nonetheless be: what is the actual distribution of the random private states which are produced in "usual" quantum key distribution protocols? Indeed, however wide the range of models our results apply to, it would be interesting to know whether or not the outputs of error correction and privacy amplification procedures which are performed in practice fall into this general framework.

3.6 Summary

We have dedicated this chapter to the study of states of the form

$$\gamma^{\pm} = p^{+} \cdot \Phi^{+} \otimes \sigma^{\pm} + p^{-} \cdot \Phi^{-} \otimes \sigma^{\mp}.$$

We have seen that the maximally entangled states can allow for strictly more distillable entanglement than what would be achieved distinguishing the shield locally, in particular even if the shield states are known to hide the data from local observer. However, we have also seen that in large dimension this behaviour disappears, as shown in the last section, and the states are actually generically data hiding. In the very specific case when we can upper bound the distillable entanglement, we can upgrade the statement to asymptotic indistinguishability, where even arbitrarily many copies are allowed to distinguish the states. In the next chapter we will use the proven indistinguishability, to show that these correlations cannot be mediated by an untrusted party, indicating that the difference between entanglement distillation and key distillation induced by data hiding states is a purely bipartite property, and does exist in a general network scenario.

Chapter 4

Quantum Repeaters for Key Distillation

An immediate application of quantum teleportation is entanglement swapping. Consider Alice and Bob sharing a maximally entangled state $\Phi_{A^{\mathrm{in}}B^{\mathrm{out}}}$, and Imagine that Alice also shares a maximally entangled state with another system A' so that the total initial state is $\Phi_{A'A} \otimes \Phi_{A^{\mathrm{in}}B^{\mathrm{out}}}$. Applying the teleportation protocol as explained in Section 1.5, will result in an identity channel from A to B^{out}, and thus in the state $\Phi_{A'B^{\mathrm{out}}}$, "swapping" the roles of A' and A^{in}. However, the entanglement swapping protocol does not perform any operation on A' and thus there is no reason for it to be held by Alice. System A' could be held by a third party that, as manifest from the input state, has never interacted with Bob. Nonetheless the result is a maximally entangled state with Bob. Quantum teleportation thus allows parties that cannot interact directly to share maximally entangled states, and thus any quantum information, as long as there is a third party to mediate the interaction. The protocol can be repeated indefinitely, with many parties sharing maximally entangled states in series or, more generally, in a network, working together to provide a maximally entangled state between two designated end points. The end points are our from now on Alice and Bob, while the intermediate parties are called repeater stations.

The task of distilling target states between Alice and Bob in a repeater is plagued by the same noise in the quantum communication that would affect Alice and Bob in direct entanglement distillation. However, if we consider the distillation of maximally entangled states, this scenario is not more difficult that entanglement distillation. Namely, the best protocol is the one that distils maximally entangled states globally and then applies entanglement swapping. Because the rate of entanglement distillation with Charlie reduces to the bipartite case, initially the setting has been interesting only when modelling the noise also in the bipartite operations [Bri+98; Dür+99]. As soon as the fidelity of the maximally entangled state that can be achieved between the links is capped, the teleportation protocols collects the noise from all the links, exponentially decreasing the fidelity in the the number of links. Still to overcome the increased noise after the teleportation protocol, the best known protocols simply involve further entanglement distillation after the entanglement swapping[Dür+99; Jia+07], or using codes equivalent to entanglement distillation [Ben+96c; KL96; Mur+16].

However, if the goal is to distill perfect key across the repeater station, the situation is non-trivial already when considering perfect operations. The results reviewed in the previous chapter, displaying a strong separation between distillable key and distillable entanglement, open the possibility that a similar separation might exist in quantum networks. No example of such separation exist, on the contrary, it has been shown that some states might contain a lot of key between the nodes, but still almost no key can be extracted across a repeater station. In this chapter we will expand this no-go result to

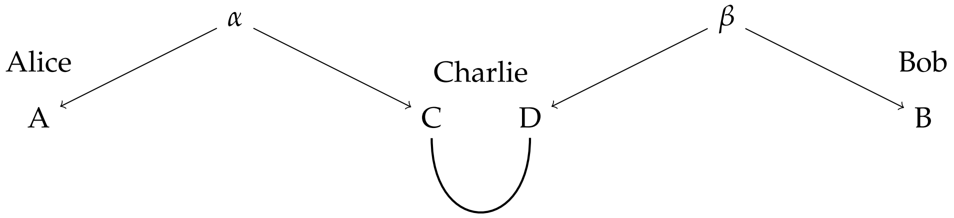


Figure 4.1: Source of entanglement for Alice, Charlie and Bob. Alice and Bob share no entanglement themselves, but they both share entanglement with Charlie, who can act globally on his two systems to mediate the generation of entanglement between Alice and Bob, as is the case of entanglement swapping.

general classes of states and protocols. Our results point at a potentially important pitfall to be aware of in the implementation of real QKD networks. A network might have a good key rate between adjacent nodes and have good operations at the repeater stations, but this is not enough to guarantee a good key rate between distant nodes. Our results are another step pointing toward the distillable entanglement being the only relevant resource for repeating quantum information. If this turned out to be true, then small deviations from the designed distributed states might have a large effect on the key rate between non adjacent nodes.

4.1 Repeater Entanglement Distillation

Ultimately, we would like to have a full understanding of key distillation in a general network scenario. In this chapter we will be studying the limitations in the rates of key that can be achieved. Thus, while realistic repeaters have multiple stations, we can always reduce to single repeater stations by grouping all the stations together, this only gives more power to the network and can only increase the rate. The reduction to a single station thus provides upper bounds without loss of generality. We will consider the simplest scenario beyond bipartite distillation, namely where we add a single third party Charlie to mediate the distillation between two parties Alice and Bob. This setting has been introduced in [Bäu+15]. The help of Charlie makes sense if Alice and Bob cannot generate useful entanglement by themselves, and thus we will assume that the source produces states that are independent for Alice and for Bob. More precisely, we will assume that the source produces product states $\alpha \otimes \beta$, of entangled states $\alpha \in \mathcal{D}(\mathrm{AC})$ shared between Alice and Charlie, and $\beta \in \mathcal{D}(\mathrm{DB})$ shared between Charlie and Bob, as displayed in Figure 4.1. Using only tripartite operations, Alice and Bob are supposed to distill the maximal amount entanglement, and later the maximal amount of perfect perfect key, where the task is made possible because Charlie can act globally on his share of the states. However, while Charlie is essential to achieve the goal, in the key distillation scenario he is also untrusted, therefore the key must be secret also from Charlie. Equivalently, we will say that Charlie's systems are given to the eavesdropper at the end of the protocol, namely that he is traced out. We will now generalize the classes of bipartite operations we have seen to the tripartite setting, and formally define the repeater distillable entanglement as a warm up to the repeater distillable key.

Multipartite operations and measurements

The generalization of separable states to multipartite systems is straightforward if we look at states that are fully separable across all parties. For three parties, the fully separable

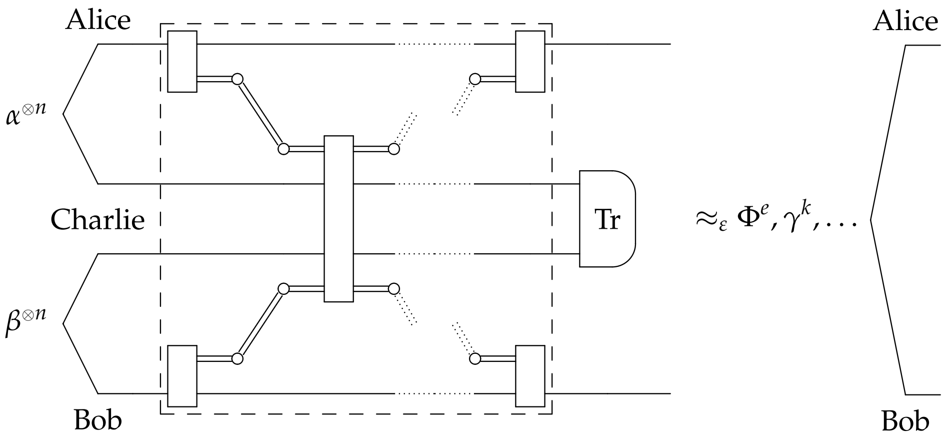


Figure 4.2: Quantum circuit for single-node repeater protocols distilling key. The dashed box is a tripartite LOCC protocol. The double lines are the classical communication. Because the communication with Charlie is two-way, it also implies two-way communication between Alice and Bob, and thus we omitted their communication lines.

states of Alice, Charlie and Bob with respective systems A, C and B are defined as

$$\mathcal{S}(\mathrm{A{:}C{:}B}) := \operatorname{conv}(\mathcal{D}(\mathrm{A}) \otimes \mathcal{D}(\mathrm{C}) \otimes \mathcal{D}(\mathrm{B})).$$

As for the bipartite case, they define the set of tripartite separable operations as those channels $\Lambda \in \mathsf{CPTP}(\mathrm{A^{in}{:}C^{in}{:}B^{in}} \rangle \mathrm{A^{out}{:}C^{out}{:}B^{out}})$ such that for all systems A, C and B

$$\mathrm{id}_{\mathrm{ACB}} \otimes \Lambda(\mathcal{S}(\mathrm{AA^{in}{:}CC^{in}{:}BB^{in}})) \subseteq \mathcal{S}(\mathrm{AA^{out}{:}CC^{out}{:}BB^{out}}). \tag{4.1}$$

We will not need a similar generalization for PPT operations.

As usual, we denote with $\mathsf{SEP}(\mathrm{A^{in}{:}C^{in}{:}B^{in}} \rangle \mathrm{A^{out}{:}C^{out}{:}B^{out}})$ the set of operations satisfying Equation (4.1), and with $\mathsf{SEP}(\mathrm{A{:}C{:}B})$ the set or all separable operations on ACB. This still defines a class of tripartite operations. Namely SEP contains local operations, which are simply

$$\mathsf{LO}(\mathrm{A{:}C{:}B}) := \mathsf{CPTP}(\mathrm{A}) \otimes \mathsf{CPTP}(\mathrm{C}) \otimes \mathsf{CPTP}(\mathrm{B}),$$

and it is closed under composition and tensor products, with an analogous generalization of these properties definition from Section 1.4.2, which we leave implicit. If L is a class of tripartite operations like $\mathsf{SEP}(\mathrm{A{:}C{:}B})$, we call repeater operations, or tripartite operations from Alice, Charlie and Bob to Alice and Bob, those operations $\mathsf{L}(\mathrm{A{:}C{:}B} \rangle \mathrm{A^{out}{:}\mathbb{C}{:}B^{out}})$ where Charlie is removed at the end.

Bipartite LOCC was constructed explicitly in Section 1.3.2. In Section 1.4.2 though, we remarked that the various bipartite LOCC classes are the smallest classes of operations that are closed under local operations and contain a specific combination of classical communication. We thus define tripartite LOCC axiomatically as the minimal class of operations $\mathsf{LOCC}(\mathrm{A^{in}{:}C^{in}{:}B^{in}} \rangle \mathrm{A^{out}{:}C^{out}{:}B^{out}})$ containing two-way classical communication from Charlie to Alice and Bob, where again the definition of classical communication from Section 1.4.2 generalizes straightforwardly to the multipartite setting. Notice that just requiring Charlie to be connected to Alice and Bob with two-way communication, implies Alice and Bob being connected with two-way communication. Later we will restrict the communication of Charlie to be only a sender to Alice and Bob, and not a receiver. In that case, we will explicitly mention that Alice and Bob are connected by two-way communication.

Repeater distillable entanglement

Let us first consider the rate of entanglement distillation achieved by Alice, Charlie and Bob, namely by a repeater with a single node. The repeater distillable entanglement is defined in analogy to the other distillation rates, and in particular the to distillable entanglement defined in Section 2.4. For any class of tripartite operations L the best finite rate for entanglement distillation under repeater operations is

$$E^{\varepsilon}_{DR,\mathsf{L}}(\alpha,\beta) := \sup\left\{\log e : \Phi^{\log e} \approx_{\varepsilon} \mathsf{L}(\mathbb{C}^e{:}\mathbb{C}{:}\mathbb{C}^e \langle \mathrm{A{:}CD{:}B}) \circ (\alpha \otimes \beta)\right\}$$

and the repeater distillable entanglement is the the best achievable rate

$$E_{DR,\mathsf{L}}(\alpha,\beta) = \lim_{\varepsilon\to 0} \limsup_{n\to\infty} \frac{1}{n} E^{\varepsilon}_{DR}(\alpha^{\otimes n}, \beta^{\otimes n}).$$

(E_R is the common notation we use for the relative entropy of entanglement and so we avoid it here). Notice how the output state is a bipartite state, as imposed by the fact that the output system is trivial ($\mathbb{C}$) at Charlie. The physical rates are of course the ones achievable with LOCC operations:

$$E_{DR}(\alpha,\beta) \equiv E_{DR,\mathsf{LOCC}(\mathrm{A{:}CD{:}B})}(\alpha,\beta) \tag{4.2}$$

Because we can group Charlie with either Alice or Bob, namely because we have the inclusion into the bipartite LOCC as

$$\mathsf{LOCC}(\mathrm{A{:}CD{:}B}) \subseteq \mathsf{LOCC}(\mathrm{A{:}CDB}), \mathsf{LOCC}(\mathrm{ACD{:}B}),$$

then we have

$$E_{DR}(\alpha,\beta) \leqslant E_{D,\mathsf{LOCC}(\mathrm{A{:}CB})}(\alpha\otimes\beta) = E_{D,\mathsf{LOCC}(\mathrm{A{:}C})}(\alpha)$$
$$E_{DR}(\alpha,\beta) \leqslant E_{D,\mathsf{LOCC}(\mathrm{AC{:}B})}(\alpha\otimes\beta) = E_{D,\mathsf{LOCC}(\mathrm{C{:}B})}(\beta)$$

giving the trivial upper bound

$$E_{DR}(\alpha,\beta) \leqslant \min\left\{E_D(\alpha), E_D(\beta)\right\}.$$

As anticipated this upper bound is matched by the minimum of the distillable entanglements being also an achievable rate

$$\min\left\{E_D(\alpha), E_D(\beta)\right\} \leqslant E_{DR}(\alpha,\beta).$$

Indeed if α and β are states, which we can always assume to be isotropic states that are ε-close to respectively Φ^e and $\Phi^{e'}$, coming from an entanglement distillation protocol between the links, then it is easily checked that the teleportation protocols yields a state 2ε-close to $\Phi^{\min\{e,e'\}}$[1]. Therefore if e and e' are achievable, $\min\{e,e'\}$ is also achievable.

Therefore, for the task of entanglement distillation using perfect operations, the repeater setting is not more difficult than entanglement distillation, and the optimal noise-free protocol for the quantum repeater is the one that performs entanglement distillation between Alice and Charlie, and between Charlie and Bob, followed by entanglement swapping. Because the rate of entanglement distillation with Charlie reduces to the bipartite case, initially repeaters have been studied only when modelling the noise also in the operations [Bri+98; Dür+99]. However we will see that this is not the case if we are tasked with distilling key across the repeater station.

[1] More precisely the steps are: embed the smaller one in the larger dimension (it will not be an isotropic state any more in general), use the larger isotropic state to teleport, map back down to the smaller dimension (the state will still not be isotropic in general). We can optionally twirl again to produce an isotropic state.

4.2 Repeater Key Distillation

The repeater key rates from [Bäu+15] are now defined changing the target states of the repeater distillable entanglement. Recall that we now have three parties: Alice and Charlie (A and C) share α and Charlie and Bob (D and B) share β. Tripartite PPT operations (whatever they are) will not be allowed, because they will be able to produce infinite amounts of key, as explained in Section 3.1. Thus we require the repeater operations to be contained in SEP. Let L be a class of tripartite operations included in SEP, then the best finite rate for key distillation under repeater operations is

$$K^{\varepsilon}_{R,\mathsf{L}}(\alpha,\beta) := \sup\left\{\log\kappa : \Upsilon(\Phi^{\log\kappa}, \mathrm{A}_{\text{🛡}}\mathrm{B}_{\text{🛡}}) \approx_{\varepsilon} \mathsf{L}(\mathbb{C}^{\kappa}\mathrm{A}_{\text{🛡}}{:}\mathrm{C}{:}\mathbb{C}^{\kappa}\mathrm{B}_{\text{🛡}}\langle \mathrm{A}{:}\mathrm{CD}{:}\mathrm{B}) \circ (\alpha\otimes\beta)\right\}, \quad (4.3)$$

where we recall from Section 3.1.3 that $\Upsilon(\Phi^{\log\kappa}, \mathrm{A}_{\text{🛡}}\mathrm{B}_{\text{🛡}})$ is the set of all private states with key size κ. The repeater distillable key is then defined as a the best achievable rate

$$K_{R,\mathsf{L}}(\rho) := \lim_{\varepsilon\to 0}\limsup_{n\to\infty}\frac{1}{n}K^{\varepsilon}_{R,\mathsf{L}}(\alpha^{\otimes n},\beta^{\otimes n}).$$

We reserve K_R for the repeater key rate defined by LOCC with two-way communication between the parties. A schematic LOCC repeater protocol is given in Figure 4.2.

As in the case of entanglement distillation by joining the parties we get the trivial upper bound

$$K_R(\alpha,\beta) \leqslant \min\{K_D(\alpha), K_D(\beta)\}.$$

This upper bound is the rate achieved by distilling the key between the nodes, and then using the one-time pad to send one of the keys from Charlie to the opposite party. As far as we known this is the only way to achieve this rate, and thus it requires a trusted repeater station Charlie, where the information held by Charlie is not traced out. Trusted repeater nodes are being considered [Sal+10], but while they might be viable options for private networks that have other security guarantees, they will provide no security in a public network. The information theoretic security guarantees are lost if we do not give the residual state of the repeater station to the eavesdropper.

The rate will of course be larger than the entanglement distillation rate

$$\min\{E_D(\alpha), E_D(\beta)\} = E_{DR}(\alpha,\beta) \leqslant K_R(\alpha,\beta)$$

But there are also instances where the inequality is strict, as a direct consequence of the difference between distillable key and distillable entanglement. In the case where $E_D(\alpha) \geqslant \log|\mathrm{D}|$, namely when one link can teleport the other fully, then the trivial teleportation protocol on β gives us

$$E_{DR}(\alpha,\beta) = E_D(\beta) \leqslant K_D(\beta) = K_R(\alpha,\beta)$$

thus displaying the same edge case behaviour as for distillable key and distillable entanglement. The repeater distillable entanglement and the repeater distillable key can be made arbitrarily different by teleporting private states with low distillable entanglement, and the repeater distillable key can be large even when the repeater distillable entanglement is zero by teleporting PPT approximate private states. Therefore the question about the difference between key and entanglement in the repeater setting has to be more subtle.

A way to avoid these extreme edge cases is to consider the case were the state in both links is the same. More precisely a well posed question is to ask whether there can a be strict inequalities in

$$E_D(\rho) = E_{DR}(\rho,\rho) \leqslant K_R(\rho,\rho) \leqslant K_D(\rho) \quad (4.4)$$

with the hope that there exists a state for which the first inequality is strict and arbitrarily large. So far, including the work presented in this thesis, we only have no-go results around perfect key, namley for ρ being either a private state or an approximate private state, indicating that there is no gap, and that the repeater distillable key collapses to the repeater distillable entanglement. Most importantly, these no-go results are not a consequence of some noise emerging from imperfect operations. They are information theoretic results about the impossibility for Charlie, to mediate correlations that are hidden to local observers.

In short, while for repeater entanglement distillation the active area of research studies the effect of noisy operations, for quantum repeater key there are open questions even with perfect operations. In this chapter we will see these no-go results, which are obtained by connecting distinguishability and repeater key distillation, and then applying the connection to the examples from of the previous chapter, namely to the undistillable private states and to the random private states. For the private states that have equal distillable entanglement and distillable key (the $BNOT$, locking and flower private states of Section 3.4) there is nothing particular to say, the trivial bounds on the repeater rate give

$$E_D(\gamma) = K_R(\gamma, \gamma) = K_D(\gamma) \tag{4.5}$$

Notice that this is for the two-way repeater key rate, but the same holds for the one-way repeater rate, as these private states are distillable with one-way LOCC. We will see the remaining private states in the upcoming sections.

4.2.1 Upper bounds

To show the no-go results, we will upper bound the repeater distillable key of private and approximate private states. To do this, we use some general upper bounds inspired from [Bäu+15]. The original upper bound on $K_R(\alpha, \beta)$ provided there is in terms of a regularized LOCC-restricted relative entropy distance to quadri-partite separable states [Bäu+15, Theorem 4]. For any states α on AC and β on DB

$$K_R(\alpha, \beta) \leqslant D^{\infty}_{\mathsf{LOCC}(\underline{\mathrm{A}}\underline{\mathrm{B}}:\underline{\mathrm{C}}\underline{\mathrm{D}})}(\alpha \otimes \beta \| \mathcal{S}(\mathrm{A{:}C{:}D{:}B})). \tag{4.6}$$

The intuition is that if Alice and Bob are joined together, the produced private state at the output looks like a maximally entangled state (because the twisting can be lifted with global operations), while if the input was separable then . It means that $\alpha \otimes \beta$ can be viewed as a bipartite state, namely if we now make $\alpha \in \mathcal{D}(\mathrm{A{:}C})$ and $\beta \in \mathcal{D}(\mathrm{A'C'})$, then the repeater protocols can be transferred to this setting and can be used to distinguish $\alpha \otimes \beta$ from $\mathcal{S}(\mathrm{A{:}A'{:}C{:}C'})$ by applying the protocol, untwisting and checking if the state at $\mathrm{AA'}$ is separable.

We can improve this upper bound with a simple observation. A crucial passage of the original proof is that no repeater protocol will ever output entanglement between Alice and Bob if the inputs are substituted with fully quadri-partite separable states. Formally the proof uses that

$$\mathcal{S}\big(\mathrm{A}^{\mathrm{out}}{:}\mathrm{B}^{\mathrm{out}}\big) \supseteq \mathsf{LOCC}\big(\mathrm{A}^{\mathrm{out}}{:}\mathbb{C}{:}\mathrm{B}^{\mathrm{out}}\big\langle \mathrm{A{:}CD{:}B}\big) \circ \mathcal{S}(\mathrm{A{:}C{:}D{:}B}).$$

However, only separability of the output is used, which holds already if we make a single input link is separable. Namely the inclusion still holds if instead we feed either $\mathcal{S}(\mathrm{A{:}CDB})$ or $\mathcal{S}(\mathrm{ACD{:}B})$ to the tripartite LOCC maps, thus these classes of separable states also give valid upper bounds. This works more generally for any class of states that gets mapped into separable states by the repeater protocols, as formalized by the bound below. To state it we need a more flexible regularization that arguably misses the point of the standard regularization, but we will be able to recover a standard regularization later after some relaxations.

Let L be a class of channels and $\mathcal{K}$ a class of states. For ease of notation define the maximal regularization of the L relative entropy from $\mathcal{K}$ as

$$D^{\overline{\infty}}_{\mathsf{L}(\mathrm{H})}(\rho\|\mathcal{K}(\mathrm{H})) := \limsup_{n\to\infty}\frac{1}{n}D_{\mathsf{L}(\mathrm{H}^{\otimes n})}(\rho^{\otimes n}\|\mathcal{K}(\mathrm{H}^{\otimes n})). \tag{4.7}$$

This will in general not be a useful definition for upper bounds, because if the normal regularization is not guaranteed, then estimating the sequence of relative entropies for finite number of copies gives no information about the regularization as an upper bound. However, in the following bounds where we use the max-regularization, we can always make a relaxation to a regularizable quantity.

Before we show the improvement mentioned above, we present another bound. We can prove a similar but new bound that does not join Alice and Bob in a single party. Instead of allowing global operations at Alice and Bob, we can simply let their system stay quantum in the distinguishability. We present it first because it is conceptually simpler.

Theorem 43 (Unpublished). *Let L be a class of repeater operations contained in SEP, and let $\mathcal{K}$ be any class of tripartite states such that it is mapped into separable bipartite states by the L repeater protocol, namely such that (for all systems)*

$$\mathcal{S}\big(\mathrm{A}^{\mathrm{out}}{:}\mathrm{B}^{\mathrm{out}}\big) \supseteq \mathsf{L}\big(\mathrm{A}^{\mathrm{out}}{:}\mathbb{C}{:}\mathrm{B}^{\mathrm{out}}\big\langle \mathrm{A}{:}\mathrm{C}{:}\mathrm{B}\big) \circ \mathcal{K}(\mathrm{A}{:}\mathrm{C}{:}\mathrm{B}).$$

Then, for any states α on AC *and β on* DB, *we have*

$$K_{R,\mathsf{L}}(\alpha,\beta) \leqslant D^{\overline{\infty}}_{\mathsf{L}(\mathrm{A}:\underline{\mathrm{C}}\underline{\mathrm{D}}:\mathrm{B})}(\alpha\otimes\beta\|\mathcal{K}(\mathrm{A}{:}\mathrm{CD}{:}\mathrm{B})).$$

Proof. Let $\sigma \in \mathcal{K}(\mathrm{A}^{\otimes n}{:}\mathrm{C}^{\otimes n}\mathrm{D}^{\otimes n}{:}\mathrm{B}^{\otimes n})$, let $\mathrm{A}_{\text{🔑}} = \mathrm{B}_{\text{🔑}} = \mathbb{C}^k$, and let

$$\Lambda \in \mathsf{L}(\mathrm{A}^{\otimes n}{:}\mathrm{C}^{\otimes n}\mathrm{D}^{\otimes n}{:}\mathrm{B}^{\otimes n})|\mathrm{A}_{\text{🔑}}\mathrm{A}_{\text{🛡}} : \mathbb{C} : \mathrm{B}_{\text{🔑}}\mathrm{B}_{\text{🛡}}) \subseteq \mathsf{L}(\mathrm{A}^{\otimes n}{:}\underline{\mathrm{C}}^{\otimes n}\underline{\mathrm{D}}^{\otimes n}{:}\mathrm{B}^{\otimes n}),$$

be a repeater protocol that distills the approximate private state $\tilde{\gamma}^{\log k}$. Then, by the assumption that σ is mapped into separable states by Λ, we have

$$\begin{aligned} D_{\mathsf{L}\left(\mathrm{A}^{\otimes n}:\underline{\mathrm{C}}^{\otimes n}\underline{\mathrm{D}}^{\otimes n}:\mathrm{B}^{\otimes n}\right)}(\alpha^{\otimes n}\otimes\beta^{\otimes n}\|\sigma) &\geqslant D(\Lambda(\alpha^{\otimes n}\otimes\beta^{\otimes n})\|\Lambda(\sigma)) \\ &\geqslant D(\tilde{\gamma}^{\log k}\|\mathcal{S}(\mathrm{A}_{\text{🔑}}\mathrm{A}_{\text{🛡}}{:}\mathrm{B}_{\text{🔑}}\mathrm{B}_{\text{🛡}})) \end{aligned}$$

Ideally, at this point we would further lower bound by the relative entropy of entanglement and use its asymptotic continuity to change $\tilde{\gamma}^{\log k}$ into $\gamma^{\log k}$ at the cost of some factor that goes to zero in the limits $n \to \infty$ and $\varepsilon \to 0$. However, the dimensions of the shield systems of $\gamma^{\log k}$ are in principle unbounded, so that we cannot argue directly that these factors go to zero. We need to remove the shield systems first, exploiting that by definition $\gamma^{\log k}$ is a twisted version of a maximally entangled state $\Phi^{\log k}$. This is the same argument used in [Hor+09, Theorem 9]. We give the statement and a contained proof below for completeness. For any state ρ that is ε-close to a private state γ^m,

$$E_R(\rho) \geqslant (1-2\varepsilon)m - g(\varepsilon). \tag{4.8}$$

Proof. Let us denote by Π the map that inverts the twisting, see Section 3.1.3. Namely the map that inverts the controlled unitary of the private state and traces out the shield, leaving a maximally entangled state. Then, by monotonicity of the trace distance, we have:

$$\Pi(\tilde{\gamma}^m) \approx_\varepsilon \Pi(\gamma^m) = \Phi^m.$$

Furthermore, while $\Pi \circ \mathcal{S}(\mathrm{A}_{🔑}\mathrm{A}_{🛡}{:}\mathrm{B}_{🔑}\mathrm{B}_{🛡})$ contains entangled states, it is still a convex set. We can apply Π by monotonicity of the relative entropy, thus we find that for any separable state $\sigma \in \mathcal{S}(\mathrm{A}_{🔑}\mathrm{A}_{🛡}{:}\mathrm{B}_{🔑}\mathrm{B}_{🛡})$:

$$\begin{aligned} D(\rho\|\sigma) &\geqslant D(\Pi(\rho)\|\Pi(\sigma)) \\ &\geqslant D(\Pi(\rho)\|\Pi \circ \mathcal{S}(\mathrm{A}_{🔑}\mathrm{A}_{🛡}{:}\mathrm{B}_{🔑}\mathrm{B}_{🛡})). \end{aligned}$$

We can now use the asymptotic continuity of the relative entropy from convex sets, as explained in Section 2.3.

$$D(\rho\|\sigma) \geqslant D(\Phi^m\|\Pi \circ \mathcal{S}(\mathrm{A}_{🔑}\mathrm{A}_{🛡}{:}\mathrm{B}_{🔑}\mathrm{B}_{🛡})) - \varepsilon 2m - g(\varepsilon)$$

It was then proven in [Hor+09, Lemma 7], that

$$D(\Phi^m\|\Pi \circ \mathcal{S}(\mathrm{A}_{🔑}\mathrm{A}_{🛡}{:}\mathrm{B}_{🔑}\mathrm{B}_{🛡})) \geqslant m$$

and thus

$$D(\rho\|\sigma) \geqslant (1-2\varepsilon)m - g(\varepsilon).$$

Taking the infimum over separable states ends the proof. □

By Equation (4.8) we thus have

$$\frac{1}{n} D_{\mathsf{L}(\mathrm{A}^{\otimes n}:\underline{\mathrm{C}}^{\otimes n}\underline{\mathrm{D}}^{\otimes n}:\mathrm{B}^{\otimes n})}(\alpha^{\otimes n} \otimes \beta^{\otimes n} \parallel \sigma) \geqslant (1-2\varepsilon)\frac{\log k}{n} - \frac{1}{n} g(\varepsilon).$$

Taking the supremum over all channels gives

$$\frac{1}{n} D_{\mathsf{L}(\mathrm{A}^{\otimes n}:\underline{\mathrm{C}}^{\otimes n}\underline{\mathrm{D}}^{\otimes n}:\mathrm{B}^{\otimes n})}(\alpha^{\otimes n} \otimes \beta^{\otimes n} \parallel \sigma) \geqslant (1-2\varepsilon)K^{\varepsilon}_{R,\mathsf{L}}(\alpha^{\otimes n}, \beta^{\otimes n}) - \frac{1}{n} g(\varepsilon).$$

Taking the limit supremum for $n \to \infty$ and the limit for $\varepsilon \to 0$ ends the proof. □

We now show the actual generalization of the original upper bound of Equation (4.6). We do not have examples that show a separation between this and the new bound, but they are in principle incomparable.

Theorem 44 ([Bäu+15]). *Let* L *be a class of repeater operations contained in* SEP*, and let* $\mathcal{K}$ *be any class of tripartite states such that it is mapped into separable bipartite states by the* L *repeater protocol, namely such that (for all systems)*

$$\mathcal{S}(\mathrm{A}^{\mathrm{out}}{:}\mathrm{B}^{\mathrm{out}}) \supseteq \mathsf{L}(\mathrm{A}^{\mathrm{out}}{:}\mathrm{C}{:}\mathrm{B}^{\mathrm{out}}\langle \mathrm{A}{:}\mathrm{C}{:}\mathrm{B}) \circ \mathcal{K}(\mathrm{A}{:}\mathrm{C}{:}\mathrm{B}).$$

Let L' *be a bipartite class of operations such that* $\mathsf{L}(\mathrm{A}{:}\mathrm{C}{:}\mathrm{B}) \subseteq \mathsf{L}'(\mathrm{AB}{:}\mathrm{C})$*, then, for any states* α *on* AC *and* β *on* DB*, we have*

$$K_{R,\mathsf{L}}(\alpha, \beta) \leqslant D^{\overline{\infty}}_{\mathsf{L}(\underline{\mathrm{A}}\underline{\mathrm{B}}:\underline{\mathrm{C}}\underline{\mathrm{D}})}(\alpha \otimes \beta \| \mathcal{K}(\mathrm{A}{:}\mathrm{CD}{:}\mathrm{B})).$$

The proof is exactly the same as the one given in [Bäu+15], with the difference that the regularization of the relative entropy is not guaranteed and thus the limit supremum must be used instead.

Proof. Let $\sigma \in \mathcal{K}(\mathrm{A}^{\otimes n}{:}\mathrm{C}^{\otimes n}\mathrm{D}^{\otimes n}{:}\mathrm{B}^{\otimes n})$, let $\mathrm{A}_{\text{🔑}} = \mathrm{B}_{\text{🔑}} = \mathbb{C}^k$, and let

$$\Lambda \in \mathsf{L}(\mathrm{A}^{\otimes n}{:}\mathrm{C}^{\otimes n}\mathrm{D}^{\otimes n}{:}\mathrm{B}^{\otimes n})|\mathrm{A}_{\text{🔑}}\mathrm{A}_{\text{🛡}} : \mathbb{C} : \mathrm{B}_{\text{🔑}}\mathrm{B}_{\text{🛡}}) \subseteq \mathsf{L}'(\mathrm{A}^{\otimes n}\mathrm{B}^{\otimes n}{:}\underline{\mathrm{C}}^{\otimes n}\underline{\mathrm{D}}^{\otimes n}),$$

be a repeater protocol that distill the approximate private state $\tilde{\gamma}^{\log k}$. Then let $\mathcal{M} \in \mathsf{M}(\underline{\mathrm{A}}_{\text{🔑}}\underline{\mathrm{B}}_{\text{🔑}}\underline{\mathrm{A}}_{\text{🛡}}\underline{\mathrm{B}}_{\text{🛡}})$ be a global measurement at Alice and Bob. We then have

$$\mathcal{M} \circ \Lambda \in \mathsf{L}'(\underline{\mathrm{A}}^{\otimes n}\underline{\mathrm{B}}^{\otimes n}{:}\underline{\mathrm{C}}^{\otimes n}\underline{\mathrm{D}}^{\otimes n}),$$

and

$$D_{\mathsf{L}'(\underline{\mathrm{A}}^{\otimes n}\underline{\mathrm{B}}^{\otimes n}:\underline{\mathrm{C}}^{\otimes n}\underline{\mathrm{D}}^{\otimes n})}(\alpha^{\otimes n} \otimes \beta^{\otimes n} \| \sigma) \geqslant D(\mathcal{M} \circ \Lambda(\alpha^{\otimes n} \otimes \beta^{\otimes n}) \| \mathcal{M} \circ \Lambda(\sigma)).$$

Since this holds for any measurement $\mathcal{M}$ we have

$$D_{\mathsf{L}'(\underline{\mathrm{A}}^{\otimes n}\underline{\mathrm{B}}^{\otimes n}:\underline{\mathrm{C}}^{\otimes n}\underline{\mathrm{D}}^{\otimes n})}(\alpha^{\otimes n} \otimes \beta^{\otimes n} \| \sigma) \geqslant D_{\mathsf{M}(\underline{\mathrm{A}}_{\text{🔑}}\underline{\mathrm{B}}_{\text{🔑}}\underline{\mathrm{A}}_{\text{🛡}}\underline{\mathrm{B}}_{\text{🛡}})}(\Lambda(\alpha^{\otimes n} \otimes \beta^{\otimes n}) \| \Lambda(\sigma)).$$

This time the lower bound on such measured relative entropy was proven in [Bäu+15, Lemma 3], showing that, because $\Lambda(\sigma)$ is separable,

$$D_{\mathsf{M}(\underline{\mathrm{A}}_{\text{🔑}}\underline{\mathrm{B}}_{\text{🔑}}\underline{\mathrm{A}}_{\text{🛡}}\underline{\mathrm{B}}_{\text{🛡}})}(\Lambda(\alpha^{\otimes n} \otimes \beta^{\otimes n}) \| \Lambda(\sigma)) \geqslant (1-\varepsilon)m - h(\varepsilon),$$

and thus

$$\frac{1}{n} D_{\mathsf{L}'(\underline{\mathrm{A}}^{\otimes n}\underline{\mathrm{B}}^{\otimes n}:\underline{\mathrm{C}}^{\otimes n}\underline{\mathrm{D}}^{\otimes n})}(\alpha^{\otimes n} \otimes \beta^{\otimes n} \parallel \sigma) \geqslant (1-\varepsilon)\frac{\log k}{n} - \frac{1}{n}h(\varepsilon).$$

Taking the supremum over all channels gives

$$\frac{1}{n} D_{\mathsf{L}'(\underline{\mathrm{A}}^{\otimes n}\underline{\mathrm{B}}^{\otimes n}:\underline{\mathrm{C}}^{\otimes n}\underline{\mathrm{D}}^{\otimes n})}(\alpha^{\otimes n} \otimes \beta^{\otimes n} \parallel \sigma) \geqslant (1-\varepsilon)K^{\varepsilon}_{R,\mathsf{L}}(\alpha^{\otimes n}, \beta^{\otimes n}) - \frac{1}{n}h(\varepsilon).$$

Taking the limit supremum for $n \to \infty$ and the limit for $\varepsilon \to 0$ ends the proof. □

As discussed in Section 2.4, only the Theorem 44 bound is known to be regularizable, and only for convex classes $\mathcal{K}$ that are closed under measurement operators in L, when L is closed under tensor product. This happens in particular for $\mathcal{S}(\mathrm{A{:}C{:}D{:}B})$, but fails already for $\mathcal{S}(\mathrm{A{:}CDB})$ and $\mathcal{S}(\mathrm{ACD{:}B})$. As promised we now make the regularizable relaxation. If L and $\mathcal{K}$ are closed under tensor product, it is easily verified that then

$$\begin{aligned} K_{R,\mathsf{L}}(\alpha, \beta) &\leqslant D^{\infty}_{\mathsf{L}'(\underline{\mathrm{A}}\underline{\mathrm{B}}:\underline{\mathrm{C}}\underline{\mathrm{D}})}(\alpha \otimes \beta \| \sigma), \\ K_{R,\mathsf{L}}(\alpha, \beta) &\leqslant D^{\infty}_{\mathsf{L}(\mathrm{A}:\underline{\mathrm{C}}\underline{\mathrm{D}}:\mathrm{B})}(\alpha \otimes \beta \| \sigma). \end{aligned} \tag{4.9}$$

where the existence of the regularization comes from super-additivity, see Section 2.4. Namely, if we are allowed to take tensor product states in the minimization, then we know that the regularization exists for all classes of channels closed under tensor product (see again Section 2.4). In particular we have for any separable state $\tilde{\alpha} \in \mathcal{S}(\mathrm{A{:}C})$ and $\tilde{\beta} \in \mathcal{S}(\mathrm{D{:}B})$

$$\begin{aligned} K_R(\alpha, \beta) &\leqslant D^{\infty}_{\mathsf{LOCC}(\underline{\mathrm{A}}\underline{\mathrm{B}}:\underline{\mathrm{C}}\underline{\mathrm{D}})}(\alpha \otimes \beta \| \tilde{\alpha} \otimes \beta), \\ K_R(\alpha, \beta) &\leqslant D^{\infty}_{\mathsf{LOCC}(\underline{\mathrm{A}}\underline{\mathrm{B}}:\underline{\mathrm{C}}\underline{\mathrm{D}})}(\alpha \otimes \beta \| \alpha \otimes \tilde{\beta}). \end{aligned}$$

These bounds solve the "factor of 2" issue about tightness of the original upper bound in some obvious simple cases. For example, suppose that both inputs are maximally entangled states, $\alpha = \beta = \Phi^m$, then the bounds are tight, both sides being equal to m, but the right hand side of Equation (4.6) yields $2m$. Intuitively, while the original bound measures the distinguishability of both input states from separable, the new bounds can measure only the distinguishability of a single state and consider the other one as an assisting resource to the measurement.

By both joining Alice and Bob and relaxing to partial measurements, the two bounds of Equation (4.9) have a common upper bound, which is what we will use in the next section. Namely, we have

$$\begin{matrix} D^{\overline{\infty}}_{\mathsf{L}'(\underline{\mathrm{A}}\underline{\mathrm{B}}:\underline{\mathrm{C}}\underline{\mathrm{D}})}(\alpha \otimes \beta \| \mathcal{K}(\mathrm{A{:}CD{:}B})) \\ D^{\overline{\infty}}_{\mathsf{L}(\mathrm{A}:\underline{\mathrm{C}}\underline{\mathrm{D}}:\mathrm{B})}(\alpha \otimes \beta \| \mathcal{K}(\mathrm{A{:}CD{:}B})) \end{matrix} \leqslant D^{\overline{\infty}}_{\mathsf{L}'(\mathrm{AB}:\underline{\mathrm{C}}\underline{\mathrm{D}})}(\alpha \otimes \beta \| \mathcal{K}(\mathrm{A{:}CD{:}B})).$$

If L and $\mathcal{K}$ are closed under tensor product, then

$$K_{R,\mathsf{L}}(\alpha, \beta) \leqslant D^{\infty}_{\mathsf{L}'(\mathrm{AB}:\underline{\mathrm{C}}\underline{\mathrm{D}})}(\alpha \otimes \beta \| \sigma). \tag{4.10}$$

Finally notice that the common lower relative entropy

$$D^{\overline{\infty}}_{\mathsf{L}(\underline{\mathrm{A}}:\underline{\mathrm{C}}\underline{\mathrm{D}}:\underline{\mathrm{B}})}(\alpha \otimes \beta \| \mathcal{K}(\mathrm{A{:}CD{:}B})) \leqslant \begin{matrix} D^{\overline{\infty}}_{\mathsf{L}'(\underline{\mathrm{A}}\underline{\mathrm{B}}:\underline{\mathrm{C}}\underline{\mathrm{D}})}(\alpha \otimes \beta \| \mathcal{K}(\mathrm{A{:}CD{:}B})), \\ D^{\overline{\infty}}_{\mathsf{L}(\mathrm{A}:\underline{\mathrm{C}}\underline{\mathrm{D}}:\mathrm{B})}(\alpha \otimes \beta \| \mathcal{K}(\mathrm{A{:}CD{:}B})). \end{matrix}$$

is the one that, as an upper bound, would really squeeze the repeater distillable key close to the repeater distillable entanglement, because it is the one that can drop if data-hiding states are produced at Alice and Bob. Still the known bounds allow us to join the results from the previous chapter, and transform the indistinguishability results into undistillability results in the repeater setting. This will be the content of the remainder of the thesis.

4.3 One-way repeater

The content of this section comes from [1].

We consider the one-way repeater where Charlie can send classical communication but cannot receive, while Alice and Bob can still communicate with each other (and cannot be done through Charlie any more) [Bäu+15]. We thus define $\mathsf{LOCC}_{\leftarrow\rightarrow}(\mathrm{A{:}C{:}B})$ as the minimal class of operations that contains the above communication, and consider the repeater distillable key defined by this class, for which we reserve the notation

$$K_R^{\leftarrow\rightarrow}(\alpha, \beta) := K_{R,\mathsf{LOCC}_{\leftarrow\rightarrow}(\mathrm{A{:}CD{:}B})}(\alpha, \beta)$$

when the context is clear. Joining Alice and Bob we find that, this class is contained in the bipartite one-way operations from Charlie to Alice/Bob

$$\mathsf{LOCC}_{\leftarrow\rightarrow}(\mathrm{A{:}C{:}B}) \subseteq \mathsf{LOCC}_{\rightarrow}(\mathrm{C{:}AB}).$$

and thus using Equation (4.10) we have

$$K_{R,\mathsf{LOCC}_{\leftarrow\rightarrow}(\mathrm{A{:}CD{:}B})}(\alpha, \beta) \leqslant D^{\infty}_{\mathsf{LOCC}_{\rightarrow}(\underline{\mathrm{C}}\underline{\mathrm{D}}:\mathrm{AB})}(\alpha \otimes \beta \| \sigma).$$

where σ is any appropriate state that becomes separable under any repeater distillation protocol. We can now use Equation (1.5) from Section 1.4.3, to simplify the above bound. Namely, we can lift the measurement at Alice and Bob from the bipartite one-way LOCC using monotonicity of the relative entropy, thus obtaining the following corollary.

Corollary 45. *For any states α on* AC *and β on* DB *and any separable state σ in $\mathcal{S}(\mathrm{A{:}CDB})$ or $\mathcal{S}(\mathrm{ACD{:}B})$:*

$$K_{R,\mathsf{LOCC}_{\leftarrow\rightarrow}(\mathrm{A{:}CD{:}B})}(\alpha, \beta) \leqslant D^{\infty}_{\mathsf{M}(\underline{\mathrm{C}}\underline{\mathrm{D}})\otimes \mathrm{id}_{\mathrm{AB}}}(\alpha \otimes \beta \| \sigma).$$

We now connect it to private states with the bound in terms of the distillable entanglement in Section 3.3. For our case Corollary 23 gives the following statement. Let $m = \log|\mathrm{A}_{\text{🔑}}| = \log|\mathrm{B}_{\text{🔑}}|$ and let ρ be any key-correlated state on $\mathrm{A}_{\text{🔑}}\mathrm{B}_{\text{🔑}}\mathrm{A}_{\text{🛡}}\mathrm{B}_{\text{🛡}}$ (with support only on the maximally correlated subspace of $\mathrm{A}_{\text{🔑}}\mathrm{B}_{\text{🔑}}$).

$$\begin{aligned} E_{D,\mathsf{LOCC}_{\rightarrow}(\mathrm{A}_{\text{🔑}}\mathrm{A}_{\text{🛡}}:\mathrm{B}_{\text{🔑}}\mathrm{B}_{\text{🛡}})}(\rho) &\geqslant \chi_{\mathsf{LOCC}_{\rightarrow}(\underline{\mathrm{A}}_{\text{🔑}}\underline{\mathrm{A}}_{\text{🛡}}:\mathrm{B}_{\text{🔑}}\mathrm{B}_{\text{🛡}})}\left(\left\{\tfrac{1}{2^m}, \mathcal{Z}^i_{\mathrm{B}_{\text{🔑}}}(\rho)\right\}\right) \\ &= \chi_{\mathsf{M}(\underline{\mathrm{A}}_{\text{🔑}}\underline{\mathrm{A}}_{\text{🛡}})\otimes\mathrm{id}_{\mathrm{B}_{\text{🔑}}\mathrm{B}_{\text{🛡}}}}\left(\left\{\tfrac{1}{2^m}, \mathcal{Z}^i_{\mathrm{B}_{\text{🔑}}}(\rho)\right\}\right) \end{aligned}$$

where the equality is due to $\mathsf{M}\otimes\mathrm{id}$ being both a restriction of $\mathsf{LOCC}_{\rightarrow}$ and a relaxation due to monotonicity of the Holevo information. We can now exploit the measurement being local to simplify this bounds. The result is the following theorem:

Theorem 46. *For any key-correlated state ρ on $\mathrm{A}_{\text{🔑}}\mathrm{B}_{\text{🔑}}\mathrm{A}_{\text{🛡}}\mathrm{B}_{\text{🛡}}$, it holds:*

$$D_{\mathsf{M}(\underline{\mathrm{A}}_{\text{🔑}}\underline{\mathrm{A}}_{\text{🛡}})\otimes\mathrm{id}_{\mathrm{B}_{\text{🔑}}\mathrm{B}_{\text{🛡}}}}(\rho\parallel\hat{\rho}) \leqslant E_D^{\rightarrow}(\rho) \tag{4.11}$$

$$D^{\infty}_{\mathsf{M}(\underline{\mathrm{A}}_{\text{🔑}}\underline{\mathrm{A}}_{\text{🛡}})\otimes\mathrm{id}_{\mathrm{B}_{\text{🔑}}\mathrm{B}_{\text{🛡}}}}(\rho\parallel\hat{\rho}) \leqslant E_D^{\rightarrow}(\rho) \tag{4.12}$$

If $\hat{\rho}$ is also separable then:

$$D^{\infty}_{\mathsf{M}(\underline{\mathrm{A}}_{\text{🔑}}\underline{\mathrm{A}}_{\text{🛡}})\otimes\mathrm{id}_{\mathrm{B}_{\text{🔑}}\mathrm{B}_{\text{🛡}}}}(\rho\|\hat{\rho}) = E_D^{\rightarrow}(\rho). \tag{4.13}$$

Proof. It is straightforward to check that $\mathcal{Z}^i_{\mathrm{B}_{\text{🔑}}}(\hat{\rho}) = \hat{\rho}$. Let $\mathcal{M}\in\mathsf{M}(\underline{\mathrm{A}}_{\text{🔑}}\underline{\mathrm{A}}_{\text{🛡}})$ be any measurement at Alice. Because the measurement is local at Alice, it commutes with the unitary $\mathcal{Z}^i_{\mathrm{B}_{\text{🔑}}}$ at Bob, thus we have:

$$D(\mathcal{M}\otimes\mathcal{Z}^i_{\mathrm{B}_{\text{🔑}}}(\rho)\|\mathcal{M}(\hat{\rho})) = D(\mathcal{M}\otimes\mathcal{Z}^i_{\mathrm{B}_{\text{🔑}}}(\rho)\|\mathcal{M}\otimes\mathcal{Z}^i_{\mathrm{B}_{\text{🔑}}}(\hat{\rho})) = D(\mathcal{M}(\rho)\|\mathcal{M}(\hat{\rho}))$$

where we omitted the identity maps and used the unitary invariance of the relative entropy. By Corollary 23 then:

$$\begin{aligned} E_D^{\rightarrow}(\rho) &\geqslant \chi_{\mathsf{M}(\underline{\mathrm{A}}_{\text{🔑}}\underline{\mathrm{A}}_{\text{🛡}})\otimes\mathrm{id}_{\mathrm{B}_{\text{🔑}}\mathrm{B}_{\text{🛡}}}}\left(\left\{\tfrac{1}{2^m}, \mathcal{Z}^i_{\mathrm{B}_{\text{🔑}}}(\rho)\right\}\right) \\ &= \sup_{\mathcal{M}\in\mathsf{M}(\underline{\mathrm{A}}_{\text{🔑}}\underline{\mathrm{A}}_{\text{🛡}})} \frac{1}{2^m}\sum_i D(\mathcal{M}\otimes\mathcal{Z}^i_{\mathrm{B}_{\text{🔑}}}(\rho)\|\mathcal{M}(\hat{\rho})) \\ &= \sup_{\mathcal{M}\in\mathsf{M}(\underline{\mathrm{A}}_{\text{🔑}}\underline{\mathrm{A}}_{\text{🛡}})} \frac{1}{2^m}\sum_i D(\mathcal{M}(\rho)\|\mathcal{M}(\hat{\rho})) \\ &= \sup_{\mathcal{M}\in\mathsf{M}(\underline{\mathrm{A}}_{\text{🔑}}\underline{\mathrm{A}}_{\text{🛡}})\otimes\mathrm{id}_{\mathrm{B}_{\text{🔑}}\mathrm{B}_{\text{🛡}}}} D(\mathcal{M}(\rho)\|\mathcal{M}(\hat{\rho})) \\ &= D_{\mathsf{M}(\underline{\mathrm{A}}_{\text{🔑}}\underline{\mathrm{A}}_{\text{🛡}})\otimes\mathrm{id}_{\mathrm{B}_{\text{🔑}}\mathrm{B}_{\text{🛡}}}}(\rho\parallel\hat{\rho}) \end{aligned}$$

proving Equation (4.11). For Equation (4.12) we have:

$$E_D^{\rightarrow}(\rho) = \frac{1}{n}E_D^{\rightarrow}(\rho^{\otimes n}) \geqslant \frac{1}{n}D_{\mathsf{M}(\underline{\mathrm{A}}_{\text{🔑}}\underline{\mathrm{A}}_{\text{🛡}})}(\rho^{\otimes n}\parallel\hat{\rho}^{\otimes n}) \qquad \forall\, n$$

because the distillable entanglement is already regularized and $\rho^{\otimes n}$ is still a key-correlated state. Since $D_{\mathsf{M}(\underline{\mathrm{A}}_{\text{🔑}}\underline{\mathrm{A}}_{\text{🛡}})}$ is super-additive, taking the limit $n\rightarrow\infty$ proves Equation (4.12). Equality in Equation (4.13) follows because if $\hat{\rho}$ is separable, then we get the opposite inequality from Equation (2.18):

$$\begin{aligned} D^{\infty}_{\mathsf{M}(\underline{\mathrm{A}}_{\text{🔑}}\underline{\mathrm{A}}_{\text{🛡}})}(\rho\|\hat{\rho}) &= D^{\infty}_{\mathsf{LOCC}_{\rightarrow}(\underline{\mathrm{A}}_{\text{🔑}}\underline{\mathrm{A}}_{\text{🛡}}:\mathrm{B}_{\text{🔑}}\mathrm{B}_{\text{🛡}})}(\rho\|\hat{\rho}) \\ &\geqslant D^{\infty}_{\mathsf{LOCC}_{\rightarrow}(\underline{\mathrm{A}}_{\text{🔑}}\underline{\mathrm{A}}_{\text{🛡}}:\mathrm{B}_{\text{🔑}}\mathrm{B}_{\text{🛡}})}(\rho\|\mathcal{S}(\mathrm{A}_{\text{🔑}}\mathrm{A}_{\text{🛡}}:\mathrm{B}_{\text{🔑}}\mathrm{B}_{\text{🛡}})) \\ &\geqslant E_D^{\rightarrow}(\rho) \\ &\geqslant D^{\infty}_{\mathsf{M}(\underline{\mathrm{A}}_{\text{🔑}}\underline{\mathrm{A}}_{\text{🛡}})}(\rho\|\hat{\rho}) \end{aligned}$$

where in the first equality we used the same argument as for the Holevo information. The upper bound on the distillable entanglement $D^{\infty}_{\mathsf{LOCC}_{\rightarrow}(\underline{\mathrm{A}}:\underline{\mathrm{B}})}(\rho\|\mathcal{S}(\mathrm{A{:}B}))$ for general states $\rho \in \mathcal{D}(\mathrm{AB})$ was proven in [LW14]. □

Namely, implicit in the theorem is the fact that the optimal $\mathsf{LOCC}_{\rightarrow}$ measurement is independent of the phase flip, which is what allowed us to remove the phase flip and regularize. We can now directly combine Theorem 46 with Corollary 45 in the corollary below by choosing $\sigma = \hat{\alpha} \otimes \hat{\beta}$, finally connecting the repeater distillable key with the bipartite distillable entanglement.

Corollary 47. *For any key-correlated states α and β with at least one separable key-attacked state, it holds:*

$$K_{R,\mathsf{LOCC}_{\leftarrow\rightarrow}(\mathrm{A{:}CD{:}B})}(\alpha,\beta) \leqslant E_{D,\mathsf{LOCC}_{\rightarrow}(\mathrm{CD{:}AB})}(\alpha \otimes \beta).$$

So far, the bound of Equation (4.6) could only be estimated via a relaxation that only works for states that are PPT. With this bound we can now show no-go results for the gap between the repeater distillable key and repeater distillable entanglement (Equation (4.4)), for states beyond the ones that are close to separable after partial transposition, like the Fourier private bits. Still because the only known bounds on distillable entanglement are for the PPT distillable entanglement, we cannot fully exploit Corollary 47 for states that are far from PPT. Up to technical factors, Corollary 47 does show though that in Equation (4.4) that the repeater distillable key is determined by the repeater distillable entanglement for key-correlated states with separable key-attacked state:

$$E_D^{\rightarrow}(\rho) \leqslant K_R^{\leftarrow\rightarrow}(\rho,\rho) \leqslant 2E_D^{\rightarrow}(\rho).$$

Examples

We can now give examples of private states with asymptotically zero repeater key, that were out of reach for [Bäu+15]. For the Swap private states and the Fourier private states we can immediately apply Corollary 47 to Corollary 26:

$$K_R^{\leftarrow\rightarrow}(\gamma_S,\gamma_S) \leqslant 2\log\left(1+\frac{1}{d}\right), \tag{4.14}$$

$$K_R^{\leftarrow\rightarrow}(\gamma_{\mathbb{U}^{\Gamma}},\gamma_{\mathbb{U}^{\Gamma}}) \leqslant 2\log\left(1+\frac{1}{\sqrt{d}}\right). \tag{4.15}$$

Since all private states are NPT (Non-positive under Partial Transposition) [Hor+09], this gives the first examples of NPT states with a high distillable key but low one-way repeater key rate. The swap private bit was implemented experimentally for $d = 2$ [Dob+11]. The key was distilled at a rate $K \approx 0.69$, enough to break the bound at $E_N(\gamma) = \log\frac{3}{2} \approx 0.58$. However, because of the factor of 2, an implementation with $d = 4$ at the same key rate is required for the same proof of concept. Still, scaling up the implementation should be experimentally feasible, since in $d = 4$ the swap operator is tensor product of qubit gates.

For the approximate private states, we cannot use the bound on distillable entanglement directly, because this is actually zero. Therefore we need to use monotonicity first to reduce to an actual private state, since the approximate private states we consider are obtained by mixing private states with $\mathsf{LOCC}_{\rightarrow}$ noise; then we can use Corollary 47:

$$\begin{aligned} K_R^{\leftarrow\rightarrow}(\tilde{\gamma}_{\mathbb{U}^{\Gamma}},\tilde{\gamma}_{\mathbb{U}^{\Gamma}}) &\leqslant 2\log\left(1+\frac{1}{\sqrt{d}}\right), \\ K_R^{\leftarrow\rightarrow}(\tilde{\gamma}_{\Gamma},\tilde{\gamma}_{\Gamma}) &\leqslant 2\frac{1+\log e}{1+\sqrt{d}}. \end{aligned} \tag{4.16}$$

Proof. For the PPT Fourier private states, using monotonicity from $\gamma_{\mathbb{U}^\Gamma}$ we get the same upper bound as in Equation (4.14).

For the PPT invariant private states we need to use monotonicity from γ_Γ, which was was shown in Lemma 29 to satisfy

$$E_{D,\mathsf{PPT}}(\gamma_\Gamma) \leqslant \frac{1+\log e}{1+\sqrt{d}}.$$

Now, we use one-way LOCC monotonicity of $K_R^{\leftarrow\rightarrow}$, Corollary 47 and the above bound, in this order, to show the claim:

$$K_R^{\leftarrow\rightarrow}(\tilde{\gamma}_\Gamma, \tilde{\gamma}_\Gamma) \leqslant K_R^{\leftarrow\rightarrow}(\gamma_\Gamma, \gamma_\Gamma) \leqslant 2E_D^{\rightarrow}(\gamma_\Gamma) \leqslant 2\frac{1+\log e}{1+\sqrt{d}} \qquad \square$$

However for the Fourier private bits this bound is not optimal. Until [1], the PPT Fourier private states were the only example in the literature for which the repeater key rate could be upper bounded by a computable quantity. Because Charlie is traced out at the end of the repeater key distillation protocol, the repeater key rate is invariant under transposition of the input on Charlie's systems, thus giving the following upper bound on the repeater key rate [Bäu+15]:

$$K_R(\alpha, \beta) \leqslant \min\{K_D(\alpha^\Gamma), K_D(\beta^\Gamma)\}. \tag{4.17}$$

The previous upper bound on $K_R(\tilde{\gamma}_{\mathbb{U}^\Gamma}, \tilde{\gamma}_{\mathbb{U}^\Gamma})$ was computed in [Bäu+15, Theorem 5] by estimating an upper bound on $K_D(\tilde{\gamma}_{\mathbb{U}^\Gamma}^\Gamma)$:

$$K_R(\tilde{\gamma}_{\mathbb{U}^\Gamma}, \tilde{\gamma}_{\mathbb{U}^\Gamma}) \leqslant O\left(\frac{\log d}{\sqrt{d}}\right).$$

which is still not optimal; used properly, Equation (4.17) yields the improved bound:

$$K_R(\tilde{\gamma}_{\mathbb{U}^\Gamma}, \tilde{\gamma}_{\mathbb{U}^\Gamma}) \leqslant \frac{1}{\sqrt{d}+1}.$$

Proof. By Equation (4.17) and convexity of the relative entropy of entanglement we find:

$$\begin{aligned} K_R(\tilde{\gamma}_{\mathbb{U}^\Gamma}, \tilde{\gamma}_{\mathbb{U}^\Gamma}) &\leqslant K_D(\tilde{\gamma}_{\mathbb{U}^\Gamma}^\Gamma) \leqslant E_R(\tilde{\gamma}_{\mathbb{U}^\Gamma}^\Gamma) \\ &\leqslant \frac{1}{1+\frac{1}{\sqrt{d}}}\left(E_R(\hat{\gamma}_{\mathbb{U}^\Gamma}) + \frac{1}{\sqrt{d}}E_R(\gamma_{\mathbb{U}})\right). \end{aligned}$$

However $\hat{\gamma}_{\mathbb{U}^\Gamma}$ is separable and, according to Equation (3.22), $E_R(\gamma_{\mathbb{U}}) = 1$. Therefore:

$$K_R(\tilde{\gamma}_{\mathbb{U}^\Gamma}, \tilde{\gamma}_{\mathbb{U}^\Gamma}) \leqslant \frac{1}{\sqrt{d}+1}\,. \qquad \square$$

This bound is still better than Equation (4.16) and holds for two-way protocols, but it will not work for the PPT invariant state. Corollary 47 only holds for private states, but we have shown how to use it for states that are close to private states, if the noise is local. We can extend the bound to all states if instead of restricting the input states, we restrict the distillation protocols. This is the content of the next section.

4.3.1 Key-swapping

Recall that we call strictly irreducible the private states with separable key-attacked state. For the purpose of this section let us denote such states with $\langle\gamma\rangle$ or $\gamma^{\langle m\rangle}$.We now show that the one-way distillable entanglement upper bound of Corollary 47 gives an upper bound for all states, not just key-correlated ones, for all protocols that try to distill key across the repeater by first distilling perfect key between the links in the form of strictly irreducible private states shared with Charlie. First, we define the following rate for such protocols.

Definition 48. *For all states α on* AC *and $\tilde{\alpha}$ on* $\tilde{\mathrm{C}}$B, *we define the one-way key-swapping distillable key as the best achievable rate:*

$$K_S(\alpha,\tilde{\alpha}) := \lim_{\delta\to 0}\lim_{\epsilon,\tilde{\epsilon}\to 0}\limsup_{n\to\infty}\frac{1}{n}K_S^{\delta,\epsilon,\tilde{\epsilon}}(\alpha^{\otimes n},\tilde{\alpha}^{\otimes n})$$

$$K_S^{\delta,\epsilon,\tilde{\epsilon}}(\alpha,\tilde{\alpha}) := \sup\left\{K\colon \begin{array}{c}\Lambda(\gamma^{\langle r\rangle}\otimes\gamma^{\langle\tilde{r}\rangle})\approx_\delta \gamma^K\\ \Gamma(\alpha)\approx_\epsilon \gamma^{\langle r\rangle},\quad \Gamma'(\tilde{\alpha})\approx_{\tilde{\epsilon}}\gamma^{\langle\tilde{r}\rangle}\end{array}\right\},$$

where the supremum is over dimensions such that

$$\begin{aligned}\Gamma &\in \mathsf{LOCC}_{\leftarrow}(\mathrm{A{:}C}\rangle \mathrm{A}^{\mathrm{in}}_{\text{🔑}}\mathrm{A}^{\mathrm{in}}_{\text{🛡}}{:}\mathrm{C}^{\mathrm{in}}_{\text{🔑}}\mathrm{C}^{\mathrm{in}}_{\text{🛡}})\\ \Gamma' &\in \mathsf{LOCC}_{\rightarrow}(\tilde{\mathrm{C}}{:}\mathrm{B}\rangle \tilde{\mathrm{C}}^{\mathrm{in}}_{\text{🔑}}\tilde{\mathrm{C}}^{\mathrm{in}}_{\text{🛡}}{:}\mathrm{B}^{\mathrm{in}}_{\text{🔑}}\mathrm{B}^{\mathrm{in}}_{\text{🛡}})\\ \Lambda &\in \mathsf{LOCC}_{\leftarrow\rightarrow}(\mathrm{A}^{\mathrm{in}}_{\text{🔑}}\mathrm{A}^{\mathrm{in}}_{\text{🛡}}{:}\mathrm{C}^{\mathrm{in}}_{\text{🔑}}\mathrm{C}^{\mathrm{in}}_{\text{🛡}}\tilde{\mathrm{C}}^{\mathrm{in}}_{\text{🔑}}\tilde{\mathrm{C}}^{\mathrm{in}}_{\text{🛡}}{:}\mathrm{B}^{\mathrm{in}}_{\text{🔑}}\mathrm{B}^{\mathrm{in}}_{\text{🛡}}\rangle \mathrm{A}^{\mathrm{out}}_{\text{🔑}}\mathrm{A}^{\mathrm{out}}_{\text{🛡}}{:}\mathbb{C}{:}\mathrm{B}^{\mathrm{out}}_{\text{🔑}}\mathrm{B}^{\mathrm{out}}_{\text{🛡}}).\end{aligned}$$

Notice that the restriction to strictly irreducible private states is only in the intermediate step, and not at the output on Alice and Bob. To write the rate concisely we had to explicit the optimization over maps. The supremum has to be taken over all dimensions, such that there exist maps and systems fulfilling the conditions inside the finite rate. The restricted protocols still include one-way entanglement distillation and swapping; thus, the new repeater key rate is still lower bounded by the minimum of the one-way distillable entanglements. With the definition in place we can now show that such rate is bounded by the distillable entanglement unconditionally.

Theorem 49. *For all input states α on* AC *and $\tilde{\alpha}$ on* $\tilde{\mathrm{C}}$B, *we have*

$$K_S(\alpha,\tilde{\alpha}) \leqslant E_D^{\rightarrow}(\alpha\otimes\tilde{\alpha}).$$

Proof. First notice that the tensor product of two strictly irreducible private states is still a strictly irreducible private state, namely $\gamma^{\langle a\rangle}\otimes\gamma^{\langle b\rangle} = \gamma^{\langle a+b\rangle}$.

Then, for the sake of the proof, let us introduce the following convenient bold shorthand notation:

$$\boldsymbol{r} := r+\tilde{r} \qquad\qquad \boldsymbol{\alpha} := \alpha\otimes\tilde{\alpha}$$

$$\boldsymbol{\gamma}^{\langle\boldsymbol{r}\rangle} := \gamma^{\langle r\rangle}\otimes\gamma^{\langle\tilde{r}\rangle} \qquad\qquad \boldsymbol{\Gamma} := \Gamma\otimes\Gamma'$$

then with $\boldsymbol{\epsilon} := \epsilon+\tilde{\epsilon}$ we have

$$\boldsymbol{\Gamma}(\boldsymbol{\alpha}^{\otimes n}) \approx_{\boldsymbol{\epsilon}} \boldsymbol{\gamma}^{\langle\boldsymbol{r}\rangle}. \tag{4.18}$$

We also define:

$$\boldsymbol{\mathcal{E}}^{\mathbf{Bell}} = \mathcal{E}^{\mathrm{Bell}}_{\mathrm{A}^{\mathrm{in}}_{\text{🔑}}\mathrm{C}^{\mathrm{in}}_{\text{🔑}}}\otimes \mathrm{id}_{\mathrm{A}^{\mathrm{in}}_{\text{🛡}}\mathrm{C}^{\mathrm{in}}_{\text{🛡}}}\otimes\mathcal{E}^{\mathrm{Bell}}_{\mathrm{B}^{\mathrm{in}}_{\text{🔑}}\tilde{\mathrm{C}}^{\mathrm{in}}_{\text{🔑}}}\otimes \mathrm{id}_{\mathrm{B}^{\mathrm{in}}_{\text{🛡}}\tilde{\mathrm{C}}^{\mathrm{in}}_{\text{🛡}}}$$

where $\mathcal{E}^{\mathrm{Bell}}$ is the reversible map of Lemma 19.

Proof idea: just like in the proof of Lemma 22, we use the distillable entanglement as an upper bound on the rate of the protocols that perform a measurement on the shield followed by hashing. However, we need to insert a step in the proof to substitute the approximate private state with exact private state, otherwise the proofs of Corollary 23 and Theorem 46 do not work. We will do this at the level of the coherent information using its asymptotic continuity.

First, let $\boldsymbol{\Gamma}$ be an intermediate one-way distillation protocol that gets $\boldsymbol{\epsilon}$-close to strictly irreducible private states $\gamma^{\langle \boldsymbol{r}\rangle}$. We apply the reversible protocol on the output of $\boldsymbol{\Gamma}$

$$E_D^{\to}(\boldsymbol{\alpha}) = \frac{1}{n}E_D^{\to}(\boldsymbol{\alpha}^{\otimes n}) \geqslant \frac{1}{n}E_D^{\to}(\boldsymbol{\mathcal{E}}^{\mathbf{Bell}} \circ \boldsymbol{\Gamma}(\boldsymbol{\alpha}^{\otimes n})).$$

Here we used that $E_D^{\to}$ is an $\mathsf{LOCC}_{\to}$ monotone, and that

$$\boldsymbol{\mathcal{E}}^{\mathbf{Bell}} \circ \boldsymbol{\Gamma} \in \mathsf{LOCC}_{\to}\big(\mathrm{C}^{\otimes n}{:}\mathrm{A}^{\otimes n}\big\rangle \mathrm{C}_{\text{🔑}}^{\mathrm{in}\otimes 2}\mathrm{C}_{\text{🛡}}^{\mathrm{in}}{:}\mathrm{A}_{\text{🔑}}^{\mathrm{in}\otimes 2}\mathrm{A}_{\text{🛡}}^{\mathrm{in}}\big) \otimes \mathsf{LOCC}_{\to}\big(\tilde{\mathrm{C}}^{\otimes n}{:}\mathrm{B}^{\otimes n}\big\rangle \tilde{\mathrm{C}}_{\text{🔑}}^{\mathrm{in}\otimes 2}\tilde{\mathrm{C}}_{\text{🛡}}^{\mathrm{in}}{:}\mathrm{B}_{\text{🔑}}^{\mathrm{in}\otimes 2}\mathrm{B}_{\text{🛡}}^{\mathrm{in}}\big)$$

where at the output of the reversible map there are two copies of the original key systems. Let now

$$\mathcal{M} \in \mathsf{M}\big(\mathrm{C}_{\text{🔑}}^{\mathrm{in}}\mathrm{C}_{\text{🛡}}^{\mathrm{in}}\tilde{\mathrm{C}}_{\text{🔑}}^{\mathrm{in}}\tilde{\mathrm{C}}_{\text{🛡}}^{\mathrm{in}}\big\rangle \mathrm{M}\big) \otimes \mathrm{id}_{\mathrm{C}_{\text{🔑}}^{\mathrm{in}}\tilde{\mathrm{C}}_{\text{🔑}}^{\mathrm{in}}\mathrm{A}_{\text{🔑}}^{\mathrm{in}\otimes 2}\mathrm{A}_{\text{🛡}}^{\mathrm{in}}\mathrm{B}_{\text{🔑}}^{\mathrm{in}\otimes 2}\mathrm{B}_{\text{🛡}}^{\mathrm{in}}}$$

be the measurement at Charlie, in the final one-way protocol that distils δ-close to a private state γ^K, (by Equation (1.5) the one-way protocol that ends with tracing out Charlie can be written as a measurement at Charlie followed by a global operation). After applying $\mathcal{M}$, we lower bound the one-way distillable entanglement with the coherent information, just like in Corollary 23:

$$E_D^{\to}(\boldsymbol{\alpha}) \geqslant \frac{1}{n}I\big(\mathrm{C}_{\text{🔑}}^{\mathrm{in}}\tilde{\mathrm{C}}_{\text{🔑}}^{\mathrm{in}}\big\rangle \mathrm{A}_{\text{🔑}}^{\mathrm{in}\otimes 2}\mathrm{A}_{\text{🛡}}^{\mathrm{in}}\mathrm{M}\mathrm{B}_{\text{🔑}}^{\mathrm{in}\otimes 2}\mathrm{B}_{\text{🛡}}^{\mathrm{in}}\big)_{\mathcal{M}\circ\boldsymbol{\mathcal{E}}^{\mathbf{Bell}}\circ\boldsymbol{\Gamma}(\boldsymbol{\alpha}^{\otimes n})}.$$

Here is where we want to change the approximate private state $\boldsymbol{\Gamma}(\boldsymbol{\alpha}^{\otimes n})$ into the exact private state $\gamma^{\langle \boldsymbol{r}\rangle}$ as mentioned before. In the form of [Win16] the asymptotic continuity of the coherent information [AF04], which we have seen before, says that:

$$|I(\mathrm{H}\rangle \mathrm{H}')_{\varrho} - I(\mathrm{H}\rangle \mathrm{H}')_{\tilde{\varrho}}| \leqslant 2\varepsilon \log |\mathrm{H}| + 2g(\varepsilon)$$

for arbitrary ε-close states $\varrho, \tilde{\varrho} \in \mathcal{D}(\mathrm{H}\mathrm{H}')$. Since combining Equation (4.18) and the monotonicity of the trace distance gives

$$\mathcal{M} \circ \boldsymbol{\mathcal{E}}^{\mathbf{Bell}}(\boldsymbol{\Gamma}(\boldsymbol{\alpha}^{\otimes n})) \approx_{\boldsymbol{\epsilon}} \mathcal{M} \circ \boldsymbol{\mathcal{E}}^{\mathbf{Bell}}(\gamma^{\langle \boldsymbol{r}\rangle}),$$

we can use the asymptotic continuity, where the dimension factor is now $\log |\mathrm{C}_{\text{🔑}}^{\mathrm{in}}\tilde{\mathrm{C}}_{\text{🔑}}^{\mathrm{in}}| = \boldsymbol{r}$. Notice that it is necessary to perform the measurement first, as to remove the unbounded shield systems from entering the dimension factor in the asymptotic continuity. We then get:

$$\begin{aligned} E_D^{\to}(\boldsymbol{\alpha}) &\geqslant \frac{1}{n}I\big(\mathrm{C}_{\text{🔑}}^{\mathrm{in}}\tilde{\mathrm{C}}_{\text{🔑}}^{\mathrm{in}}\big\rangle \mathrm{A}_{\text{🔑}}^{\mathrm{in}\otimes 2}\mathrm{A}_{\text{🛡}}^{\mathrm{in}}\mathrm{M}\mathrm{B}_{\text{🔑}}^{\mathrm{in}\otimes 2}\mathrm{B}_{\text{🛡}}^{\mathrm{in}}\big)_{\mathcal{M}\circ\boldsymbol{\mathcal{E}}^{\mathbf{Bell}}(\gamma^{\langle \boldsymbol{r}\rangle})} - \frac{1}{n}(2\boldsymbol{\epsilon}\boldsymbol{r} + g(\boldsymbol{\epsilon})) \\ &= \frac{1}{n}I\big(\mathrm{C}_{\text{🔑}}^{\mathrm{in}}\tilde{\mathrm{C}}_{\text{🔑}}^{\mathrm{in}}\big\rangle \mathrm{A}_{\text{🔑}}^{\mathrm{in}\otimes 2}\mathrm{A}_{\text{🛡}}^{\mathrm{in}}\mathrm{M}\mathrm{B}_{\text{🔑}}^{\mathrm{in}\otimes 2}\mathrm{B}_{\text{🛡}}^{\mathrm{in}}\big)_{\mathcal{M}\circ\boldsymbol{\mathcal{E}}^{\mathbf{Bell}}(\gamma^{\langle \boldsymbol{r}\rangle})} + O\Big(\frac{\boldsymbol{r}\boldsymbol{\epsilon}}{n}\Big). \end{aligned}$$

Now, as shown in Corollary 23 and Theorem 46, we can rewrite the conditional information as a relative entropy and then, by the unitary invariance of the relative entropy, correct the phase flip on the Alice/Bob side of $\gamma^{\langle \boldsymbol{r}\rangle}$. This results in:

$$E_D^{\to}(\boldsymbol{\alpha}) \geqslant \frac{1}{n}D(\mathcal{M}(\gamma^{\langle \boldsymbol{r}\rangle}) \parallel \mathcal{M}(\hat{\gamma}^{\langle \boldsymbol{r}\rangle})) + O\Big(\frac{\boldsymbol{r}\boldsymbol{\epsilon}}{n}\Big).$$

We can now finish the distillation protocol using monotonicity of the relative entropy. Let

$$\Lambda \in \mathsf{LOCC}\big(\mathrm{A}_{\text{🔑}}^{\mathrm{in}\otimes 2}\mathrm{A}_{\text{🛡}}^{\mathrm{in}}\mathrm{M}{:}\mathrm{B}_{\text{🔑}}^{\mathrm{in}\otimes 2}\mathrm{B}_{\text{🛡}}^{\mathrm{in}}\big\rangle \mathrm{A}_{\text{🔑}}^{\mathrm{out}}\mathrm{A}_{\text{🛡}}^{\mathrm{out}}{:}\mathrm{B}_{\text{🔑}}^{\mathrm{out}}\mathrm{B}_{\text{🛡}}^{\mathrm{out}}\big)$$

be the completion of the $\mathsf{LOCC}_{\leftarrow\rightarrow}$ protocol that distills δ-close to a private state γ^K, where we have given the measurement outcome to Alice without loss of generality because Alice and Bob can communicate with each other:

$$E_D^{\rightarrow}(\boldsymbol{\alpha}) \geqslant \frac{1}{n} D(\Lambda \circ \mathcal{M}(\gamma^{\langle \boldsymbol{r} \rangle}) \parallel \Lambda \circ \mathcal{M}(\hat{\gamma}^{\langle \boldsymbol{r} \rangle})) + O\Big(\frac{\boldsymbol{r}\boldsymbol{\epsilon}}{n}\Big).$$

Notice that $\Lambda \circ \mathcal{M}(\hat{\gamma}^{\langle \boldsymbol{r} \rangle})$ is a separable state in $\mathcal{S}(\mathrm{A}_{\text{🔑}}^{\mathrm{out}}\mathrm{A}_{\text{🛡}}^{\mathrm{out}}{:}\mathrm{B}_{\text{🔑}}^{\mathrm{out}}\mathrm{B}_{\text{🛡}}^{\mathrm{out}})$, because we distilled to strictly irreducible private states.

At this point using Equation (4.8), because $\Lambda \circ \mathcal{M}(\gamma^{\langle \boldsymbol{r} \rangle}) \approx_\delta \gamma^K$, we can lower bound the relative entropy as follows:

$$E_D^{\rightarrow}(\boldsymbol{\alpha}) \geqslant \frac{1}{n}((1-2\delta)K - g(\delta)) + O\Big(\frac{\boldsymbol{r}\boldsymbol{\epsilon}}{n}\Big)$$

Taking the supremum over all the distillation maps used we get

$$E_D^{\rightarrow}(\boldsymbol{\alpha}) \geqslant (1-2\delta)\frac{1}{n} K_S^{\delta,\epsilon,\tilde{\epsilon}}(\alpha^{\otimes n}, \tilde{\alpha}^{\otimes n}) - \frac{1}{n} 2\boldsymbol{\epsilon}\boldsymbol{r} - \frac{1}{n}(g(\delta) + g(\boldsymbol{\epsilon}))$$

for all number of copies n, security parameters δ, $\boldsymbol{\epsilon}$ and rate $\boldsymbol{r}$. Because in the intermediate private states we cannot distill more than if we had maximally entangled states as input, therefore $\boldsymbol{r} \leqslant n \log |\mathrm{AB}|$ and taking $\limsup$ for $n \to \infty$ we get

$$E_D^{\rightarrow}(\boldsymbol{\alpha}) \geqslant (1-2\delta) \limsup_{n\to\infty} \frac{1}{n} K_S^{\delta,\epsilon,\tilde{\epsilon}}(\alpha^{\otimes n}, \tilde{\alpha}^{\otimes n}) - 2\boldsymbol{\epsilon} \log |\mathrm{AB}|$$

Taking the limits $\boldsymbol{\epsilon} \to 0$ and $\delta \to 0$ concludes the proof. □

This section was dedicated to show that the one-way repeater distillable key is upper bounded by some form of distillable entanglement for private states, and thus to show that the key is not by itself a useful resource that can be transmitted across a repeater station, that resource is the distillable entanglement. We even generalized it to specific distillation protocols, that we called key-swapping, so that we could generalize to all input states. However, we still hit the limitation of upper bounding the distillable entanglement of the input states, which can currently be done only with the very loose relaxation to the PPT distillable entanglement, and thus works only for states close to PPT. To overcome this in the next section we will try to bound the repeater rate without the help of the distillable entanglement, using the random private states from Section 3.5 in a way that should apply to states that are very distillable under PPT operations.

4.4 Bounded repeater

Like [Bäu+15] and [1], we consider another variation of the quantum repeater key rate, namely the bounded-repeater key rate K_B. This will allow us to avoid the regularization in the bound on the repeater distillable key, and thus avoid the distillable entanglement and directly use the single copy estimates of Section 3.5. The operational interpretation for the bounded repeater goes as follows: instead of letting Charlie act jointly on arbitrarily many copies of the input, we restrict him to act only on one. This should model, for example, bounded-memory repeater stations that can only act on a finite number of copies at the same time (this should include for example memory-less repeater stations that perform their operations fast and do not allow for distillation). Alice and Bob then proceed to distill key as usual, and still apply their distillation protocols without restriction on the outcomes of the operations with Charlie. We make Charlie perform the same joint tripartite operation Λ on $\alpha \otimes \beta$ and we remove him after that, the resulting remaining state is $\Lambda(\alpha\otimes\beta)^{\otimes n}$, on which we let Alice and Bob perform any LOCC operation. Because the optimization over the tripartite channels commutes with the optimization over the bipartite channels, for any states α on AC and β on DB, we define the bounded-repeater finite key rate as

$$K_B^{\varepsilon,n}(\alpha,\beta) := \sup_{\Lambda\in\mathrm{LOCC}(\mathrm{A}:\underline{\mathrm{C}}\underline{\mathrm{D}}:\mathrm{B})} K_D^{\varepsilon}\left([\Lambda(\alpha\otimes\beta)]^{\otimes n}\right).$$

If we leave implicit that the measurement outcomes at Charlie are given to either Alice or Bob, we can simply write

$$K_B^{\varepsilon,n}(\alpha,\beta) := \sup_{\Lambda\in\mathrm{LOCC}(\mathrm{A}:\underline{\mathrm{C}}\underline{\mathrm{D}}:\mathrm{B})} K_D^{\varepsilon}\left([\Lambda(\alpha\otimes\beta)]^{\otimes n}\right).$$

Note that this time the expression of the finite rate depends on the number of copies. Namely, the finite rate for n copies is not of the form $f(\rho^{\otimes n})$ for some n-independent function f. Still the definition of achievable rate is that for all ε, there exist a sequence of protocols with finite rate limiting to it as the number of copies goes to infinity, and does not rely on the finite rate having the form $f(\rho^{\otimes n})$. The bounded-repeater key rate is defined as the best achievable rate for the bounded-repeater finite key rate above

$$K_B(\alpha,\beta) := \lim_{\varepsilon\to 0}\limsup_{n\to\infty}\frac{1}{n}K_B^{\varepsilon,n}(\alpha,\beta)$$

In [Bäu+15] and [1], the bounded-repeater key rate was defined using the equivalent expression in Lemma 50 below, which will be more convenient to use later. Lemma 50 rewrites $K_B(\alpha,\beta)$ as an optimization over bipartite distillable keys.

Lemma 50. *For any states α on* AC *and β on* DB*, we have*

$$K_B(\alpha,\beta) = \sup_{\Lambda\in\mathrm{LOCC}(\mathrm{A}:\underline{\mathrm{C}}\underline{\mathrm{D}}:\mathrm{B})} K_D(\Lambda(\alpha\otimes\beta)).$$

Proof. The inequality $\sup_\Lambda K_D(\Lambda(\alpha\otimes\beta)) \leqslant K_B(\alpha,\beta)$ comes from exchanging the limits with the supremum, so let us now concentrate on showing the other direction. By definition, $K_B(\alpha,\beta)$ is an achievable rate and thus for all $\varepsilon,\delta>0$, there exists Λ, Π, n and $R > K_B(\alpha,\beta)-\delta$ such that

$$\Pi\left(\Lambda(\alpha\otimes\beta)^{\otimes n}\right) \approx_\varepsilon \gamma^{nR}.$$

However, it was shown in [Hor+09, Lemma 6, pg 26] that if $\sigma \approx_\varepsilon \gamma^k$, then $K_D(\sigma) \geqslant (1-4\varepsilon)k-(2-\varepsilon)h(\varepsilon)$, where we used the improved bounds from [Win16; Shi17] (as the proof uses the continuity bounds to estimate the Devetak-Winter rate of σ). This means that we have

$$\begin{aligned} nR - 4\varepsilon nR - 2h(\varepsilon) &\leqslant K_D\left(\Pi\left(\Lambda(\alpha\otimes\beta)^{\otimes n}\right)\right) \\ &\leqslant K_D\left(\Lambda(\alpha\otimes\beta)^{\otimes n}\right) \\ &= nK_D(\Lambda(\alpha\otimes\beta)). \end{aligned}$$

Thus for all $\varepsilon,\delta>0$ there exists Λ such that

$$K_D(\Lambda(\alpha\otimes\beta)) \geqslant (1-4\varepsilon)K_B(\alpha,\beta)-\delta-2h(\varepsilon)$$

which implies that $\sup_\Lambda K_D(\Lambda(\alpha\otimes\beta)) \geqslant K_B(\alpha,\beta)$. □

The interpretation of Lemma 50 is that the optimal operations performed with the repeater station do not depend on the finite parameters of the protocol used by Alice and Bob to distill the key. We now move onto proving the upper bounds we will use to argue that the bounded-repeater distillable key of random private states is small.

4.4.1 Upper bounds

The content of this section comes from [1].

It is possible to prove an upper bound in terms of a single copy relative entropy measure. Namely, while the general upper bound of Theorem 43 for the repeater key rate is regularized, for the bounded-repeater rate it is possible to prove an un-regularized upper bound, which is much easier to deal with. The price to pay is that it is only in terms of a restriction to partial measurements, and cannot be reduced to a restriction on full measurements.

Theorem 51 ([Bäu+15; 1; 2]). *Let $\mathcal{K} \in \mathcal{D}(\mathrm{ACDB})$ be a set of states such that*

$$\mathsf{LOCC}\big(\mathrm{A{:}CD{:}B}\big\rangle \mathrm{A}^{\mathrm{out}}{:}\mathbb{C}{:}\mathrm{B}^{\mathrm{out}}\big) \circ \mathcal{K} \subseteq \mathcal{S}\big(\mathrm{A}^{\mathrm{out}}{:}\mathrm{B}^{\mathrm{out}}\big).$$

Then, for any states α on AC and β on DB, we have

$$K_B(\alpha, \beta) \leqslant D_{\mathsf{LOCC}(\mathrm{A{:}\underline{C}\underline{D}{:}B})}(\alpha \otimes \beta \| \mathcal{K}).$$

Proof. By assumption, for any $\Lambda \in \mathsf{LOCC}(\mathrm{A{:}CD{:}B}\rangle \mathrm{A}^{\mathrm{out}}{:}\mathbb{C}{:}\mathrm{B}^{\mathrm{out}})$ and $\sigma \in \mathcal{K}$, we have $\Lambda(\sigma) \in \mathcal{S}(\mathrm{A{:}B})$. Therefore, using Equation (1.6) for any such σ,

$$\begin{aligned} D_{\mathsf{LOCC}(\mathrm{A{:}\underline{C}\underline{D}{:}B})}(\alpha \otimes \beta || \sigma) &= \sup_{\Lambda \in \mathsf{LOCC}(\mathrm{A{:}\underline{C}\underline{D}{:}B})} D(\Lambda(\alpha \otimes \beta) || \Lambda(\sigma)) \\ &\geqslant \sup_{\substack{\mathrm{A}^{\mathrm{out}}\mathrm{B}^{\mathrm{out}} \\ \Lambda \in \mathsf{LOCC}(\mathrm{A{:}CD{:}B}\rangle \mathrm{A}^{\mathrm{out}}{:}\mathbb{C}{:}\mathrm{B}^{\mathrm{out}})}} D\big(\Lambda(\alpha \otimes \beta) || \mathcal{S}\big(\mathrm{A}^{\mathrm{out}}{:}\mathrm{B}^{\mathrm{out}}\big)\big) \\ &\geqslant \sup_{\substack{\mathrm{A}^{\mathrm{out}}\mathrm{B}^{\mathrm{out}} \\ \Lambda \in \mathsf{LOCC}(\mathrm{A{:}CD{:}B}\rangle \mathrm{A}^{\mathrm{out}}{:}\mathbb{C}{:}\mathrm{B}^{\mathrm{out}})}} K_D(\Lambda(\alpha \otimes \beta)) \\ &= \sup_{\mathsf{LOCC}(\mathrm{A{:}\underline{C}\underline{D}{:}B})} K_D(\Lambda(\alpha \otimes \beta)) \\ &= K_B(\alpha, \beta)), \end{aligned}$$

where the last equality is due to Lemma 50 and the last inequality is the relative entropy of entanglement upper bound on the distillable key [Hor+05b]. Taking the infimum over σ ends the proof. □

As for the original repeater bound, the tripartite LOCC partial measurements can be relaxed to bipartite LOCC by joining any combinations of the parties. We can still combine this bound with Corollary 23 to get a bound in terms of the distillable entanglement. Compared to Corollary 47, which also gives an upper bound in terms of the distillable entanglement, this bound is valid also for two-way LOCC. However, we trade in a multiplicative factor in the dimension of the key.

Corollary 52. *Let α on $\mathrm{A}_{\text{🔑}}\mathrm{A}_{\text{🛡}}\mathrm{C}_{\text{🔑}}\mathrm{C}_{\text{🛡}}$ and β on $\mathrm{D}_{\text{🔑}}\mathrm{D}_{\text{🛡}}\mathrm{B}_{\text{🔑}}\mathrm{B}_{\text{🛡}}$ be any pair of key-correlated states with at least one separable key-attacked state. Then:*

$$K_B(\alpha, \beta) \leqslant |\mathrm{A}_{\text{🔑}}\mathrm{B}_{\text{🔑}}| \cdot E_D(\alpha \otimes \beta)\ .$$

Proof. $\alpha \otimes \beta$ is still a key-correlated state across the cut between Alice/Bob and Charlie, therefore let us recall the statement of Corollary 23 for $\alpha \otimes \beta$ for two-way LOCC partial measurements.

$$\begin{aligned} E_D(\rho) &\geqslant \chi_{\mathsf{LOCC}(\underline{\mathrm{C}}_{\text{🔑}}\underline{\mathrm{C}}_{\text{🛡}}\underline{\mathrm{D}}_{\text{🔑}}\underline{\mathrm{D}}_{\text{🛡}}{:}\mathrm{A}_{\text{🔑}}\mathrm{A}_{\text{🛡}}\mathrm{B}_{\text{🔑}}\mathrm{B}_{\text{🛡}})}\left(\left\{\tfrac{1}{|\mathrm{A}_{\text{🔑}}\mathrm{B}_{\text{🔑}}|}, \mathcal{Z}^i_{\mathrm{A}_{\text{🔑}}\mathrm{B}_{\text{🔑}}}(\rho)\right\}\right) \\ &= \sup_{\Lambda \in \mathsf{LOCC}(\underline{\mathrm{C}}_{\text{🔑}}\underline{\mathrm{C}}_{\text{🛡}}\underline{\mathrm{D}}_{\text{🔑}}\underline{\mathrm{D}}_{\text{🛡}}{:}\mathrm{A}_{\text{🔑}}\mathrm{A}_{\text{🛡}}\mathrm{B}_{\text{🔑}}\mathrm{B}_{\text{🛡}})} \frac{1}{|\mathrm{A}_{\text{🔑}}\mathrm{B}_{\text{🔑}}|} \sum_i D(\Lambda \circ \mathcal{Z}^i_{\mathrm{B}_{\text{🔑}}}(\rho) \parallel \Lambda(\hat{\rho})) \end{aligned}$$

As a direct application to Theorem 51, we get

$$\begin{aligned}
K_B(\alpha,\beta) &\leqslant D_{\mathsf{LOCC}(\mathrm{A}_{\text{🔑}}\mathrm{A}_{\text{🛡}}:\underline{\mathrm{C}}_{\text{🔑}}\underline{\mathrm{C}}_{\text{🛡}}\underline{\mathrm{D}}_{\text{🔑}}\underline{\mathrm{D}}_{\text{🛡}}:\mathrm{B}_{\text{🔑}}\mathrm{B}_{\text{🛡}})}(\alpha\otimes\beta \,\|\, \hat{\alpha}\otimes\hat{\beta}) \\
&\leqslant D_{\mathsf{LOCC}(\underline{\mathrm{C}}_{\text{🔑}}\underline{\mathrm{C}}_{\text{🛡}}\underline{\mathrm{D}}_{\text{🔑}}\underline{\mathrm{D}}_{\text{🛡}}:\mathrm{A}_{\text{🔑}}\mathrm{A}_{\text{🛡}}\mathrm{B}_{\text{🔑}}\mathrm{B}_{\text{🛡}})}(\alpha\otimes\beta \,\|\, \hat{\alpha}\otimes\hat{\beta}) \\
&= \sup_{\Lambda\in\mathsf{LOCC}(\underline{\mathrm{C}}_{\text{🔑}}\underline{\mathrm{C}}_{\text{🛡}}\underline{\mathrm{D}}_{\text{🔑}}\underline{\mathrm{D}}_{\text{🛡}}:\mathrm{A}_{\text{🔑}}\mathrm{A}_{\text{🛡}}\mathrm{B}_{\text{🔑}}\mathrm{B}_{\text{🛡}})} D(\Lambda(\alpha\otimes\beta) \,\|\, \Lambda(\hat{\alpha}\otimes\hat{\beta})) \\
&\leqslant \sup_{\Lambda\in\mathsf{LOCC}(\underline{\mathrm{C}}_{\text{🔑}}\underline{\mathrm{C}}_{\text{🛡}}\underline{\mathrm{D}}_{\text{🔑}}\underline{\mathrm{D}}_{\text{🛡}}:\mathrm{A}_{\text{🔑}}\mathrm{A}_{\text{🛡}}\mathrm{B}_{\text{🔑}}\mathrm{B}_{\text{🛡}})} \sum_i D(\Lambda\circ(\mathcal{Z}^i_{\mathrm{A}_{\text{🔑}}\mathrm{B}_{\text{🔑}}})(\alpha\otimes\beta) \,\|\, \Lambda(\hat{\alpha}\otimes\hat{\beta})) \\
&\leqslant |\mathrm{A}_{\text{🔑}}\mathrm{B}_{\text{🔑}}|\cdot E_D(\alpha\otimes\beta)
\end{aligned}$$

□

For private bits the resulting factor is just another factor of two off from the dimensionless bound on the one-way repeater of Corollary 47. In particular, all the examples from the previous section will also have asymptotically zero two-way bounded-repeater key rate.

4.4.2 Random private states

The content of this section comes from [2].

We now user the upper bound of Theorem 51 on $K_B(\alpha,\beta)$, to show that two random private states have with high probability a small bounded-repeater key rate. Remember that random private states are states that stay distinguishable under PPT operations, and thus remain distillable under PPT operations. For these states the upper bounds we have proven in terms of the distillable entanglement will not work, because the bounds we know are all on the PPT distillable entanglement. We still expect the LOCC distillable entanglement of such states to be small, but we have no way to prove it. To prove that the real LOCC distillable entanglement of random private states is small we would need to prove a regularized version of the indistinguishability. However, because two copies of a uniformly random state are not uniformly random in the larger system, this task is hard. Here we can circumvent this problem because bounding the operations of the repeater station to single copies implies a single-copy upper bound.

Theorem 53. *Let $|\mathrm{A}_{\text{🛡}}| = |\mathrm{C}_{\text{🛡}}| = d$ and $|\mathrm{D}| = |\mathrm{B}| = k$, and let γ be a random private bits on $\mathrm{A}_{\text{🔑}}\mathrm{A}_{\text{🛡}}\mathrm{C}_{\text{🔑}}\mathrm{C}_{\text{🛡}}$ as defined by Construction 33. Then under the assumption of Conjecture 9, for any state β on DB it holds*

$$\mathbf{P}\left(K_B(\gamma,\beta) \leqslant O\left(\frac{k}{\sqrt{d}}\log dk\right)\right) \geqslant 1 - e^{-c_0 d^3},$$

where $c_0 > 0$ is a universal constant.

Proof. For convenience, let us first bound

$$\|(\gamma-\hat{\gamma})\otimes\beta\|_{\mathsf{LOCC}(\mathrm{A}_{\text{🔑}}\mathrm{A}_{\text{🛡}}:\underline{\mathrm{C}}_{\text{🔑}}\underline{\mathrm{C}}_{\text{🛡}}\underline{\mathrm{D}}:\mathrm{B})}.$$

First, we relax to a set of bipartite partial measurements, and then we use Equation (2.6) to restrict from partial measurements to measurements:

$$\begin{aligned}
\|(\gamma-\hat{\gamma})\otimes\beta\|_{\mathsf{LOCC}(\mathrm{A}_{\text{🔑}}\mathrm{A}_{\text{🛡}}:\underline{\mathrm{C}}_{\text{🔑}}\underline{\mathrm{C}}_{\text{🛡}}\underline{\mathrm{D}}:\mathrm{B})} &\leqslant \|(\gamma-\hat{\gamma})\otimes\beta\|_{\mathsf{LOCC}(\mathrm{A}_{\text{🔑}}\mathrm{A}_{\text{🛡}}\mathrm{B}:\underline{\mathrm{C}}_{\text{🔑}}\underline{\mathrm{C}}_{\text{🛡}}\underline{\mathrm{D}})} \\
&= \|(\gamma-\hat{\gamma})\otimes\beta\|_{\mathsf{LOCC}(\underline{\mathrm{A}}_{\text{🔑}}\underline{\mathrm{A}}_{\text{🛡}}\mathrm{B}:\underline{\mathrm{C}}_{\text{🔑}}\underline{\mathrm{C}}_{\text{🛡}}\underline{\mathrm{D}})} \\
&\leqslant \|(\gamma-\hat{\gamma})\otimes\beta\|_{\mathsf{SEP}(\underline{\mathrm{A}}_{\text{🔑}}\underline{\mathrm{A}}_{\text{🛡}}\underline{\mathrm{B}}:\underline{\mathrm{C}}_{\text{🔑}}\underline{\mathrm{C}}_{\text{🛡}}\underline{\mathrm{D}})}.
\end{aligned} \tag{4.19}$$

Notice that we cannot use Equation (2.6) to restrict directly from $\mathsf{LOCC}(\mathrm{A}_{\text{🔑}}\mathrm{A}_{\text{🛡}}:\underline{\mathrm{C}}_{\text{🔑}}\underline{\mathrm{C}}_{\text{🛡}}\underline{\mathrm{D}}:\mathrm{B})$ to $\mathsf{LOCC}(\underline{\mathrm{A}}_{\text{🔑}}\underline{\mathrm{A}}_{\text{🛡}}:\underline{\mathrm{C}}_{\text{🔑}}\underline{\mathrm{C}}_{\text{🛡}}\underline{\mathrm{D}}:\underline{\mathrm{B}})$ because systems $\mathrm{A}_{\text{🔑}}\mathrm{A}_{\text{🛡}}$ and B are separated. Now let $\Delta = \sigma^+ - \sigma^-$, where $\sigma^{\pm}$ are the random orthogonal states on $\mathrm{A}_{\text{🛡}}\mathrm{C}_{\text{🛡}}$ appearing in the shield of γ. Observe that by Proposition 5

$$
\begin{aligned}
&\|(\gamma-\hat{\gamma})\otimes\beta\|_{\mathsf{SEP}(\underline{\mathrm{A}}_{\text{🔑}}\underline{\mathrm{A}}_{\text{🛡}}\underline{\mathrm{B}}:\underline{\mathrm{C}}_{\text{🔑}}\underline{\mathrm{C}}_{\text{🛡}}\underline{\mathrm{D}})}\\
&\qquad\qquad = \frac{1}{4}\|(\Phi^{+}-\Phi^{-})\otimes\Delta\otimes\beta\|_{\mathsf{SEP}(\underline{\mathrm{A}}_{\text{🔑}}\underline{\mathrm{A}}_{\text{🛡}}\underline{\mathrm{B}}:\underline{\mathrm{C}}_{\text{🔑}}\underline{\mathrm{C}}_{\text{🛡}}\underline{\mathrm{D}})}\\
&\qquad\qquad \leqslant \frac{1}{2}\|\Phi^{+}\otimes\Delta\otimes\beta\|_{\mathsf{SEP}(\underline{\mathrm{A}}_{\text{🔑}}\underline{\mathrm{A}}_{\text{🛡}}\underline{\mathrm{B}}:\underline{\mathrm{C}}_{\text{🔑}}\underline{\mathrm{C}}_{\text{🛡}}\underline{\mathrm{D}})}\\
&\qquad\qquad \leqslant \frac{1}{2}(2\mathcal{R}(\Phi^{+}\otimes\beta)+1)\|\Delta\|_{\mathsf{SEP}(\underline{\mathrm{A}}_{\text{🛡}}:\underline{\mathrm{C}}_{\text{🛡}})}\\
&\qquad\qquad \leqslant \frac{1}{2}(4k-1)\|\Delta\|_{\mathsf{SEP}(\underline{\mathrm{A}}_{\text{🛡}}:\underline{\mathrm{C}}_{\text{🛡}})}\\
&\qquad\qquad \leqslant 2k\|\Delta\|_{\mathsf{SEP}(\underline{\mathrm{A}}_{\text{🛡}}:\underline{\mathrm{C}}_{\text{🛡}})}.
\end{aligned}
$$

Yet, we know from [AL14, Section 6.1], that $\|\Delta\|_{\mathsf{SEP}} \leqslant C'/\sqrt{d}$ with probability greater than $1-e^{-c_0 d^3}$. Therefore, joining with Equation (4.19), we get

$$
\mathbf{P}\left(\|(\gamma-\hat{\gamma})\otimes\beta\|_{\mathsf{LOCC}} \leqslant C\frac{k}{\sqrt{d}}\right) \geqslant 1-e^{-c_0 d^3} \tag{4.20}
$$

with $\mathsf{LOCC} \equiv \mathsf{LOCC}(\mathrm{A}_{\text{🔑}}\mathrm{A}_{\text{🛡}}:\underline{\mathrm{C}}_{\text{🔑}}\underline{\mathrm{C}}_{\text{🛡}}\mathrm{D}:\mathrm{B})$.

Under the assumption of Conjecture 9, with $\kappa = \log 4d^2k^2 = 2\log 2dk$, we use asymptotic continuity on the fact that $\hat{\gamma}\otimes\beta \in \mathcal{S}(\mathrm{A}_{\text{🔑}}\mathrm{A}_{\text{🛡}}:\mathrm{C}_{\text{🔑}}\mathrm{C}_{\text{🛡}}\mathrm{DB})$ to find

$$
\begin{aligned}
&D_{\mathsf{LOCC}(\mathrm{A}_{\text{🔑}}\mathrm{A}_{\text{🛡}}:\underline{\mathrm{C}}_{\text{🔑}}\underline{\mathrm{C}}_{\text{🛡}}\underline{\mathrm{D}}:\mathrm{B})}(\gamma\otimes\beta\|\mathcal{S}(\mathrm{A}_{\text{🔑}}\mathrm{A}_{\text{🛡}}:\mathrm{C}_{\text{🔑}}\mathrm{C}_{\text{🛡}}\mathrm{DB}))\\
&\quad = \Big|D_{\mathsf{LOCC}(\mathrm{A}_{\text{🔑}}\mathrm{A}_{\text{🛡}}:\underline{\mathrm{C}}_{\text{🔑}}\underline{\mathrm{C}}_{\text{🛡}}\underline{\mathrm{D}}:\mathrm{B})}(\gamma\otimes\beta\|\mathcal{S}(\mathrm{A}_{\text{🔑}}\mathrm{A}_{\text{🛡}}:\mathrm{C}_{\text{🔑}}\mathrm{C}_{\text{🛡}}\mathrm{DB}))\\
&\qquad - D_{\mathsf{LOCC}(\mathrm{A}_{\text{🔑}}\mathrm{A}_{\text{🛡}}:\underline{\mathrm{C}}_{\text{🔑}}\underline{\mathrm{C}}_{\text{🛡}}\underline{\mathrm{D}}:\mathrm{B})}(\hat{\gamma}\otimes\beta\|\mathcal{S}(\mathrm{A}_{\text{🔑}}\mathrm{A}_{\text{🛡}}:\mathrm{C}_{\text{🔑}}\mathrm{C}_{\text{🛡}}\mathrm{DB}))\Big|\\
&\quad \leqslant O\Big(\kappa\|(\gamma-\hat{\gamma})\otimes\beta\|_{\mathsf{LOCC}(\mathrm{A}_{\text{🔑}}\mathrm{A}_{\text{🛡}}:\underline{\mathrm{C}}_{\text{🔑}}\underline{\mathrm{C}}_{\text{🛡}}\underline{\mathrm{D}}:\mathrm{B})}\Big).
\end{aligned}
$$

We join this with Equation (4.20) giving

$$
\mathbf{P}\left(D_{\mathsf{LOCC}(\mathrm{A}_{\text{🔑}}\mathrm{A}_{\text{🛡}}:\underline{\mathrm{C}}_{\text{🔑}}\underline{\mathrm{C}}_{\text{🛡}}\underline{\mathrm{D}}:\mathrm{B})}(\gamma\otimes\beta\|\mathcal{S}) \leqslant O\left(\frac{k}{\sqrt{d}}\log dk\right)\right) \geqslant 1-e^{-c_0 d^3} \tag{4.21}
$$

where again $\mathsf{LOCC} \equiv \mathsf{LOCC}(\mathrm{A}_{\text{🔑}}\mathrm{A}_{\text{🛡}}:\underline{\mathrm{C}}_{\text{🔑}}\underline{\mathrm{C}}_{\text{🛡}}\underline{\mathrm{D}}:\mathrm{B})$ and $\mathcal{S} \equiv \mathcal{S}(\mathrm{A}_{\text{🔑}}\mathrm{A}_{\text{🛡}}:\mathrm{C}_{\text{🔑}}\mathrm{C}_{\text{🛡}}\mathrm{DB})$.

To conclude , notice that any state in $\mathcal{S}(\mathrm{A}_{\text{🔑}}\mathrm{A}_{\text{🛡}}:\mathrm{C}_{\text{🔑}}\mathrm{C}_{\text{🛡}}\mathrm{DB})$ is mapped into $\mathcal{S}(\mathrm{A}:\mathrm{B})$ by any map in $\mathsf{LOCC}(\mathrm{A}_{\text{🔑}}\mathrm{A}_{\text{🛡}}:\underline{\mathrm{C}}_{\text{🔑}}\underline{\mathrm{C}}_{\text{🛡}}\underline{\mathrm{D}}:\mathrm{B})$, therefore by Theorem 51

$$
K_B(\gamma,\beta) \leqslant D_{\mathsf{LOCC}(\mathrm{A}_{\text{🔑}}\mathrm{A}_{\text{🛡}}:\underline{\mathrm{C}}_{\text{🔑}}\underline{\mathrm{C}}_{\text{🛡}}\underline{\mathrm{D}}:\mathrm{B})}(\gamma\otimes\beta\|\mathcal{S}(\mathrm{A}_{\text{🔑}}\mathrm{A}_{\text{🛡}}:\mathrm{C}_{\text{🔑}}\mathrm{C}_{\text{🛡}}\mathrm{DB}))
$$

which together with Equation (4.21) proves the theorem. □

Theorem 53 tells us that the bounded-repeater rate of our random private states $K_B(\gamma,\beta)$ should be with high probability small, as its shield dimension d grows in size, because eventually even a maximally entangled state at β eventually cannot help. Since most states are not as robust as the maximally entangled states, we obtain an improved scaling if β is another random private state.

Theorem 54. *Let $|\mathrm{A}_{\text{🛡}}| = |\mathrm{C}_{\text{🛡}}| = d$ and $|\mathrm{D}_{\text{🛡}}| = |\mathrm{B}_{\text{🛡}}| = k$, and let γ on $\mathrm{A}_{\text{🔑}}\mathrm{A}_{\text{🛡}}\mathrm{C}_{\text{🔑}}\mathrm{C}_{\text{🛡}}$ and γ' on $\mathrm{D}_{\text{🔑}}\mathrm{D}_{\text{🛡}}\mathrm{B}_{\text{🔑}}\mathrm{B}_{\text{🛡}}$ be random private bits as defined by Construction 33. Then under the assumption of Conjecture 9 we have*

$$
\mathbf{P}\left(K_B(\gamma,\gamma') \leqslant O\left(\sqrt{\frac{k}{d}}\log kd\right)\right) \geqslant 1-e^{-c_0 d^3}-e^{-c_0' k^3\log^2 k}. \tag{4.22}
$$

where $c_0, c_0' > 0$ are universal constants.

Proof. Just like in Equation (4.19) we have:

$$\begin{aligned}&\|(\gamma-\hat{\gamma})\otimes\gamma'\|_{\mathsf{LOCC}(\mathrm{A}_{\text{🔑}}\mathrm{A}_{\text{🛡}}:\underline{\mathrm{C}}_{\text{🔑}}\underline{\mathrm{C}}_{\text{🛡}}\underline{\mathrm{D}}_{\text{🔑}}\underline{\mathrm{D}}_{\text{🛡}}:\mathrm{B}_{\text{🔑}}\mathrm{B}_{\text{🛡}})}\\ &\qquad\leqslant\|(\gamma-\hat{\gamma})\otimes\gamma'\|_{\mathsf{SEP}(\underline{\mathrm{A}}_{\text{🔑}}\underline{\mathrm{A}}_{\text{🛡}}\underline{\mathrm{B}}_{\text{🔑}}\underline{\mathrm{B}}_{\text{🛡}}:\underline{\mathrm{C}}_{\text{🔑}}\underline{\mathrm{C}}_{\text{🛡}}\underline{\mathrm{D}}_{\text{🔑}}\underline{\mathrm{D}}_{\text{🛡}})}.\end{aligned}\tag{4.23}$$

Now let $\sigma^{\pm}$ on $\mathrm{A}_{\text{🛡}}\mathrm{C}_{\text{🛡}}$ be the random orthogonal shield states of γ and define $\Delta=\sigma^{+}-\sigma^{-}$. Similarly let $\varsigma^{\pm}$ on $\mathrm{D}_{\text{🛡}}\mathrm{B}_{\text{🛡}}$ be the random orthogonal shield states of $\tilde{\gamma}$. Then by Proposition 5

$$\begin{aligned}&\|(\gamma-\hat{\gamma})\otimes\tilde{\gamma}\|_{\mathsf{SEP}(\underline{\mathrm{A}}_{\text{🔑}}\underline{\mathrm{A}}_{\text{🛡}}\underline{\mathrm{B}}_{\text{🔑}}\underline{\mathrm{B}}_{\text{🛡}}:\underline{\mathrm{C}}_{\text{🔑}}\underline{\mathrm{C}}_{\text{🛡}}\underline{\mathrm{D}}_{\text{🔑}}\underline{\mathrm{D}}_{\text{🛡}})}\\ &\qquad\leqslant\frac{1}{2}\|\Phi^{+}\otimes\Delta\otimes\tilde{\gamma}\|_{\mathsf{SEP}(\underline{\mathrm{A}}_{\text{🔑}}\underline{\mathrm{A}}_{\text{🛡}}\underline{\mathrm{B}}_{\text{🔑}}\underline{\mathrm{B}}_{\text{🛡}}:\underline{\mathrm{C}}_{\text{🔑}}\underline{\mathrm{C}}_{\text{🛡}}\underline{\mathrm{D}}_{\text{🔑}}\underline{\mathrm{D}}_{\text{🛡}})}\\ &\qquad\leqslant\frac{1}{4}\|\Phi^{+}\otimes\Delta\otimes\Phi^{+}\otimes\varsigma^{+}\|_{\mathsf{SEP}(\underline{\mathrm{A}}_{\text{🔑}}\underline{\mathrm{A}}_{\text{🛡}}\underline{\mathrm{B}}_{\text{🔑}}\underline{\mathrm{B}}_{\text{🛡}}:\underline{\mathrm{C}}_{\text{🔑}}\underline{\mathrm{C}}_{\text{🛡}}\underline{\mathrm{D}}_{\text{🔑}}\underline{\mathrm{D}}_{\text{🛡}})}\\ &\qquad\quad+\frac{1}{4}\|\Phi^{+}\otimes\Delta\otimes\Phi^{-}\otimes\varsigma^{-}\|_{\mathsf{SEP}(\underline{\mathrm{A}}_{\text{🔑}}\underline{\mathrm{A}}_{\text{🛡}}\underline{\mathrm{B}}_{\text{🔑}}\underline{\mathrm{B}}_{\text{🛡}}:\underline{\mathrm{C}}_{\text{🔑}}\underline{\mathrm{C}}_{\text{🛡}}\underline{\mathrm{D}}_{\text{🔑}}\underline{\mathrm{D}}_{\text{🛡}})}\\ &\qquad\leqslant\frac{1}{4}(2\mathcal{R}(\Phi^{+\otimes 2})+1)(2\mathcal{R}_{\mathrm{D}_{\text{🛡}}:\mathrm{B}_{\text{🛡}}}(\varsigma^{+})+1)\|\Delta\|_{\mathsf{SEP}(\underline{\mathrm{A}}_{\text{🛡}}:\underline{\mathrm{D}}_{\text{🛡}})}\\ &\qquad\quad+\frac{1}{4}(2\mathcal{R}(\Phi^{+\otimes 2})+1)(2\mathcal{R}_{\mathrm{D}_{\text{🛡}}:\mathrm{B}_{\text{🛡}}}(\varsigma^{-})+1)\|\Delta\|_{\mathsf{SEP}(\underline{\mathrm{A}}_{\text{🛡}}:\underline{\mathrm{D}}_{\text{🛡}})}\\ &\qquad=\frac{7}{2}(\mathcal{R}_{\mathrm{D}_{\text{🛡}}:\mathrm{B}_{\text{🛡}}}(\varsigma^{+})+\mathcal{R}_{\mathrm{D}_{\text{🛡}}:\mathrm{B}_{\text{🛡}}}(\varsigma^{-})+1)\|\Delta\|_{\mathsf{SEP}(\underline{\mathrm{A}}_{\text{🛡}}:\underline{\mathrm{D}}_{\text{🛡}})}\end{aligned}$$

Yet, we know from [AL14, Section 6.1] that $\|\Delta\|_{\mathsf{SEP}}\leqslant C'/\sqrt{d}$ with probability greater than $1-e^{-c_0 d^3}$, and from Corollary 32 that $\mathcal{R}(\varsigma^{\pm})\leqslant C''\sqrt{k}\log k$ with probability greater than $1-e^{-c_0 k^3\log^2 k}$. Therefore, joining the above with Equation (4.23), we get (albeit with different constants)

$$\mathbf{P}\left(\|(\gamma-\hat{\gamma})\otimes\tilde{\gamma}\|_{\mathsf{LOCC}(\mathrm{A}_{\text{🔑}}\mathrm{A}_{\text{🛡}}:\underline{\mathrm{C}}_{\text{🔑}}\underline{\mathrm{C}}_{\text{🛡}}\underline{\mathrm{D}}_{\text{🔑}}\underline{\mathrm{D}}_{\text{🛡}}:\mathrm{B}_{\text{🔑}}\mathrm{B}_{\text{🛡}})}\leqslant C\sqrt{\frac{k}{d}}\log kd\right)\geqslant 1-e^{-c_0 d^3}-e^{-c_0' k^3\log^2 k}.\tag{4.24}$$

Again under the assumption of Conjecture 9 with $\kappa=2\log 4d^2k^2$, we use the fact that $\hat{\gamma}\otimes\gamma'\in\mathcal{S}\equiv\mathcal{S}(\mathrm{A}_{\text{🔑}}\mathrm{A}_{\text{🛡}}:\mathrm{C}_{\text{🔑}}\mathrm{C}_{\text{🛡}}\mathrm{D}_{\text{🔑}}\mathrm{D}_{\text{🛡}}\mathrm{B}_{\text{🔑}}\mathrm{B}_{\text{🛡}})$, and find

$$D_{\mathsf{LOCC}}(\gamma\otimes\gamma'\|\mathcal{S})\leqslant O(\kappa\|(\gamma-\hat{\gamma})\otimes\gamma'\|_{\mathsf{LOCC}}).$$

for $\mathsf{LOCC}\equiv\mathsf{LOCC}(\mathrm{A}_{\text{🔑}}\mathrm{A}_{\text{🛡}}:\underline{\mathrm{C}}_{\text{🔑}}\underline{\mathrm{C}}_{\text{🛡}}\underline{\mathrm{D}}_{\text{🔑}}\underline{\mathrm{D}}_{\text{🛡}}:\mathrm{B}_{\text{🔑}}\mathrm{B}_{\text{🛡}})$ and $\mathcal{S}$ as above. We then join this with Equation (4.24)

$$\mathbf{P}\left(D_{\mathsf{LOCC}}(\gamma\otimes\gamma'\|\mathcal{S})\leqslant C\sqrt{\frac{k}{d}}\log k\log d\right)\geqslant 1-e^{-c_0 d^3}-e^{-c_0' k^3\log^2 k}\tag{4.25}$$

with again LOCC and $\mathcal{S}$ as above. Since any state in $\mathcal{S}$ is mapped into $\mathcal{S}(\mathrm{A}_{\text{🔑}}\mathrm{A}_{\text{🛡}}:\mathrm{B}_{\text{🔑}}\mathrm{B}_{\text{🛡}})$ by any map in LOCC, we conclude by Theorem 51 that

$$K_B(\gamma,\gamma')\leqslant D_{\mathsf{LOCC}}(\gamma\otimes\varrho\|\mathcal{S})$$

where $\mathsf{LOCC}\equiv\mathsf{LOCC}(\mathrm{A}_{\text{🔑}}\mathrm{A}_{\text{🛡}}:\underline{\mathrm{C}}_{\text{🔑}}\underline{\mathrm{C}}_{\text{🛡}}\underline{\mathrm{D}}_{\text{🔑}}\underline{\mathrm{D}}_{\text{🛡}}:\mathrm{B}_{\text{🔑}}\mathrm{B}_{\text{🛡}})$ and $\mathcal{S}\equiv\mathcal{S}(\mathrm{A}_{\text{🔑}}\mathrm{A}_{\text{🛡}}:\mathrm{C}_{\text{🔑}}\mathrm{C}_{\text{🛡}}\mathrm{D}_{\text{🔑}}\mathrm{D}_{\text{🛡}}\mathrm{B}_{\text{🔑}}\mathrm{B}_{\text{🛡}})$ as above, which together with Equation (4.25) ends the proof. □

We have shown that for random private states the bounded repeater rate will be low, under the assumption of asymptotic continuity of general restricted relative entropies.

Finally notice that the bounded repeater distillable key gives an upper bound on the bounded repeater distillable entanglement. Namely, we have for any channel Λ

$$E_D(\Lambda(\alpha \otimes \beta)) \leqslant K_D(\Lambda(\alpha \otimes \beta)) \leqslant K_B(\alpha, \beta)$$

showing that random private states have indeed some form of low distillable entanglement which will be forcibly close to the small bounded repeater distillable key for high dimension.

4.5 Summary

We have given general results showing that in Equation (4.4) the resulting behaviour is

$$E_D(\rho) = E_{DR}(\rho, \rho) \sim K_R(\rho, \rho) \ll K_D(\rho) \tag{4.26}$$

While the general case is still out of our reach, it seems more and more likely, and there might even be equality in the two repeater distillation rates.

Motivated by the question whether the repeater key rate can be zero for states with non-zero key rate, the PPT^2 conjecture was introduced in [Chr12]. In our context, it states that any repeater protocol acting on states α and β that are both PPT will be separable at Alice and Bob. If true, it would imply that the repeater key rate of two PPT states (which are the only known examples for bound entangled states) is zero, including if the two states are bound entangled states with key. That is, if α and β are PPT states, then $K_R(\alpha, \beta) = 0$ even if $K_D(\alpha), K_D(\beta) > 0$. Just like in the case of Equation (4.26), there have only been progress in support of the conjecture, but the general case is still open, see also [CMW18] and references there in.

Our results for random private states give a complementary view on the PPT^2 conjecture, by providing states far from PPT that nonetheless behave like separable states across a repeater station. Our work might thus be viewed as pointing to extensions of the PPT^2 conjecture. We would also like to mention an implication for the older NPT bound entanglement conjecture, which postulates that there exist undistillable NPT states [Dür+00; DiV+00]. Since the states that we have constructed have low LOCC-restricted relative entropy of entanglement, and since the regularised LOCC-restricted relative entropy of entanglement is an upper bound on the distillable entanglement, our states are candidates for NPT states with zero distillable entanglement.

Epilogue

Corollary 47 bounds the repeater key rate of a restricted class of states and protocols. While being restrictive, we would like to stress that the communication between Alice and Bob is two way, and also that if the two-way step is limited to bipartite distillation between the nodes, we can always apply the result to the outcomes of the distillation. In particular even if the two-way recurrence protocol is used to distill between the nodes, as in the case of heralded entanglement generation and purification, we can apply the bound on swapping key rate to the outputs of the recurrence protocol. The bound also applies to repeater key schemes based on quantum error correction. The link with outgoing communication from the station is trivially covered. For the link with incoming communication, the bound on swapping key rate applies to the output of the code (as mentioned above), since usually the code is decoded or corrected at the station rendering it a bipartite distillation protocol. As such, we can apply our bound in some way to most repeater schemes (see also [Mur+16] and references therein for an overview of the repeater schemes) and where it applies, any attempt to improve the rate of key distillation above that of entanglement distillation will not work. For example, attempting to use the noisy processing protocol[RGK05] would yield no advantage. We are not aware that there exist any protocol that contains a truly two-way tripartite step. Finally, we note that, because optimal one-way protocols exist when close to the target states, the optimal two-way protocols are composed of a two-way "lift-off" protocol followed by a one-way "conclusion" protocol [KW04].

We leave as an open problem whether Corollary 47 generalizes to all states and protocols, including two-way communication. Such a result would show that all entangled states with zero distillable entanglement, including those with distillable key, have zero repeater key rate. Another open problem, called the PPT^2 conjecture [Chr12], asks whether swapping PPT states in all dimensions always yields separable states. If the conjecture is true, then it would imply that all PPT states have zero repeater key rate. In that, the results here presented support the conjecture. Since our results are asymptotic in nature, they give insight on the PPT^2 conjecture that cannot be achieved with the study of swapping specific states in specific dimensions.

The connection made between key distillation, entanglement distillation and quantum data hiding raises the possibility of finding a rate at which data hiding states can be distilled, H_D (which we refrain from defining formally). Namely, in performing entanglement distillation on private states, it may be possible to retain the undistillable correlations into data hiding states with zero distillable entanglement so that they could be used as a resource, such that

$$K_D(\rho) = H_D(\rho) + E_D(\rho).$$

We have also shown that to separated parties random private states are indistinguishable from separable states. Ultimately we would like to show is that this remains true asymptotically, namely under regularization, as this would imply actually upper bounding both the distillable entanglement and the repeater distillable key with such indistinguishability. To prove such a result new tools are needed that can deal with the non-uniform randomness of tensor product or random states. Let us emphasize once more that the “usual” bounds on quantities such as the quantum repeater key rate or the distillable entanglement, are based on the partial transposition and would thus be useless for random private states. Finally, we acknowledge that the states between adjacent nodes in a network will generally be specifically designed states rather than random states. However private states will generally also not be the designed input states. Private states are to be thought as the result of performing key distillation, with the shield including, for instance, all the classical communication and the randomness used in the classical part of the protocols. The output of quantum key distribution protocols thus is indeed a random variable over private states. This argument has its limitations though. If we consider states for which the distillable key and the distillable entanglement are the same, the output will still be a random variable, but it will not be a random private state because, well, it is distillable. It would be interesting to determine, what is the distribution of private states generated by common quantum key distribution protocols.

Appendix - Asymptotic continuity conjecture

In Section 2.3 we have conjectured that any restricted relative entropy is asymptotically continuous. Here we display a failed attempt at the proof. This section heavily builds on the proof of [LW14, Proposition 3, page 10].

In the asymptotic continuity proof of $D_{\mathsf{L}}(\cdot\|\mathcal{K})$ for measurements, a simple but crucial step is the smoothing of $\mathcal{K}$. Namely the first step is defining $\mathcal{K}_x := (1-x)\mathcal{K} + x\tau$ and $\sigma_x := (1-x)\sigma + x\tau$ for any state σ, where τ is the maximally mixed state. It is then a few steps to show that

$$D_{\mathsf{L}}(\varrho\|\mathcal{K}) \leqslant D_{\mathsf{L}}(\varrho\|\mathcal{K}_x) \leqslant D_{\mathsf{L}}(\varrho\|\mathcal{K}) - \log(1-x). \tag{27}$$

The main part of the proof then takes the outcome probabilities $\operatorname{tr} M_i\sigma_x$ and rescales them to $\operatorname{tr} M_i\sigma_x / \operatorname{tr} M_i \leqslant 1$, where the factor gets absorbed into another term, the asymptotic continuity of which is not affected. The smoothing of σ then provides a lower bound on the value of $\operatorname{tr} M_i\sigma_x / \operatorname{tr} M_i$ and asymptotic continuity can be proven. This breaks for general channels.

We can assume that our class of channels is such that, if $\Lambda \in \mathsf{L}$ is a channel in this class, then there exist a Kraus decomposition

$$\Lambda(\rho) = \sum_i K_i\rho K_i^\dagger$$

such that the following instrument, that we call Λ with records,

$$\hat{\Lambda}(\rho) = \sum_i |i\rangle\langle i| \otimes K_i\rho K_i^\dagger$$

is also in L. Because the original channel is always recovered by tracing out the record, by monotonicity of the relative entropy, the supremum in $D_{\mathsf{L}}(\cdot\|\mathcal{K})$ can always be achieved by channels with records. Going back to the proof argument, we could not find a way to rescale $K_i\sigma_x K_i^\dagger$ as to provide a lower bound on the minimal eigenvalue, while keeping the maximal eigenvalue lower that 1. We can recover the asymptotic continuity provided by the smoothing step, by also smoothing the maps with smoothed, completable versions of the Kraus operators, extracted from the weak measurements of [KKB11], as shown in the proof below. However using weak Kraus operators does not provide a similar guarantee as the smoothing of the states. Namely the analogue of Equation (27) is missing and the proof does not work. This missing step could be thought as a warning that indeed the jump, that would make $D_{\mathsf{L}}(\cdot\|\mathcal{K})$ not asymptotically continuous, could happen exactly in the completion of the smoothed map.

Proof of the step with weak measurements. Let us consider maps $\hat{\Lambda}$ of the form

$$\hat{\Lambda}(\rho) = \sum_i |i\rangle\langle i| \otimes \hat{K}_i\rho\hat{K}_i^\dagger.$$

Using the polar decomposition of $\hat{K}_i$ we can further write

$$\hat{\Lambda}(\rho) = \sum_i |i\rangle\langle i| \otimes U_i K_i \rho K_i U_i^\dagger.$$

where K_i are positive semidefinite operators on H, and U_i are projections/isometries that can be reversed[2], meaning that if we define the map

$$\Lambda(\rho) = \sum_i |i\rangle\langle i| \otimes K_i \rho K_i$$

then we have

$$D(\hat{\Lambda}(\varrho)\|\hat{\Lambda}(\varsigma)) = D(\Lambda(\varrho)\|\Lambda(\varsigma)).$$

for any states ϱ and ς. We now need to smooth Λ into a map $\tilde{\Lambda}$, so that we can have a lower bound on the eigenvalues produced on σ_x. To do this, we use the pseudo-weak measurement construction of [KKB11]. Let $\varepsilon > 0$, and define

$$\varepsilon_i := \varepsilon \left\|K_i^2\right\|_\infty \qquad K := \sum_i \left\|K_i^2\right\|_\infty$$
$$E := 1 + \textstyle\sum_i \varepsilon_i = 1 + \varepsilon K,$$

and then define the positive-semidefinite operators

$$\tilde{K}_i := (K_i^2 + \varepsilon_i \mathbb{1})^{\frac{1}{2}} = (K_i^2 + \varepsilon \left\|K_i^2\right\|_\infty \mathbb{1})^{\frac{1}{2}}.$$

Notice that $\tilde{K}_i^2/E$ sum to the identity, thus

$$\tilde{\Lambda}(\varrho) = \sum_i |i\rangle\langle i| \otimes \tilde{K}_i \frac{\rho}{E} \tilde{K}_i$$

is a valid channel. The original channel can be recovered with the Kraus operators

$$R_{ki} = |k\rangle\langle i| \otimes (\delta_{ki} + \varepsilon_i)^{\frac{1}{2}} K_k \tilde{K}_i^{-1}$$

defining the channel

$$\mathrm{P}(\varrho) = \sum_{ki} R_{ki} \rho R_{ki}^\dagger.$$

Namely it is straightforward to check that for all states ϱ

$$\mathrm{P} \circ \tilde{\Lambda}(\varrho) = \Lambda(\varrho)$$

and thus

$$D(\tilde{\Lambda}(\varrho)\|\tilde{\Lambda}(\varrho)) \geqslant D(\Lambda(\varrho)\|\Lambda(\varrho)).$$

We can now begin. For any constant $a_i > 0$ we have

$$\begin{aligned} D_{\{\tilde{\Lambda}\}}(\varrho\|\sigma_x) &= \sum_i \mathrm{tr}\left[\tilde{K}_i \frac{\varrho}{E} \tilde{K}_i \left[\log\left(\tilde{K}_i \frac{\varrho}{E} \tilde{K}_i\right) - \log\left(\tilde{K}_i \frac{\sigma_x}{E} \tilde{K}_i\right)\right]\right] \\ &= \sum_i \mathrm{tr}\left[\tilde{K}_i \frac{\varrho}{E} \tilde{K}_i \left[\log(\tilde{K}_i \varrho \tilde{K}_i) - \log(\tilde{K}_i \sigma_x \tilde{K}_i)\right]\right] \end{aligned}$$

[2] We can write any operator as $\hat{K} = \sum_{k=0}^r s_k |l_k\rangle\langle r_k|$ where $r \leqslant d$ is the rank, $s_k > 0$ are the non-zero singular values, and $\{|l_k\rangle\}$ as well as $\{|r_k\rangle\}$ are orthogonal to each other. Then $K = \sum_{k=0}^r s_k |r_k\rangle\langle r_k|$ is positive semidefinite and $U = \sum_{k=0}^r |l_k\rangle\langle r_k|$. It then holds that $\hat{K} = UK$ and $K = U^\dagger \hat{K}$, meaning that the action of U can be reversed even if U itself is not invertible.

$$
\begin{aligned}
&= \sum_i \operatorname{tr}\left[\tilde{K}_i \frac{\varrho}{E} \tilde{K}_i \log\left(a_i \tilde{K}_i \varrho \tilde{K}_i\right)\right] - \sum_i \operatorname{tr}\left[\tilde{K}_i \frac{\varrho}{E} \tilde{K}_i \log\left(a_i \tilde{K}_i \sigma_x \tilde{K}_i\right)\right] \\
&=: \mathrm{II}(\varrho) - \mathrm{I}(\varrho)
\end{aligned}
$$

where we defined the sums $\mathrm{II}(\varrho)$ and $\mathrm{I}(\varrho)$ just like in [LW14].

For the σ-dependent term $\mathrm{I}(\varrho)$ we have

$$
\begin{aligned}
|\mathrm{I}(\varrho) - \mathrm{I}(\varsigma)| &= \left| \sum_i \operatorname{tr}\left[\tilde{K}_i \frac{\varrho - \varsigma}{E} \tilde{K}_i \log\left(a_i \tilde{K}_i \sigma_x \tilde{K}_i\right)\right] \right| \\
&\leqslant \sum_i \left| \operatorname{tr}\left[\tilde{K}_i \frac{\varrho - \varsigma}{E} \tilde{K}_i \log\left(a_i \tilde{K}_i \sigma_x \tilde{K}_i\right)\right] \right| \\
&\leqslant \sum_i \left\| \tilde{K}_i \frac{\varrho - \varsigma}{E} \tilde{K}_i \right\|_1 \left\| \log\left(a_i \tilde{K}_i \sigma_x \tilde{K}_i\right) \right\|_\infty
\end{aligned}
$$

We now have

$$
\tilde{K}_i \sigma_x \tilde{K}_i \leqslant \|\tilde{K}_i\|_\infty^2 \|\varsigma_x\|_\infty \leqslant \|\tilde{K}_i\|_\infty^2 = (1+\varepsilon)\|K_i^2\|_\infty
$$

and

$$
\tilde{K}_i \sigma_x \tilde{K}_i \geqslant x \tilde{K}_i \tau \tilde{K}_i = \frac{x}{d} \tilde{K}_i^2 \geqslant \frac{x}{d} \varepsilon \|K_i^2\|_\infty .
$$

By choosing $c \leqslant 1$ and fixing

$$
a_i := \frac{c}{(1+\varepsilon)\|K_i^2\|_\infty} \tag{28}
$$

the above implies

$$
\mathbb{1} \geqslant c\mathbb{1} \geqslant a_i \tilde{K}_i \sigma_x \tilde{K}_i \geqslant c \frac{x}{d} \frac{\varepsilon}{1+\varepsilon} .
$$

We can now use this to further bound $|\mathrm{I}(\varrho) - \mathrm{I}(\varsigma)|$:

$$
\begin{aligned}
|\mathrm{I}(\varrho) - \mathrm{I}(\varsigma)| &\leqslant \sum_i \|\tilde{K}_i(\varrho - \varsigma)\tilde{K}_i\|_1 \left\|\log\left(a_i \tilde{K}_i \sigma_x \tilde{K}_i\right)\right\|_\infty \\
&\leqslant \sum_i \|\tilde{K}_i(\varrho - \varsigma)\tilde{K}_i\|_1 \left|\log \frac{cx\varepsilon}{d(1+\varepsilon)}\right| \\
&= \|\Lambda(\varrho - \varsigma)\|_1 \log \frac{d(1+\varepsilon)}{cx\varepsilon} .
\end{aligned} \tag{29}
$$

Let η be the operator function $\eta(t) = -t \log t$; then for the second term $\mathrm{II}(\varrho)$ we have:

$$
\begin{aligned}
|\mathrm{II}(\varrho) - \mathrm{II}(\varsigma)| &= \left| \sum_i \frac{1}{a_i} \operatorname{tr}\left[\eta(a_i \tilde{K}_i \varrho \tilde{K}_i) - \eta(a_i \tilde{K}_i \varsigma \tilde{K}_i)\right] \right| \\
&\leqslant \frac{1}{a_i} \sum_i \left| \operatorname{tr}\left[\eta(a_i \tilde{K}_i \varrho \tilde{K}_i) - \eta(a_i \tilde{K}_i \varsigma \tilde{K}_i)\right] \right|
\end{aligned}
$$

We now define:

$$
\tilde{\varrho}_i = a_i \tilde{K}_i \varrho \tilde{K}_i^\dagger \qquad\qquad \tilde{\varsigma}_i = a_i \tilde{K}_i \varsigma \tilde{K}_i^\dagger .
$$

Notice that $|\operatorname{tr}[\eta(\tilde{\varrho}_i) - \eta(\tilde{\varrho}_i)]|$ is an entropy-like difference for which we would like to get a Fannes-like upper bound, but ϱ_i and ς_i are not normalized. However we do have $c\mathbb{1} \geqslant \tilde{\varrho}_i, \tilde{\varsigma}_i \geqslant 0$, and since ϱ and ς have rank at most d, the rank of ϱ_i and ς_i will also be at most d. One can check that the argument used in [NC02] still works for unnormalized operators if we chose $c \leqslant \frac{1}{2}$. Namely, let for a moment ρ and σ be positive-semidefinite rank-d operators satisfying $c\mathbb{1} \geqslant \rho, \sigma \geqslant 0$, and let $\rho^\downarrow = (r_1 \ldots r_d)$ and $\sigma^\downarrow = (s_1 \ldots s_d)$

be the vectors of eigenvalues in descending order for ρ and σ respectively. Notice that they do not need to be operators on a d dimensional space, or have support on the same d-dimensional subspace. Then

$$|\operatorname{tr}\eta(\rho)-\operatorname{tr}\eta(\sigma)|\leqslant\left\|\rho^{\downarrow}-\sigma^{\downarrow}\right\|_1\log d+\eta\left(\left\|\rho^{\downarrow}-\sigma^{\downarrow}\right\|_1\right).$$

Therefore we find

$$|\mathrm{II}(\varrho)-\mathrm{II}(\varsigma)|\leqslant\sum_i\frac{1}{a_i}\left(\left\|\tilde{\varrho}_i^{\downarrow}-\tilde{\varsigma}_i^{\downarrow}\right\|_1\log d+\eta(\left\|\tilde{\varrho}_i^{\downarrow}-\tilde{\varsigma}_i^{\downarrow}\right\|_1)\right).$$

Furthermore we have $\left\|\tilde{\varrho}_i^{\downarrow}-\tilde{\varsigma}_i^{\downarrow}\right\|_1\leqslant\|\tilde{\varrho}_i-\tilde{\varsigma}_i\|_1$ and

$$\begin{aligned}\|\tilde{\varrho}_i-\tilde{\varsigma}_i\|_1&\leqslant\|\tilde{\varrho}_i\|_1+\|\tilde{\varsigma}_i\|_1\\&\leqslant a_i\|\tilde{K}_i^2\|_\infty\|\varrho\|_1+a_i\|\tilde{K}_i^2\|_\infty\|\varsigma_i\|_1\\&\leqslant\frac{2c}{1+\varepsilon}\leqslant 2c\end{aligned}$$

Since $y\log d+\eta(y)$ is monotone for $y\leqslant\frac{d}{e}$ (where we can assume $d\geqslant 2$), if we further reduce c down to $c=\frac{1}{e}$ we can then write

$$\begin{aligned}&|\mathrm{II}(\varrho)-\mathrm{II}(\varsigma)|\\&\leqslant\sum_i\frac{1}{a_i}(\|\tilde{\varrho}_i-\tilde{\varsigma}_i\|_1\log d+\eta(\|\tilde{\varrho}_i-\tilde{\varsigma}_i\|_1))\\&=E\|\tilde{\Lambda}(\varrho-\varsigma)\|_1\log d+\sum_i\frac{1}{a_i}\eta(\|\tilde{\varrho}_i-\tilde{\varsigma}_i\|_1).\end{aligned}$$

We now call $A:=\sum_i a_i^{-1}$ and by the concavity of η we have

$$\begin{aligned}&|\mathrm{II}(\varrho)-\mathrm{II}(\varsigma)|\\&\leqslant E\|\tilde{\Lambda}(\varrho-\varsigma)\|_1\log d+A\sum_i\frac{a_i^{-1}}{A}\eta(\|\tilde{\varrho}_i-\tilde{\varsigma}_i\|_1)\\&\leqslant E\|\tilde{\Lambda}(\varrho-\varsigma)\|_1\log d+A\eta\left(\sum_i\frac{a_i^{-1}}{A}\|\tilde{\varrho}_i-\tilde{\varsigma}_i\|_1\right)\\&=E\|\tilde{\Lambda}(\varrho-\varsigma)\|_1\log d+A\eta\left(\frac{E}{A}\|\tilde{\Lambda}(\varrho-\varsigma)\|_1\right).\\&=E\|\tilde{\Lambda}(\varrho-\varsigma)\|_1\left(\log d+\log\frac{E}{A}+\log\|\tilde{\Lambda}(\varrho-\varsigma)\|_1\right).\end{aligned}$$

□

A refined version of the asymptotic continuity conjecture could look like

Conjecture 55. *Let $|\mathcal{D}(\mathrm{H})|=d$ and let $\mathcal{K}\in\mathcal{D}(\mathrm{H})$ be a set of states star shaped around the maximally mixed state $\tau=\mathbb{1}/d$. Let L be a class of quantum channels on $\mathcal{D}(\mathrm{H})$ such that if $\rho\mapsto\sum_i K_i\rho K_i^\dagger\in\mathsf{L}$ then $\rho\mapsto\sum_i|i\rangle\langle i|\otimes K_i\rho K_i^\dagger\in\mathsf{L}$. Then for any states ϱ and ς on H satisfying $\epsilon:=\|\varrho-\varsigma\|_{\mathsf{L}}\leqslant\frac{1}{2e}$ we have*

$$\left|D_{\mathsf{L}}(\varrho\|\mathcal{K})-D_{\mathsf{L}}(\varsigma\|\mathcal{K})\right|\leqslant\epsilon\log 4e^2d^3+2\eta(\epsilon)-\log(1-\epsilon)$$

where $d=|\mathrm{H}|$.

Bibliography

[AF04] R. **Alicki** and M. **Fannes**. “Continuity of quantum conditional information”. In: *Journal of Physics A: Mathematical and General* 37.5 (2004), p. L55. DOI: 10.1088/0305-4470/37/5/L01. arXiv: quant-ph/0312081.

[AL14] G. **Aubrun** and C. **Lancien**. “Locally restricted measurements on a multipartite quantum system: data hiding is generic”. In: (2014). arXiv: 1406.1959 [quant-ph]. URL: http://www.rintonpress.com/journals/qiconline.html#v15n56.

[AS06] G. **Aubrun** and S. J. **Szarek**. “Tensor products of convex sets and the volume of separable states on *N* qudits”. In: *Physical Review A* 73.2 (2006), p. 022109. DOI: 10.1103/PhysRevA.73.022109. arXiv: quant-ph/0503221.

[AS17] G. **Aubrun** and S. **Szarek**. “Dvoretzky’s Theorem and the Complexity of Entanglement Detection”. In: *Discrete Analysis* 2017.1 (2017). DOI: 10.19086/da.1242. arXiv: 1510.00578 [quant-ph].

[ASY14] G. **Aubrun**, S. J. **Szarek**, and D. **Ye**. “Entanglement thresholds for random induced states”. In: *Communications on pure and applied mathematics* 67.1 (2014), pp. 129–171. DOI: 10.1002/cpa.21460. arXiv: 1106.2264 [quant-ph].

[Aud+02] K. **Audenaert**, B. **De Moor**, K. G. H. **Vollbrecht**, and R. F. **Werner**. “Asymptotic relative entropy of entanglement for orthogonally invariant states”. In: *Physical Review A* 66.3 (2002), p. 032310. DOI: 10.1103/PhysRevA.66.032310. arXiv: quant-ph/0204143.

[Aud07] K. M. **Audenaert**. “A sharp continuity estimate for the von Neumann entropy”. In: *Journal of Physics A: Mathematical and Theoretical* 40.28 (2007), p. 8127. DOI: 10.1088/1751-8113/40/28/S18. arXiv: quant-ph/0610146.

[Bad+03] P. **Badziag**, M. **Horodecki**, A. **Sen**, U. **Sen**, et al. “Locally accessible information: How much can the parties gain by cooperating?” In: *Physical review letters* 91.11 (2003), p. 117901. DOI: 10.1103/PhysRevLett.91.117901. arXiv: quant-ph/0304040.

[Bäu+15] S. **Bäuml**, M. **Christandl**, K. **Horodecki**, and A. **Winter**. “Limitations on quantum key repeaters”. In: *Nature communications* 6 (2015), p. 6908. DOI: 10.1038/ncomms7908. arXiv: 1402.5927 [quant-ph].

[BB84] C. H. **Bennett** and G. **Brassard**. “An update on quantum cryptography”. In: *Workshop on the Theory and Application of Cryptographic Techniques*. Springer. 1984, pp. 475–480. DOI: 10.1007/3-540-39568-7_39.

[Ben+83] C. H. **Bennett**, G. **Brassard**, S. **Breidbart**, and S. **Wiesner**. “Quantum cryptography, or unforgeable subway tokens”. In: *Advances in Cryptology*. Springer. 1983, pp. 267–275. DOI: 10.1007/978-1-4757-0602-4_26.

[Ben+93] 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". In: *Physical review letters* 70.13 (1993), p. 1895. DOI: 10.1103/PhysRevLett.70.1895.

[Ben+96a] C. H. **Bennett**, H. J. **Bernstein**, S. **Popescu**, and B. **Schumacher**. "Concentrating partial entanglement by local operations". In: *Physical Review A* 53.4 (1996), p. 2046. DOI: 10.1103/PhysRevA.53.2046. arXiv: quant-ph/9511030.

[Ben+96b] C. H. **Bennett**, G. **Brassard**, S. **Popescu**, B. **Schumacher**, J. A. **Smolin**, and W. K. **Wootters**. "Purification of noisy entanglement and faithful teleportation via noisy channels". In: *Physical review letters* 76.5 (1996), p. 722. DOI: 10.1103/PhysRevLett.76.722. arXiv: quant-ph/9511027.

[Ben+96c] C. H. **Bennett**, D. P. **DiVincenzo**, J. A. **Smolin**, and W. K. **Wootters**. "Mixed-state entanglement and quantum error correction". In: *Physical Review A* 54.5 (1996), p. 3824. DOI: 10.1103/PhysRevA.54.3824. arXiv: quant-ph/9604024.

[Ben+99] C. H. **Bennett**, D. P. **DiVincenzo**, C. A. **Fuchs**, T. **Mor**, E. **Rains**, P. W. **Shor**, J. A. **Smolin**, and W. K. **Wootters**. "Quantum nonlocality without entanglement". In: *Physical Review A* 59.2 (1999), p. 1070. DOI: 10.1103/PhysRevA.59.1070. arXiv: quant-ph/9804053.

[BFT17] M. **Berta**, O. **Fawzi**, and M. **Tomamichel**. "On variational expressions for quantum relative entropies". In: *Letters in Mathematical Physics* 107.12 (2017), pp. 2239–2265. DOI: 10.1007/s11005-017-0990-7. arXiv: 1512.02615 [quant-ph].

[Bha13] R. **Bhatia**. *Matrix analysis*. Vol. 169. Springer Science & Business Media, 2013. DOI: 10.1007/978-1-4612-0653-8.

[BP10] F. G. **Brandão** and M. B. **Plenio**. "A reversible theory of entanglement and its relation to the second law". In: *Communications in Mathematical Physics* 295.3 (2010), pp. 829–851. DOI: 10.1007/s00220-010-1003-1. arXiv: 0710.5827 [quant-ph].

[Bri+98] H.-J. **Briegel**, W. **Dür**, J. I. **Cirac**, and P. **Zoller**. "Quantum repeaters: the role of imperfect local operations in quantum communication". In: *Physical Review Letters* 81.26 (1998), p. 5932. DOI: 10.1103/PhysRevLett.81.5932. arXiv: quant-ph/9803056.

[Cal+98] A. **Calderbank**, E. **Rains**, P. **Shor**, and N. **Sloane**. "Quantum error correction via codes over GF (4)". In: *IEEE Transactions on Information Theory* 44.4 (1998), pp. 1369–1387. DOI: 10.1109/18.681315. arXiv: quant-ph/9608006.

[CF17] M. **Christandl** and R. **Ferrara**. "Private States, Quantum Data Hiding, and the Swapping of Perfect Secrecy". In: *Physical review letters* 119.22 (2017), p. 220506. DOI: 10.1103/PhysRevLett.119.220506. arXiv: 1609.04696 [quant-ph].

[CFL18] M. **Christandl**, R. **Ferrara**, and C. **Lancien**. "Random private quantum states". In: (2018). arXiv: 1801.02861 [quant-ph].

[Che+16] L. **Chen**, L. **Chen**, S. **Jordan**, Y.-K. **Liu**, D. **Moody**, R. **Peralta**, R. **Perlner**, and D. **Smith-Tone**. *Report on post-quantum cryptography*. 8105. US Department of Commerce, National Institute of Standards and Technology, 2016. DOI: 10.6028/NIST.IR.8105.

[Chr12] M. **Christandl**. "PPT square conjecture". BIRS workshop: Operator structures in quantum information theory, Problem G. 2012. URL: https://www.birs.ca/workshops/2012/12w5084/report12w5084.pdf.

[CKR09] M. **Christandl**, R. **König**, and R. **Renner**. "Postselection technique for quantum channels with applications to quantum cryptography". In: *Physical review letters* 102.2 (2009), p. 020504. DOI: 10.1103/PhysRevLett.102.020504. arXiv: 0809.3019 [quant-ph].

[CMW18] M. **Christandl**, A. **Müller-Hermes**, and M. M. **Wolf**. "When Do Composed Maps Become Entanglement Breaking?" In: *arXiv preprint arXiv:1807.01266* (2018). arXiv: 1807.01266 [quant-ph].

[CSW12] M. **Christandl**, N. **Schuch**, and A. **Winter**. "Entanglement of the antisymmetric state". In: *Communications in Mathematical Physics* 311.2 (2012), pp. 397–422. DOI: 10.1007/s00220-012-1446-7. arXiv: 0910.4151 [quant-ph].

[CW04] M. **Christandl** and A. **Winter**. ""Squashed entanglement": an additive entanglement measure". In: *Journal of mathematical physics* 45.3 (2004), pp. 829–840. DOI: 10.1063/1.1643788. arXiv: quant-ph/0308088.

[CW05] M. **Christandl** and A. **Winter**. "Uncertainty, monogamy and locking of quantum correlations". In: *Information Theory, 2005. ISIT 2005. Proceedings. International Symposium on*. IEEE. 2005, pp. 879–883. DOI: 10.1109/TIT.2005.853338. arXiv: quant-ph/0501090.

[Dat09] N. **Datta**. "Max-relative entropy of entanglement, alias log robustness". In: *International Journal of Quantum Information* 7.02 (2009), pp. 475–491. DOI: 10.1142/S0219749909005298. arXiv: 0807.2536 [quant-ph].

[DH99] M. J. **Donald** and M. **Horodecki**. "Continuity of relative entropy of entanglement". In: *Physics Letters A* 264.4 (1999), pp. 257–260. DOI: 10.1016/S0375-9601(99)00813-0. arXiv: quant-ph/9910002.

[DHR02] M. J. **Donald**, M. **Horodecki**, and O. **Rudolph**. "The uniqueness theorem for entanglement measures". In: *Journal of Mathematical Physics* 43.9 (2002), pp. 4252–4272. DOI: 10.1063/1.1495917. arXiv: quant-ph/0105017.

[DiV+00] D. P. **DiVincenzo**, P. W. **Shor**, J. A. **Smolin**, B. M. **Terhal**, and A. V. **Thapliyal**. "Evidence for bound entangled states with negative partial transpose". In: *Physical Review A* 61.6 (2000), p. 062312. DOI: 10.1103/PhysRevA.61.062312. arXiv: quant-ph/9910026.

[DiV+04] D. P. **DiVincenzo**, M. **Horodecki**, D. W. **Leung**, J. A. **Smolin**, and B. M. **Terhal**. "Locking classical correlations in quantum states". In: *Physical Review Letters* 92.6 (2004), p. 067902. DOI: 10.1103/PhysRevLett.92.067902. arXiv: quant-ph/0303088.

[DLT02] D. P. **DiVincenzo**, D. W. **Leung**, and B. M. **Terhal**. "Quantum data hiding". In: *IEEE Transactions on Information Theory* 48.3 (2002), pp. 580–598. DOI: 10.1109/18.985948. arXiv: quant-ph/0103098.

[Dob+11] K. **Dobek**, M. **Karpiński**, R. **Demkowicz-Dobrzański**, K. **Banaszek**, and P. **Horodecki**. "Experimental extraction of secure correlations from a noisy private state". In: *Physical review letters* 106.3 (2011), p. 030501. DOI: 10.1103/PhysRevLett.106.030501. arXiv: 1010.4575 [quant-ph].

[Dür+00] W. **Dür**, J. **Cirac**, M. **Lewenstein**, and D. **Bruß**. "Distillability and partial transposition in bipartite systems". In: *Physical Review A* 61.6 (2000), p. 062313. DOI: 10.1103/PhysRevA.61.062313. arXiv: quant-ph/9910022.

[Dür+99] W. **Dür**, H.-J. **Briegel**, J. **Cirac**, and P. **Zoller**. "Quantum repeaters based on entanglement purification". In: *Physical Review A* 59.1 (1999), p. 169. DOI: 10.1103/PhysRevA.59.169. arXiv: quant-ph/9808065.

[DW05] I. **Devetak** and A. **Winter**. "Distillation of secret key and entanglement from quantum states". In: *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*. Vol. 461. 2053. The Royal Society. 2005, pp. 207–235. DOI: 10.1098/rspa.2004.1372. arXiv: quant-ph/0306078.

[DY08] I. **Devetak** and J. **Yard**. "Exact cost of redistributing multipartite quantum states". In: *Physical Review Letters* 100.23 (2008), p. 230501. DOI: 10.1103/PhysRevLett.100.230501.

[Eke91] A. K. **Ekert**. "Quantum cryptography based on Bell's theorem". In: *Physical review letters* 67.6 (1991), p. 661. DOI: 10.1103/PhysRevLett.67.661.

[EW02] T. **Eggeling** and R. F. **Werner**. "Hiding classical data in multipartite quantum states". In: *Physical Review Letters* 89.9 (2002), p. 097905. DOI: 10.1103/PhysRevLett.89.097905. arXiv: quant-ph/0203004.

[Fan73] M. **Fannes**. "A continuity property of the entropy density for spin lattice systems". In: *Communications in Mathematical Physics* 31.4 (1973), pp. 291–294. DOI: 10.1007/BF01646490.

[Fek23] M. **Fekete**. "Über die Verteilung der Wurzeln bei gewissen algebraischen Gleichungen mit ganzzahligen Koeffizienten". In: *Mathematische Zeitschrift* 17.1 (1923), pp. 228–249. DOI: 10.1007/BF01504345.

[Fil14] S. N. **Filippov**. "PPT-inducing, distillation-prohibiting, and entanglement-binding quantum channels". In: *Journal of Russian Laser Research* 35.5 (2014), pp. 484–491. DOI: 10.1007/s10946-014-9451-2. arXiv: 1409.4036 [quant-ph].

[GB02] L. **Gurvits** and H. **Barnum**. "Largest separable balls around the maximally mixed bipartite quantum state". In: *Physical Review A* 66.6 (2002), p. 062311. DOI: 10.1103/PhysRevA.66.062311. arXiv: quant-ph/0204159.

[Got98] D. **Gottesman**. "The Heisenberg representation of quantum computers". In: *Proc. XXII International Colloquium on Group Theoretical Methods in Physics, 1998*. 1998, pp. 32–43. arXiv: quant-ph/9807006.

[Gur03] L. **Gurvits**. "Classical deterministic complexity of Edmonds' Problem and quantum entanglement". In: *Proceedings of the thirty-fifth annual ACM symposium on Theory of computing*. ACM. 2003, pp. 10–19. DOI: 10.1145/780542.780545. arXiv: quant-ph/0303055.

[HH04] T. **Hiroshima** and M. **Hayashi**. "Finding a maximally correlated state: Simultaneous Schmidt decomposition of bipartite pure states". In: *Physical Review A* 70.3 (2004), p. 030302. DOI: 10.1103/PhysRevA.70.030302. arXiv: quant-ph/0405107.

[HH99] M. **Horodecki** and P. **Horodecki**. "Reduction criterion of separability and limits for a class of distillation protocols". In: *Physical Review A* 59.6 (1999), p. 4206. DOI: 10.1103/PhysRevA.59.4206. arXiv: quant-ph/9708015.

[HHH00] P. **Horodecki**, M. **Horodecki**, and R. **Horodecki**. "Binding entanglement channels". In: *Journal of Modern Optics* 47.2-3 (2000), pp. 347–354. DOI: 10.1080/09500340008244047. arXiv: quant-ph/9904092.

[HHH01] M. **Horodecki**, P. **Horodecki**, and R. **Horodecki**. "Separability of n-particle mixed states: necessary and sufficient conditions in terms of linear maps". In: *Physics Letters A* 283.1-2 (2001), pp. 1–7. arXiv: quant-ph/9605038.

[HHH98] M. **Horodecki**, P. **Horodecki**, and R. **Horodecki**. “Mixed-state entanglement and distillation: is there a “bound” entanglement in nature?” In: *Physical Review Letters* 80.24 (1998), p. 5239. DOI: 10.1103/PhysRevLett.80.5239. arXiv: quant-ph/9801069.

[HLW06] P. **Hayden**, D. W. **Leung**, and A. **Winter**. “Aspects of generic entanglement”. In: *Communications in Mathematical Physics* 265.1 (2006), pp. 95–117. DOI: 10.1007/s00220-006-1535-6. arXiv: quant-ph/0407049.

[Hol73] A. S. **Holevo**. “Statistical problems in quantum physics”. In: *Proceedings of the second Japan-USSR Symposium on probability theory*. Springer. 1973, pp. 104–119. DOI: 10.1007/BFb0061483.

[Hor+05a] K. **Horodecki**, M. **Horodecki**, P. **Horodecki**, and J. **Oppenheim**. “Locking entanglement with a single qubit”. In: *Physical review letters* 94.20 (2005), p. 200501. DOI: 10.1103/PhysRevLett.94.200501. arXiv: quant-ph/0404096.

[Hor+05b] K. **Horodecki**, M. **Horodecki**, P. **Horodecki**, and J. **Oppenheim**. “Secure key from bound entanglement”. In: *Physical review letters* 94.16 (2005), p. 160502. DOI: 10.1103/PhysRevLett.94.160502. arXiv: quant-ph/0309110.

[Hor+08] K. **Horodecki** et al. “Low-dimensional bound entanglement with one-way distillable cryptographic key”. In: *IEEE Transactions on Information Theory* 54.6 (2008), pp. 2621–2625. DOI: 10.1109/TIT.2008.921709. arXiv: quant-ph/0506203.

[Hor+09] K. **Horodecki**, M. **Horodecki**, P. **Horodecki**, and J. **Oppenheim**. “General paradigm for distilling classical key from quantum states”. In: *IEEE Transactions on Information Theory* 55.4 (2009), pp. 1898–1929. DOI: 10.1109/TIT.2008.2009798. arXiv: quant-ph/0506189.

[Hor+16] K. **Horodecki**, P. **Ćwikliński**, A. **Rutkowski**, and M. **Studziński**. “Irreducible private states”. In: (2016). arXiv: 1612.08938 [quant-ph].

[Hor97] P. **Horodecki**. “Separability criterion and inseparable mixed states with positive partial transposition”. In: *Physics Letters A* 232.5 (1997), pp. 333–339. DOI: 10.1016/S0375-9601(97)00416-7. arXiv: quant-ph/9703004.

[HSR03] M. **Horodecki**, P. W. **Shor**, and M. B. **Ruskai**. “Entanglement breaking channels”. In: *Reviews in Mathematical Physics* 15.06 (2003), pp. 629–641. DOI: 10.1142/S0129055X03001709. arXiv: quant-ph/0302031.

[Jia+07] L. **Jiang**, J. M. **Taylor**, N. **Khaneja**, and M. D. **Lukin**. “Optimal approach to quantum communication using dynamic programming”. In: *Proceedings of the National Academy of Sciences* 104.44 (2007), pp. 17291–17296. DOI: 10.1073/pnas.0703284104. arXiv: arXiv:0710.5808 [quant-ph].

[KKB11] M. **Kleinmann**, H. **Kampermann**, and D. **Bruß**. “Asymptotically perfect discrimination in the local operation and classical communication paradigm”. In: *Physical Review A* 84.4 (2011), p. 042326. DOI: 10.1103/PhysRevA.84.042326. arXiv: 1105.5132 [quant-ph].

[KL96] E. **Knill** and R. **Laflamme**. “Concatenated quantum codes”. In: (1996). arXiv: quant-ph/9608012.

[KW04] D. **Kretschmann** and R. F. **Werner**. “Tema con variazioni: quantum channel capacity”. In: *New Journal of Physics* 6.1 (2004), p. 26. DOI: 10.1088/1367-2630/6/1/026. arXiv: quant-ph/0311037.

[LPW18] L. **Lami**, C. **Palazuelos**, and A. **Winter**. “Ultimate Data Hiding in Quantum Mechanics and Beyond”. In: *Communications in Mathematical Physics* (2018). DOI: 10.1007/s00220-018-3154-4. arXiv: 1703.03392 [quant-ph].

[LW14] K. **Li** and A. **Winter**. "Relative entropy and squashed entanglement". In: *Communications in Mathematical Physics* 326.1 (2014), pp. 63–80. DOI: 10.1007/s00220-013-1871-2. arXiv: 1210.3181 [quant-ph].

[MM+13] E. **Meckes**, M. **Meckes**, et al. "Spectral measures of powers of random matrices". In: *Electronic communications in probability* 18 (2013). DOI: 10.1214/ECP.v18-2551. arXiv: 1210.2681 [math.PR].

[Mon13] A. **Montanaro**. "Weak multiplicativity for random quantum channels". In: *Communications in Mathematical Physics* 319.2 (2013), pp. 535–555. DOI: 10.1007/s00220-013-1680-7. arXiv: 1112.5271 [quant-ph].

[MRW16] A. **Müller-Hermes**, D. **Reeb**, and M. M. **Wolf**. "Positivity of linear maps under tensor powers". In: *Journal of Mathematical Physics* 57.1 (2016), p. 015202. DOI: 10.1063/1.4927070. arXiv: 1502.05630 [quant-ph].

[Mur+16] S. **Muralidharan**, L. **Li**, J. **Kim**, N. **Lütkenhaus**, M. D. **Lukin**, and L. **Jiang**. "Optimal architectures for long distance quantum communication". In: *Scientific reports* 6 (2016), p. 20463. DOI: 10.1038/srep20463. arXiv: 1509.08435 [quant-ph].

[MW99] U. M. **Maurer** and S. **Wolf**. "Unconditionally secure key agreement and the intrinsic conditional information". In: *IEEE Transactions on Information Theory* 45.2 (1999), pp. 499–514. DOI: 10.1109/18.748999.

[MWW09] W. **Matthews**, S. **Wehner**, and A. **Winter**. "Distinguishability of quantum states under restricted families of measurements with an application to quantum data hiding". In: *Communications in Mathematical Physics* 291.3 (2009), pp. 813–843. DOI: 10.1007/s00220-009-0890-5. arXiv: 0810.2327 [quant-ph].

[MZ10] L. **Moravčíková** and M. **Ziman**. "Entanglement-annihilating and entanglement-breaking channels". In: *Journal of Physics A: Mathematical and Theoretical* 43.27 (2010), p. 275306. DOI: 10.1088/1751-8113/43/27/275306. arXiv: 1006.2502 [quant-ph].

[NC02] M. A. **Nielsen** and I. **Chuang**. *Quantum computation and quantum information*. AAPT, 2002. ISBN: 978-1-107-00217-3.

[OSS14] M. **Ozols**, G. **Smith**, and J. A. **Smolin**. "Bound entangled states with a private key and their classical counterpart". In: *Physical review letters* 112.11 (2014), p. 110502. DOI: 10.1103/PhysRevLett.112.110502. arXiv: 1305.0848 [quant-ph].

[Par70] J. L. **Park**. "The concept of transition in quantum mechanics". In: *Foundations of Physics* 1.1 (1970), pp. 23–33. DOI: 10.1007/BF00708652.

[Pet07] D. **Petz**. *Quantum information theory and quantum statistics*. Springer Science & Business Media, 2007.

[PH10] Ł. **Pankowski** and M. **Horodecki**. "Low-dimensional quite noisy bound entanglement with a cryptographic key". In: *Journal of Physics A: Mathematical and Theoretical* 44.3 (2010), p. 035301. DOI: 10.1088/1751-8113/44/3/035301. arXiv: 1008.1226 [quant-ph].

[Pia09] M. **Piani**. "Relative entropy of entanglement and restricted measurements". In: *Physical review letters* 103.16 (2009), p. 160504. DOI: 10.1103/PhysRevLett.103.160504. arXiv: 0904.2705 [quant-ph].

[Pin60] M. S. **Pinsker**. "Information and information stability of random variables and processes". In: (1960).

[Rai01] E. M. **Rains**. “A semidefinite program for distillable entanglement”. In: *IEEE Transactions on Information Theory* 47.7 (2001), pp. 2921–2933. DOI: 10.1109/18.959270. arXiv: quant-ph/0008047.

[Rai99a] E. M. **Rains**. “Bound on distillable entanglement”. In: *Physical Review A* 60.1 (1999), p. 179. DOI: 10.1103/PhysRevA.60.179. arXiv: quant-ph/9809082.

[Rai99b] E. M. **Rains**. “Rigorous treatment of distillable entanglement”. In: *Physical Review A* 60.1 (1999), p. 173. DOI: 10.1103/PhysRevA.60.173. arXiv: quant-ph/9809078.

[Ren07] R. **Renner**. “Symmetry of large physical systems implies independence of subsystems”. In: *Nature Physics* 3.9 (2007), p. 645. DOI: 10.1038/nphys684. arXiv: quant-ph/0703069.

[RGK05] R. **Renner**, N. **Gisin**, and B. **Kraus**. “Information-theoretic security proof for quantum-key-distribution protocols”. In: *Physical Review A* 72.1 (2005), p. 012332. DOI: 10.1103/PhysRevA.72.012332. arXiv: quant-ph/0502064.

[RS07] J. M. **Renes** and G. **Smith**. “Noisy processing and distillation of private quantum states”. In: *Physical review letters* 98.2 (2007), p. 020502. DOI: 10.1103/PhysRevLett.98.020502. arXiv: quant-ph/0603262.

[RW03] R. **Renner** and S. **Wolf**. “New bounds in secret-key agreement: The gap between formation and secrecy extraction”. In: *International Conference on the Theory and Applications of Cryptographic Techniques*. Springer. 2003, pp. 562–577. DOI: 10.1007/3-540-39200-9_35.

[Sal+10] L. **Salvail**, M. **Peev**, E. **Diamanti**, R. **Alléaume**, N. **Lütkenhaus**, and T. **Länger**. “Security of trusted repeater quantum key distribution networks”. In: *Journal of Computer Security* 18.1 (2010), pp. 61–87. DOI: 10.3233/JCS-2010-0373. arXiv: 0904.4072 [quant-ph].

[Sch60] J. **Schwinger**. “Unitary operator bases”. In: *Proceedings of the National Academy of Sciences* 46.4 (1960), pp. 570–579. DOI: 10.1073/pnas.46.4.570.

[Shi17] M. E. **Shirokov**. “Tight uniform continuity bounds for the quantum conditional mutual information, for the Holevo quantity, and for capacities of quantum channels”. In: *Journal of Mathematical Physics* 58.10 (2017), p. 102202. DOI: 10.1063/1.4987135. arXiv: 1512.09047 [quant-ph].

[Sho99] P. W. **Shor**. “Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer”. In: *SIAM review* 41.2 (1999), pp. 303–332. DOI: 10.1137/S0097539795293172. arXiv: quant-ph/9508027.

[SS07] G. **Smith** and J. A. **Smolin**. “Degenerate quantum codes for Pauli channels”. In: *Physical review letters* 98.3 (2007), p. 030501. DOI: 10.1103/PhysRevLett.98.030501. arXiv: quant-ph/0604107.

[Tan86] W.-S. **Tang**. “On positive linear maps between matrix algebras”. In: *Linear algebra and its applications* 79 (1986), pp. 33–44. DOI: 10.1016/0024-3795(86)90290-9.

[TDL01] B. M. **Terhal**, D. P. **DiVincenzo**, and D. W. **Leung**. “Hiding bits in Bell states”. In: *Physical review letters* 86.25 (2001), p. 5807. DOI: 10.1103/PhysRevLett.86.5807. arXiv: quant-ph/0011042.

[Tuc02] R. R. **Tucci**. “Entanglement of distillation and conditional mutual information”. In: (2002). arXiv: quant-ph/0202144.

[Tuc99] R. R. **Tucci**. “Quantum entanglement and conditional information transmission”. In: (1999). arXiv: quant-ph/9909041.

[Ved+97] V. **Vedral**, M. B. **Plenio**, M. A. **Rippin**, and P. L. **Knight**. “Quantifying entanglement”. In: *Physical Review Letters* 78.12 (1997), p. 2275. DOI: 10.1103/PhysRevLett.78.2275. arXiv: quant-ph/9702027.

[Ver14] S. **Verdú**. “Total variation distance and the distribution of relative information”. In: *Information Theory and Applications Workshop (ITA), 2014*. IEEE. 2014, pp. 1–3. DOI: 10.1109/ITA.2014.6804281.

[VT99] G. **Vidal** and R. **Tarrach**. “Robustness of entanglement”. In: *Physical Review A* 59.1 (1999), p. 141. DOI: 10.1103/PhysRevA.59.141. arXiv: quant-ph/9806094.

[VW02] G. **Vidal** and R. F. **Werner**. “Computable measure of entanglement”. In: *Physical Review A* 65.3 (2002), p. 032314. DOI: 10.1103/PhysRevA.65.032314. arXiv: quant-ph/0102117.

[WD16] X. **Wang** and R. **Duan**. “Improved semidefinite programming upper bound on distillable entanglement”. In: *Physical Review A* 94.5 (2016), p. 050301. DOI: 10.1103/PhysRevA.94.050301. arXiv: 1601.07940 [quant-ph].

[Wer89] R. F. **Werner**. “Quantum states with Einstein-Podolsky-Rosen correlations admitting a hidden-variable model”. In: *Physical Review A* 40.8 (1989), p. 4277. DOI: 10.1103/PhysRevA.40.4277.

[Wey27] H. **Weyl**. “Quantenmechanik und gruppentheorie”. In: *Zeitschrift für Physik* 46.1-2 (1927), pp. 1–46. DOI: 10.1007/BF02055756.

[WH02] J. **Walgate** and L. **Hardy**. “Nonlocality, asymmetry, and distinguishing bipartite states”. In: *Physical Review Letters* 89.14 (2002), p. 147901. DOI: 10.1103/PhysRevLett.89.147901. arXiv: quant-ph/0202034.

[Wie83] S. **Wiesner**. “Conjugate coding”. In: *ACM Sigact News* 15.1 (1983), pp. 78–88. DOI: 10.1145/1008908.1008920.

[Win16] A. **Winter**. “Tight uniform continuity bounds for quantum entropies: conditional entropy, relative entropy distance and energy constraints”. In: *Communications in Mathematical Physics* 347.1 (2016), pp. 291–313. DOI: 10.1007/s00220-016-2609-8. arXiv: 1507.07775 [quant-ph].

[WZ82] W. K. **Wootters** and W. H. **Zurek**. “A single quantum cannot be cloned”. In: *Nature* 299.5886 (1982), pp. 802–803. DOI: 10.1038/299802a0.

[Yan+09] 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”. In: *IEEE Transactions on Information Theory* 55.7 (2009), pp. 3375–3387. DOI: 10.1109/TIT.2009.2021373. arXiv: 0704.2236 [quant-ph].

[Zha07] Z. **Zhang**. “Estimating mutual information via Kolmogorov distance”. In: *IEEE Transactions on Information Theory* 53.9 (2007), pp. 3280–3282. DOI: 10.1109/TIT.2007.903122.