Connections between Quantum Key Distribution and Quantum Data Hiding

Mads Friis Frand-Madsen

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

Mads Friis Frand-Madsen
Department of Mathematical Sciences
University of Copenhagen
Universitetsparken 5
2100 København Ø

mffm@math.ku.dk
madsffm@gmail.com

Advisor:

Matthias Christandl
Department of Mathematical Sciences
University of Copenhagen

Assessment Committee:

Laura Mančinska
Department of Mathematical Sciences
University of Copenhagen

Karol Horodecki
Institute of Informatics
University of Gdansk

Eric Chitambar
Electrical and Computer Engineering
University of Illinois

ISBN: 978-87-7125-070-1

Submitted on 31st of January, 2023
Edited for print on 16th of March, 2023

Acknowledgements

Først, tak til min vejleder Matthias, tak til komiteen og tak til QMATH for at give mig muligheden for at tage en PhD. Det har været hårdt, det har været spændende og det har været svært, så særligt tak for din støtte og vejledning undervejs, Matthias.

Og fremmest, tak til Johannes, Maxime og Marie for god stemning over dårlig kaffe i kælderen på HCØ. Tak til Marie og Maxime for at lade mig drikke øl med løbeklubben, selv om jeg næsten aldrig er med ude at løbe. Tak til Lise for at moderere mit ego og altid grine højt. Tak til Ramona for gode snakke og akademiske diskussioner, og for at vise mig de schweiziske bjerge. Tak til Jakob og Simon for at være overbærende med den tidspressede udgave af mig, der dukkede op den sidste måned af skriveprocessen. Og tak til Simon og Paula for at finde flere fejl i mine første udkast til afhandlingen, end jeg burde indrømme her.

Contributions

The content of this thesis is based on work made during the past three years of my PhD with QMATH, the Centre for the Mathematics of Quantum Theory, at the University of Copenhagen.

This thesis includes unpublished work done in collaboration with my advisor, Matthias Christandl. The original contributions of this thesis are:

- A further establishment of the connection between privacy, entanglement, and quantum data hiding through the introduction of a *rate of hiding classical information.*
- A *generalization of the previous framework on the relationship between privacy and data hiding* [1].
- A *preliminary analysis of current commercially available quantum hardware* in terms of its relative performance in a randomness extraction scenario [2] in comparison with classical processing.

Abstract

We consider a classical communication setup with two spatially separated parties sharing a quantum system. In such a setup, entanglement is considered a valuable resource useful for a plethora of undertakings; for one, the generation of secret key, which is the primary concern of our work. In 2004, Horodecki et al. showed that the amount of secret key that can be extracted from noisy bipartite quantum states may exceed the amount of distillable maximally entangled quantum bits. Their work was based on the intuition of quantum data hiding but did not offer a quantitative relationship to this phenomenon. More recently, the connection between quantum key distribution and quantum data hiding was further developed by Christandl and Ferrara [1], who also used the connection as a tool to bound the rate at which secret key can be distributed in a network scenario.

In this thesis, we consider the question of the distinguishability of quantum states given an imperfect quantum memory. We define a rate at which secure classical data can be extracted from partially secure data with respect to an eavesdropper with imperfect quantum memory. We prove an upper bound and a lower bound on this rate under general assumptions in terms of entropic quantities. Finally, we introduce a rate at which classical data can be hidden from an eavesdropper with imperfect quantum memory, and we relate this back to the notion of quantum data hiding.

In a one-way classical communication scenario for quantum key distribution, we prove a lower bound on the distillable key of a generic state in terms of correlation and orthogonality of related quantum states. We elaborate on this by showing that the security of a key can be understood in terms of the orthogonality of related quantum states. Finally, we also consider the problem of extending the distance of quantum key distribution through intermediate stations, a setting referred to as a quantum key repeater. Here, we exhibit situations, where we can lower bound the performance of the optimal key repeater strategy in terms of local state discrimination. This strategy outperforms the naive approach of first distilling maximally entangled states between the individual parties followed by entanglement swapping in certain special cases.

Finally, we consider the current practical implications of the gap between the distinguishability of quantum states given perfect or imperfect quantum memory. In 2006, a large gap between the theoretical performance of an eavesdropper with classical or quantum side information was shown in a randomness extraction protocol [2]. Using current commercially available quantum hardware, we were, however, not able to achieve an advantage using a quantum protocol when compared with an error-free classical protocol in this scenario.

Resumé

Vi betragter en model med konventionel kommunikkation mellem to parter, som deler et kvantesystem. I en sådan model betragter vi sammenfiltring som en værdifuld ressource, der kan bruges til adskillige formål, eksempelvis at generere en hemmelig nøgle, hvilket er det primære fokus i vores arbejde. I 2004 viste Horodecki et al. at mængden af nøgle, som kan udvindes fra et delt kvantesystem kan overstige mængden af maksimalt sammenfiltrede kvantebits, der kan udvindes. Deres arbejde var baseret på intuition fra skjult kvantedata, men gav ikke en kvantitativ beskrivelse af relationen til dette fænomen. For nylig blev forbindelsen mellem kvantemekanisk distribution af hemmelige nøgler og skjult kvantedata yderligere udviklet af Christandl of Ferrara [1], der brugte forbindelsen som et værktøj til at begrænse raten hvormed en nøgle kan distribueres i et netværk.

I denne afhandling betragter vi spørgsmålet om skelnen mellem kvantetilstande givet uperfekt kvantehukommelse. Vi definerer en rate, hvormed sikker konventionel data kan udvindes fra delvist sikker konventionel data med hensyn til en fjendtlig, lyttende part. Vi viser en øvre begrænsning og en nedre begrænsning på denne rate under generelle antagelser i termer af entropiske størrelser. Endelig introducerer vi en rate, hvormed konventionel data kan skjules fra en fjendtlig, lyttende part med uperfekt hukommelse, og vi relaterer denne størrelse tilbage til begrebet om skjult kvantedata.

I en model for kvantemekanisk distribution af nøgle med en-vejs kommunikation viser vi en nedre begrænsning på mængden af nøgle, der kan udvindes i termer af korrelation og ortogonalitet af relaterede kvantetilstande. Vi uddyber dette ved at vise, at sikkerheden af en nøgle kan forstås i termer af ortogonalitet af relaterede kvantetilstande. Endelig betragter vi også problemet med at øge afstanden for kvantemekanisk distribution af nøgle gennem mellemliggende stationer, en model der refereres til som en gentager af nøgle. Her viser vi situationer, hvor vi kan nedre begrænse ydeevnen af en optimal gentager af nøgle i termer af lokal skelnen mellem tilstande. Strategien udkonkurrerer den naive tilgang, nemlig hvor parterne først udvinder maksimalt sammenfiltrede tilstande efterfulgt af et byt af sammenfiltring, i visse specielle tilfælde.

Endelig betragter vi de nuværende praktiske implikationer af forskellen mellem at skelne kvantetilstande givet perfekt eller uperfekt kvantehukommelse. I 2006 blev en stor forskel mellem den teoretiske ydeevne af en fjendlig, lyttende part med konventionel information eller kvanteinformation vist i en protokol for udvindelse af tilfældig data [2]. Med nuværende kommercielt tilgængelig hardware har vi ikke været i stand til at eftervise en fordel ved at bruge en kvantemekanisk protokol sammenlignet med fejlfri konventionel protokol i dette scenarie.

Contents

Introduction

Let us consider a communication setup involving the usual suspects, Alice and Bob, and let us suppose that they are concerned with the privacy of their communication. In the modern-day world, most digital communication is encrypted using a key, which is a random string of zeros and ones. If Alice and Bob share such a key, then Alice can use it to encrypt a message before sending it to Bob, and when Bob receives the encrypted message he again uses the key to decrypt and read the message. Whenever the key is as long as the message, there is a protocol where the privacy of the message is equivalent to the privacy of the key; in other words, if an eavesdropper, let us call her Eve, intercepts the encrypted message, then having little information about the key shared by Alice and Bob will ensure that she has little information about the contents of the message sent from Alice to Bob. The issue in this protocol is the requirement of a key as long as the message is; with the ever-increasing amount of digital communication and the scarcity of shared secret keys, this protocol is not practically feasible as a standalone solution to the problem of private communication!

One approach to addressing the issue of key length is to recognize the computational limitations of a malicious party, say, an eavesdropper as above. Suppose Alice and Bob share a short key. Then there are deterministic and efficient protocols for generating a long key based on some presumably computationally hard mathematical problem. Whenever Eve has negligible information about the short key, and she is computationally limited, then Eve will also have next to no information about the long key assuming the computational hardness of certain mathematical problems. Although this resolves the issue of the key length, when Alice and Bob reduce their level of paranoia to computational security, they still need a short shared secret key to begin with.

The question at hand for Alice and Bob is how to come up with an initial shared secret key. They find an answer to their perils in public-key cryptography. Here, we consider protocols using a pair of related keys, namely, a private key and a public key. Using only a short secret key, Bob may generate a pair of keys using a protocol based on some presumably computationally hard mathematical problem. The public key is distributed openly and Alice may use it to encrypt a message and send it to Bob, while Bob retains the private key for decryption of any received message. As long as Bob has a short secret key, there are computationally secure protocols for public-key generation assuming the computational hardness of certain mathematical problems. Using the protocol from above, Alice can now generate a long computationally secure key from a short initial key and send it to Bob using his public key. Upon receiving the encrypted long secret key, Bob may decrypt it using his private key; almost magically, Alice and Bob have now achieved a shared long secret key! This shows that having relaxed their level of paranoia to computational security, Alice and Bob are able to communicate privately with the modest requirement that Alice and Bob each have a short secret key unknown to the eavesdropper.

However, there is good reason for the rising level of paranoia when it comes to cryptographic matters. More than 20 years ago, it was shown that some mathematical problems used in cryptographic protocols due to their conjectured hardness are efficiently solvable on a quantum computer using Shor's algorithm [3]. In theory, this renders the solution provided above to Alice and Bob's problem of ensuring private communication insecure, and this realization is gradually becoming a practical security concern. Depending on the level of security demanded by Alice and Bob, we can argue that the proposed protocol was never actually secure; if we were to project the advances in computational power into the future, it will eventually be possible for an eavesdropper to extract the private keys used in public-key encryption today. The temporal aspect of security is by no means new, however, advances in the development of quantum hardware and quantum algorithms suggest that current cryptographic protocols are not as secure as expected.

Fortunately, Shor's algorithm was discovered more than 20 years ago much earlier than its implementation for cryptographically relevant parameters is possible. It gives us time. One approach is to stay within the paradigm of computational security and devise cryptographic protocols based on other mathematical problems conjectured to be quantum computationally hard. Keeping the temporal aspect of security in mind, this approach enables Alice and Bob to communicate privately in the presence of an eavesdropper with quantum processing power.

While emerging quantum technology poses a threat to the privacy of communication, it also provides a solution! Quantum mechanics allows Alice and Bob to obtain a shared secret key at a distance without making computational assumptions on the eavesdropper [4, 5]. We refer to such protocols as quantum key distribution protocols. The promise of quantum key distribution is to remove the temporal aspect of security, however, it comes at a very literal cost. There are high technological demands of devices for quantum key distribution, and with high levels of noise, it is safe to assume that the generation of secret keys will be relatively expensive in a foreseeable future. With these challenges in mind, it is exceedingly relevant to consider the optimality of quantum key distribution protocols. It has been shown that a secret key can be obtained from certain quantum states, where no pure entanglement can be achieved [6]. This discovery was inspired by a phenomenon referred to as quantum data hiding, namely, that some quantum states shared by two spatially separated parties are hard to distinguish with local access but easy to distinguish given access to the entire system. This indicates that we should look beyond protocols for pure entanglement distillation in our search for optimal protocols for quantum key distribution. As has always been the case, the choice of key generation protocol should reflect the desired level of security, where quantum computationally secure cryptography proposes a feasible solution to security in a foreseeable future, whereas quantum key distribution gives a time-independent security promise. However, the storage of the key or even its deletion after use may pose practical security concerns beyond those of the actual quantum key distribution protocol.

We begin by describing the formalism necessary in order to discuss fundamental concepts of finite dimensional quantum systems and states of quantum systems. We proceed to give a thorough description of operations on the state of a quantum system, in particular bipartite quantum systems. With the elementary definitions in place, we review various measures of distance and briefly discuss their relation to state discrimination. In order to support the flow of the following chapters, we state some fundamental results from quantum information and their adaptation to our setup.

In Chapter 2 we consider the task of obtaining secure data with respect to an eavesdropper. Initially, we suppose the protagonist Xavier has a partially secret key modeled by a random variable X, and Eve has information about X encoded into the state of a quantum system E. With no assumptions concerning capabilities of Eve, the task of obtaining a secret key with respect to Eve is referred to as strong randomness extraction. We reformulate the setup in terms of a classical-quantum state of a bipartite quantum system XE and refer to the task as secure state distillation. We begin by proving that the asymptotic rate of secure state distillation is given by the conditional entropy of X given E. With these introductory observations in mind, we discuss to what extent Xavier can improve his rate of secure state distillation given general assumptions on the capabilities of Eve. Here, we prove general upper and lower bounds on the rate at which Xavier can distill secure bits of information, and provide examples showing that Xavier can indeed distill more secure bits assuming, say, Eve has no quantum memory. We finish the chapter by introducing hiding states with respect to a restricted eavesdropper, where the classical data encoded into the state of system X could be readily inferred by an eavesdropper with no restrictions, while the restricted Eve remains oblivious concerning the state of system X.

In the next chapter, Chapter 3, we focus our attention on a setup with two parties Alice and Bob sharing a quantum system AB, and furthermore, they can perform local quantum operations and communicate classically. In this setup, we discuss the connections between private key, entanglement, and encoding classical data in quantum systems. We begin by proving a lower bound on the distillable private key of a certain class of states and proceed to apply this bound in the context of a quantum key repeater. Furthermore, our bound on private key distillation in conjunction with a known lower bound on entanglement distillation [1] provides a first indication of a connection between private key, entanglement, and hidden classical data. We shift gears in the second part of Chapter 3 in order to present the setup in terms of a classical-quantum state as in Chapter 2. To encode classical data into the state of a bipartite quantum system AB shared by Alice and Bob, we apply phase gates to a subsystem of AB. This results in a classical-quantum state of system XAB, where Alice and Bob are viewed as the eavesdropper restricted to local quantum operations and classical communication. Here, we show that the encoded classical data is retrievable provided global access to system AB, if and only if Alice and Bob can individually extract (possibly uncorrelated) secret keys. In the affirmative case, we show that the secret keys are completely uncorrelated if the state of system AB is separable, and there are indeed separable states showcasing this feat. This leads us to a definition of the phase hiding rate of a state, and we briefly discuss its possible connection to the difference between distillable private key and distillable entanglement.

Finally, in Chapter 4 we consider the performance of current quantum hardware compared with its classical counterpart in the following scenario. First, we sample a binary string of length n uniformly at random, and we allow an eavesdropper to read the string but only store $\log n$ (qu)bits of information. Then we perform a particular strong randomness extraction protocol [2], and hence ask the eavesdropper to guess one bit of the resulting string. Theoretically, an eavesdropper with a quantum computer always succeeds in this task (when n is even), while a classical eavesdropper may as well resort to guessing at random when n is large [2]. In an effort to identify a practically relevant example of an advantage using current quantum hardware, we compare the performance of a theoretically optimal quantum strategy on the IonQ quantum computer with the performance of various classical strategies.

Chapter 1

Preliminaries

We begin by describing the mathematical formalism necessary to discuss quantum mechanics in an information-theoretic context. We will throughout this thesis only be concerned with quantum systems with finitely many degrees of freedom, which we model by a finite-dimensional complex Hilbert space. We proceed to describe the notion of quantum states and channels before we move on to the more computationally relevant concepts of quantum bits and quantum gates. We give the definitions of various measures of distance and relate these concepts to the operationally relevant task of state discrimination. Finally, we define the von Neumann entropy and other related entropic quantities, which will play an integral role in our work.

1.1 Finite Dimensional Complex Hilbert Spaces

We adopt many of the choices of notation and terminology from [7], however, we deviate whenever it is necessary and whenever a simplification is possible.

A complex Hilbert space H is a vector space over a complex field with an inner product $\langle \cdot | \cdot \rangle_H$, such that the norm induced by the inner product makes H complete. We will only be concerned with finite-dimensional Hilbert spaces, and we denote the dimension of H by $d_H \in \mathbb{N}$.

1.1.1 Tensor Products of Hilbert Spaces

From any two finite-dimensional complex Hilbert spaces H, H', we can construct a new finite-dimensional complex Hilbert space, namely, their tensor product. The tensor product space $H \otimes H'$ is given by all linear combinations of $u \otimes u'$ for $u \in H$, $u' \in H'$ with the identifications of elements that ensure

$$H \times H' \to H \otimes H', \qquad (u, u') \mapsto u \otimes u'$$

is a bi-linear map. The tensor product $H \otimes H'$ is again a finite-dimensional complex Hilbert space with inner product given by $\langle u \otimes u', v \otimes v' \rangle_{H \otimes H'} = \langle u, v \rangle_H \langle u', v' \rangle_{H'}$ for all $u, v \in H$, $u', v' \in H'$. We use the shorthand notation $HH' := H \otimes H'$, and the tensor product space is of dimension $d_{HH'} = d_H d_{H'}$. On occasion, we will consider tensor products of several finite dimensional complex Hilbert spaces; here, we will consider the tensor product of

$u \in H_1H_2$ and $v \in H_1'H_2'$ as an element of $H_1H_1'H_2H_2'$, and in such cases, we will denote the implicit reordering by adding subscripts $u_{H_1H_2} \otimes v_{H_1'H_2'} \in H_1H_1'H_2H_2'$.

1.1.2 Linear Operators and Linear Maps

We denote the set of linear operators from one finite-dimensional complex Hilbert space H to another H' by $\mathrm{L}\left(H, H'\right)$, and for an operator $K \in \mathrm{L}\left(H, H'\right)$ we denote the adjoint operator by $K^\dagger \in \mathrm{L}\left(H', H\right)$. For $K \in \mathrm{L}\left(H, H'\right)$ and $K' \in \mathrm{L}\left(H', H''\right)$ we denote their composition by $K'K \in \mathrm{L}\left(H, H''\right)$.

For a finite dimensional complex Hilbert space H we denote by $\mathrm{L}\left(H\right) \coloneqq \mathrm{L}\left(H, H\right)$, and we denote by $\mathbb{1}_H \in \mathrm{L}\left(H\right)$ the identity operator. We say that an operator $K \in \mathrm{L}\left(H, H'\right)$ is isometric, if $K^\dagger K = \mathbb{1}_H$. An operator $K \in \mathrm{L}\left(H\right)$ is said to be normal if $KK^\dagger = K^\dagger K$. Furthermore, it is said to be unitary if K is invertible with $K^{-1} = K^\dagger$, that is, $KK^\dagger = K^\dagger K = \mathbb{1}_H$, and we denote by $\mathrm{U}\left(H\right)$ the set of all unitary operators on H. We say that K is Hermitian if $K = K^\dagger$, and K is additionally said to be positive semi-definite if $K = \tilde{K}^\dagger \tilde{K}$ for some $\tilde{K} \in \mathrm{L}\left(H\right)$, which we denote by $K \geq 0$. More generally, we write $K \geq K'$ if $K - K' \geq 0$ for any operators $K, K' \in \mathrm{L}\left(H\right)$. We say that K is a projection if K is a Hermitian operator satisfying $K^2 = K$.

For $K \in \mathrm{L}\left(H\right)$ we denote by $|K|$ the positive operator satisfying $KK^\dagger = |K|^2$. A Hermitian operator $K \in \mathrm{L}\left(H\right)$ has a spectral decomposition with real eigenvalues, so in particular it can be decomposed into a difference $K = K_+ - K_-$, where $K_+, K_- \geq 0$ and K_+, K_- are orthogonal, that is, $K_+K_- = K_-K_+ = 0$. If K is Hermitian, then $|K| = K_+ + K_-$.

Now note that $\mathrm{L}\left(H\right)$ is itself a finite-dimensional complex Hilbert space with inner product given by $\langle K, \tilde{K}\rangle = \mathrm{Tr}(K^\dagger \tilde{K})$ for $K, \tilde{K} \in \mathrm{L}\left(H\right)$. The norm induced from the inner product on $\mathrm{L}\left(H\right)$ is given by

$$\left\|K\right\|_2 \coloneqq \sqrt{\langle K, K\rangle} = \sqrt{\mathrm{Tr}\, K^\dagger K} = \sqrt{\mathrm{Tr}\,|K|^2}, \qquad K \in \mathrm{L}\left(H\right)$$

which we shall refer to as the 2-norm. Furthermore, we define the 1-norm on $\mathrm{L}\left(H\right)$ by

$$\left\|K\right\|_1 \coloneqq \mathrm{Tr}\,|K|, \qquad K \in \mathrm{L}\left(H\right),$$

and it is also referred to as the trace norm. For $\varepsilon > 0$ and two linear operators $K, K' \in \mathrm{L}\left(H\right)$ satisfying $\left\|K - K'\right\|_1 \leq \varepsilon$, we write K $\approx_\varepsilon K'$. Finally, it holds that

$$\left\|K\right\|_2 \leq \left\|K\right\|_1 \leq \sqrt{d_H}\left\|K\right\|_2$$

for all $K \in \mathrm{L}\left(H\right)$.

We will refer to a linear operator from $\mathrm{L}\left(H\right)$ to $\mathrm{L}\left(H'\right)$ as a linear map, and we denote by $\mathrm{LM}\left(H, H'\right)$ the set of all linear maps. We denote the identity map by $\mathrm{id}_H \in \mathrm{LM}\left(H, H\right)$, and the trace of a linear operator is denoted by $\mathrm{Tr}_H \in \mathrm{LM}\left(H, \mathbb{C}\right)$. When it is clear from the context, we will simply write $\mathrm{id} = \mathrm{id}_H$ and $\mathrm{Tr} = \mathrm{Tr}_H$. We note that the composition of linear maps is again a linear map, that is,

$$\Lambda_2 \circ \Lambda_1 = \Lambda \in \mathrm{LM}\left(H, H''\right), \qquad \Lambda_1 \in \mathrm{LM}\left(H, H'\right), \Lambda_2 \in \mathrm{LM}\left(H', H''\right),$$

and so for subsets $\Omega_1 \subseteq \mathrm{LM}\,(H, H')$, $\Omega_2 \subseteq \mathrm{LM}\,(H', H'')$ we define the composition

$$\Omega_2 \circ \Omega_1 := \left\{ \Lambda_2 \circ \Lambda_1 \,\middle|\, \Lambda_1 \in \mathrm{LM}\left(H, H'\right), \Lambda_2 \in \mathrm{LM}\left(H', H''\right) \right\} \subseteq \mathrm{LM}\left(H, H''\right).$$

Any linear map $\Lambda \in \mathrm{LM}\,(H, H')$ has a Kraus representation given by

$$\Lambda\,(X) = \sum_{i \in I} K_i X \tilde{K}_i^\dagger, \quad X \in \mathrm{L}\,(H)\,, \tag{1.1}$$

where I is a finite set of indices, and $K_i, \tilde{K}_i \in \mathrm{L}\,(H, H')$. The adjoint map $\Lambda^\dagger \in \mathrm{LM}\,(H', H)$ is given by

$$\Lambda^\dagger\,(Y) = \sum_{i \in I} K_i^\dagger Y \tilde{K}_i, \quad Y \in \mathrm{L}\left(H'\right).$$

A linear map $\Lambda \in \mathrm{LM}\,(H, H')$ is said to be positive, if $\Lambda\,(X) \geq 0$ for all positive semi-definite $X \in \mathrm{L}\,(H)$. Furthermore, we say that $\Lambda \in \mathrm{LM}\,(H, H')$ is completely positive, if $\Lambda \otimes \mathrm{id}_{H''} \in \mathrm{LM}\,(HH'', H'H'')$ is positive for all finite dimensional complex Hilbert spaces H''. It holds that Λ is completely positive if and only if

$$\Lambda\,(X) = \sum_{i \in I} K_i X K_i^\dagger, \quad X \in \mathrm{L}\,(H)$$

for linear operators $K_i \in \mathrm{L}\,(H, H')$ known as the Kraus operators. We denote by $\mathrm{CP}\,(H, H')$ the set of all completely positive linear maps. Along similar lines, we say that Λ is trace-preserving, if $\mathrm{Tr}_{H'} \Lambda\,(X) = \mathrm{Tr}_H X$ for all $X \in \mathrm{L}\,(H)$. It holds that Λ with Kraus representation given by (1.1) is trace-preserving if and only if $\sum_{i \in I} \tilde{K}_i^\dagger K_i = \mathbb{1}_H$. Finally, we denote by $\mathrm{CPTP}\,(H, H')$ the set of completely positive and trace-preserving linear maps. The set of completely positive maps and the set of completely positive and trace-preserving maps are both closed under composition, that is,

$$\begin{aligned} \mathrm{CP}\left(H', H''\right) \circ \mathrm{CP}\left(H, H'\right) &\subseteq \mathrm{CP}\left(H, H''\right), \\ \mathrm{CPTP}\left(H', H''\right) \circ \mathrm{CPTP}\left(H, H'\right) &\subseteq \mathrm{CPTP}\left(H, H''\right). \end{aligned}$$

If follows from Stinespring's dilation theorem that $\Lambda \in \mathrm{LM}\,(H, H')$ is completely positive and trace preserving (CPTP), if and only if there exists a finite-dimensional complex Hilbert space H'' and an isometric operator $V \in \mathrm{L}\,(H, H'H'')$ such that $\Lambda\,(X) = (\mathrm{id}_{H'} \otimes \mathrm{Tr}_{H''}) \left(VXV^\dagger\right)$.

For finite dimensional complex Hilbert spaces H_1, H_1' and H_2, H_2' we note that $\mathrm{L}\,(H_1, H_1')$ and $\mathrm{L}\,(H_2, H_2')$ are themselves finite dimensional complex Hilbert spaces. This allows us to form the tensor product of linear operators $\mathrm{L}\,(H_1, H_1') \otimes \mathrm{L}\,(H_2, H_2')$, which is isomorphic to $\mathrm{L}\,(H_1H_2, H_1'H_2')$ due to the association

$$K_1 \otimes K_2 \mapsto K, \qquad K\,(u \otimes v) = K_1 u \otimes K_2 v \text{ for } u \in H_1, v \in H_2,$$

and the fact that any $K \in \mathrm{L}\,(H_1H_2, H_1'H_2')$ can be written as a linear combination of tensor products $K_1 \otimes K_2$. Abusing this observation, we will write $K_1 \otimes K_2 = K \in \mathrm{L}\,(H_1H_2, H_1'H_2')$ for any $K_1 \in \mathrm{L}\,(H_1, H_1')$, $K_2 \in \mathrm{L}\,(H_2, H_2')$. As a noteworthy example of this, we will write $\mathbb{1}_{H_1} \otimes \mathbb{1}_{H_2} = \mathbb{1}_{H_1H_2}$.

Similarly, we have that the tensor product $\mathrm{LM}\left(H_1, H_1'\right) \otimes \mathrm{LM}\left(H_2, H_2'\right)$ is isomorphic to $\mathrm{LM}\left(H_1 H_2, H_1' H_2'\right)$ by an analogous association. Again abusing the notation, we will for $\Lambda_1 \in \mathrm{LM}\left(H_1, H_1'\right)$, $\Lambda_2 \in \mathrm{LM}\left(H_2, H_2'\right)$ write $\Lambda_1 \otimes \Lambda_2 = \Lambda$ for $\Lambda \in \mathrm{LM}\left(H_1 H_2, H_1' H_2'\right)$ given by $\Lambda\left(X, Y\right) = \Lambda_1\left(X\right) \otimes \Lambda_2\left(Y\right)$ for $X \in \mathrm{L}\left(H_1\right)$, $Y \in \mathrm{L}\left(H_2\right)$. This time we may along the same lines note that this abuse of notation allows us to write, say, $\mathrm{Tr}_{H_1 H_2} = \mathrm{Tr}_{H_1} \otimes \mathrm{Tr}_{H_2}$.

With the above notation in mind, we define for subsets $\Omega_1 \subseteq \mathrm{LM}\left(H_1, H_1'\right)$, $\Omega_2 \subseteq \mathrm{LM}\left(H_2, H_2'\right)$ the set of tensor products of linear maps by

$$\Omega_1 \otimes \Omega_2 \coloneqq \left\{\Lambda_1 \otimes \Lambda_2 \,|\, \Lambda_1 \in \Omega_1, \Lambda_2 \in \Omega_2\right\} \subseteq \mathrm{LM}\left(H_1 H_2, H_1' H_2'\right).$$

1.1.3 Dirac's bra-ket Notation and the Computational Basis

We will adhere to Dirac's bra-ket notation for a finite-dimensional complex Hilbert space H. First note that there exist an isomorphism associating elements $u \in H$ to linear operators $|u\rangle \in \mathrm{L}\left(\mathbb{C}, H\right)$, where the linear operator is given by $|u\rangle : z \mapsto zu$. We denote by $\langle v| \coloneqq |v\rangle^\dagger \in \mathrm{L}\left(H, \mathbb{C}\right)$, and note that $\langle v|$ is equivalently defined as the map $\langle v| : H \to \mathbb{C}$ given by $|u\rangle \mapsto \langle v\,|\,u\rangle$, where $\langle v\,|\,u\rangle \coloneqq \langle v|\,|u\rangle \in \mathrm{L}\left(\mathbb{C}, \mathbb{C}\right) \cong \mathbb{C}$ is equal to the inner product of u, v. Furthermore, for $|u\rangle \in \mathrm{L}\left(\mathbb{C}, H\right)$ and $\langle v| \in \mathrm{L}\left(H', \mathbb{C}\right)$ we denote the outer product by $|u\rangle\langle v| \in \mathrm{L}\left(H, H'\right)$, and note that the set of all such operators span $\mathrm{L}\left(H, H'\right)$.

We will abuse the notation and through the isomorphism above consider $|u\rangle$ as an element of H. The inner product of $|u\rangle, |v\rangle \in H$ is then also denoted by $\langle u\,|\,v\rangle = \langle u, v\rangle$. For $|u\rangle \in H$, $|v\rangle \in H'$ and $K \in \mathrm{L}\left(H, H'\right)$ we may furthermore denote by $\langle v|\,K\,|u\rangle \coloneqq \langle v, Ku\rangle = \left\langle K^\dagger v, u\right\rangle$, and finally we denote tensor products by $|uv\rangle \coloneqq |u\rangle\,|v\rangle \coloneqq |u\rangle \otimes |v\rangle$.

Any operator $K \in \mathrm{L}\left(H, H'\right)$ has a singular value decomposition, that is, there exists an integer $k \leq d_H, d_{H'}$, singular values s_i for $i = 1, \ldots, k$ and orthonormal vectors $|u_i\rangle \in H$ and $|v_j'\rangle \in H'$ for $i, j = 1, \ldots, k$ such that $K = \sum_{i=1}^k s_i |v_i'\rangle\langle u_i|$. In particular, for a normal operator $K \in \mathrm{L}\left(H\right)$ there exists eigenvalues $e_i \in \mathbb{C}$ for $i = 1, \ldots, d_H$ and an orthonormal basis of eigenvectors $|u_i\rangle \in H$, such that $K = \sum_{i=1}^{d_H} e_i\,|u_i\rangle\langle u_i|$. Furthermore, Hermitian operators have real eigenvalues, and positive semi-definite operators have non-negative eigenvalues. Finally, we note that any two commuting normal operators are simultaneously diagonalizable.

When we consider a d-dimensional complex Hilbert space H, the computational basis is simply a choice of an orthonormal basis consisting of d vectors, which we denote by $|0\rangle, \ldots, |d-1\rangle \in H$. More generally, if we consider tensor products of complex Hilbert spaces HH', the computational basis of HH' is given by the tensor products of vectors of the computational bases of H, H', respectively.

For some linear maps their definition is basis dependent. For a d-dimensional complex Hilbert space H we may thus pick a basis, and then we may define complex conjugation $K \mapsto \overline{K}$ with respect to this basis. Furthermore, we denote by $T_H \colon K \mapsto K^T$ the transposition of K with respect to the chosen basis.

1.2 Quantum Systems

A quantum system with finitely many degrees of freedom is given by a finite-dimensional complex Hilbert space H. If the system is completely isolated, then the state of the system is described by a unit vector in H, and the operations on the state of the quantum system are unitary operators. However, when studying quantum systems that are not completely isolated from their environment, we use the more general framework, where states of a quantum system are described by linear operators on the underlying Hilbert space, and operations on states are described by linear maps. We adopt the convention of referring to a quantum system H with finitely many degrees of freedom merely as a quantum system H, where the Hilbert space H is understood to be finite-dimensional.

We begin by introducing relevant notation and terminology concerning the state of a quantum system and operations on the state of a quantum system. Afterward, we proceed to introduce concepts that are central to the study of bipartite quantum systems.

1.2.1 States and Channels

The state of a quantum system given by a finite-dimensional complex Hilbert space H is given by a positive semi-definite operator $\rho \in \mathrm{L}(H)$, which is normalized with respect to the trace operator, that is, $\operatorname{Tr} \rho = 1$. We refer to ρ as a density operator, and the set of all density operators on H is denoted by $\mathcal{D}(H)$. We will also refer to ρ as a state of the quantum system H. More generally, we denote by $\mathcal{D}_{\leq}(H)$ the set of all positive semi-definite operators, which are subnormal with respect to the trace operator.

A state $\rho \in \mathcal{D}(H)$ is said to be pure if it is of rank 1, that is, $\rho = |\psi\rangle\langle\psi|$ for some $|\psi\rangle \in H$. Otherwise, ρ is said to be a mixed state. We denote the maximally mixed state by $\omega \coloneqq \frac{1}{d_H}\mathbb{1}_H$. For any state $\rho \in \mathcal{D}(H)$ we refer to any state $\rho' \in \mathcal{D}(HH')$ with the property $\operatorname{Tr}_{H'} \rho' = \rho$ as an extension. If ρ' is a pure state, we call it a purification. Given a spectral decomposition $\rho = \sum_{i \in I} p_i |\psi_i\rangle\langle\psi_i|$ we can always construct a purification $\rho^* = |\Psi\rangle\langle\Psi| \in \mathcal{D}(HH^*)$ with $H \cong H^*$, where $|\Psi\rangle$ is given by

$$|\Psi\rangle = \sum_{i=1}^{d_H} \sqrt{p_i}\, |\psi_i\rangle\, |i\rangle\,,$$

and where $\{|i\rangle\}_{i=1}^{d_H}$ denotes an orthonormal basis of H^*. We will likewise refer to $|\Psi\rangle$ as a purification of ρ. In general, a purification is not unique, but all purifications are related by a local isometry acting only on the purifying system.

For a quantum state ρ of joint quantum system HH' we denote by $\rho_H \coloneqq \operatorname{Tr}_{H'} \rho$, and we also denote by $\rho^H \coloneqq \rho_H$. Now suppose $\{|i\rangle\}_{i \in I}$ is an orthonormal basis of H. We refer to a state $\rho \in \mathcal{D}(HH')$ given by

$$\rho = \sum_{i=1}^{d_H} p_i\, |i\rangle\langle i| \otimes \rho_i, \qquad \rho_i \in \mathcal{D}\left(H'\right),$$

where $(p_i)_{i=1}^{d_H}$ is a probability distribution, as a classical-quantum state (cq-state).
A channel from one quantum system H to another H' is a completely positive and trace-preserving linear map, and we denote the set of all channels on H with output system H' by $\mathcal{C}_{\text{all}}(H\rangle H') \coloneqq \operatorname{CPTP}(H, H')$. Furthermore, we denote the set of all channels on H by

$$\mathcal{C}_{\text{all}}(H) \coloneqq \bigcup_{H'} \mathcal{C}_{\text{all}}\left(H\rangle H'\right),$$

and finally for an arbitrary set of channels $\mathcal{C}(H) \subseteq \mathcal{C}_{\text{all}}(H)$ we denote by

$$\mathcal{C}\left(H\rangle H'\right) \coloneqq \mathcal{C}(H) \cap \mathcal{C}_{\text{all}}\left(H\rangle H'\right).$$

In many cases, we will define a channel $\Lambda \in \mathcal{C}_{\text{all}}(H\rangle H')$ only on the set of density operators, but note that the definition extends to the entire Hilbert space by linearity. Furthermore, for channels $\Lambda \in \mathcal{C}_{\text{all}}(H)$ defined only on part of a joint quantum system HH' we denote by

$$\Lambda(\rho) \coloneqq (\Lambda \otimes \mathrm{id}_{H'})(\rho), \qquad \rho \in \mathcal{D}\left(HH'\right).$$

Now consider a quantum system H and joint quantum system $H'H''$, and suppose $\{|i\rangle\}_{i\in I}$ is an orthonormal basis of H'. A quantum instrument is a channel $\Lambda \in \mathcal{C}_{\text{all}}(H\rangle H'H'')$ given by

$$\Lambda(\rho) = \sum_{i\in I} |i\rangle\langle i| \otimes \Lambda_i(\rho),$$

where $\Lambda_i \in \mathrm{CP}(H, H'')$. Note that the state resulting from applying a quantum instrument is a cq-state. We denote the set of quantum instruments by $\mathcal{I}_{\text{all}}(H\rangle H'H'')$ with the convention that the first output system H' is classical.

Finally, a measurement of system H and outcome in H' is a channel $\Lambda \in \mathcal{C}_{\text{all}}(H\rangle H')$ given by

$$\Lambda(\rho) = \sum_{i\in I} \mathrm{Tr}(M_i\rho)\, |i\rangle\langle i|,$$

where $M_i \in \mathrm{L}(H)$ are positive semi-definite operators satisfying $\sum_{i\in I} M_i = \mathbb{1}_H$. We refer to $\{M_i\}_{i\in I}$ as the Positive Operator-Valued Measurement (POVM) representation of the measurement Λ. A projective measurement has POVM representation given by a set of projectors, and a measurement on a basis $\{|j\rangle\}_{j\in J}$ of H has POVM representation given by the projectors onto $|j\rangle$ for $j \in J$. We denote the set of all measurements of the quantum system H with outcome system H' by $\mathcal{M}_{\text{all}}(H\rangle H')$, and furthermore, we denote the set of all measurements of the quantum system H by

$$\mathcal{M}_{\text{all}}(H) \coloneqq \bigcup_{H'} \mathcal{M}_{\text{all}}\left(H\rangle H'\right).$$

1.2.2 Bipartite Quantum Systems

We will now introduce the relevant terminology to model the scenario of two spatially separated parties Alice and Bob sharing a quantum system AB. The spatial separation limits the set of permissible operations by Alice and Bob to channels acting on systems A and B individually. Bringing the systems together in order to perform a global operation requires quantum communication, which we will not consider here. Instead, we allow Alice and Bob to communicate classically, and we introduce the necessary terminology in order to discuss this communication setup. To stress the spatial distance between systems A and B, we denote the bipartite quantum system by $A:B$ whenever necessary.

Before we proceed, we introduce the following general notions particularly relevant when considering bipartite quantum systems. A product operator $K \in \mathrm{L}(A:B)$ is a linear operator given by $K = K_A \otimes K_B$ for $K_A \in \mathrm{L}(A)$, $K_B \in \mathrm{L}(B)$, and we denote the set of all product operators by $\mathrm{L}_{\text{prod}}(A:B)$. More generally, a separable operator is a convex combination of product operators, and we denote the set of all separable operators by $\mathrm{L}_{\mathrm{sep}}\left(A:B\right)$. Finally, for $K \in \mathrm{L}\left(A:B\right)$ we define the partial transpose with respect to system B as $\mathrm{K}^{\Gamma} := \left(\mathrm{id}_A \otimes T_B\right)(K)$, and we say that K is positive under partial transposition (PPT) if $K^{\Gamma} \geq 0$. Although the definition of transposition is basis dependent, we note that the PPT criterion is not [8]. We denote the set of all PPT operators by $\mathrm{L}_{\mathrm{ppt}}\left(A:B\right)$.

States and Channels on Bipartite Quantum Systems

The state ρ of a bipartite quantum system AB is said to be a product state, if it holds that $\rho_{AB} = \rho_A \otimes \rho_B$. We denote the set of all product states by $\mathcal{D}_{\mathrm{prod}}\left(A:B\right)$. More generally, a state is said to be separable, if it is a convex combination of product states, and we denote the set of separable states by $\mathcal{D}_{\mathrm{sep}}\left(A:B\right)$. If the state of a bipartite quantum system is not separable, we say that it is entangled.

For any state ρ of a bipartite quantum system AB we may consider $\rho^{\Gamma} = \left(\mathrm{id}_A \otimes T_B\right)(\rho)$. If $\rho^{\Gamma} \geq 0$, then we say that ρ is a PPT state, and we denote the set of all PPT states by $\mathcal{D}_{\mathrm{ppt}}\left(A:B\right)$. Since the transposition map is positive (although not completely positive) and trace-preserving, we furthermore have

$$\mathcal{D}_{\mathrm{prod}}\left(A:B\right) \subseteq \mathcal{D}_{\mathrm{sep}}\left(A:B\right) \subseteq \mathcal{D}_{\mathrm{ppt}}\left(A:B\right) \subseteq \mathcal{D}\left(AB\right).$$

Analogously, we define a product channel $\Lambda \in \mathcal{C}_{\mathrm{all}}\left(A:B\right)$ as $\Lambda = \Lambda_A \otimes \Lambda_B$ for $\Lambda_A \in \mathcal{C}_{\mathrm{all}}\left(A\right)$, $\Lambda_B \in \mathcal{C}_{\mathrm{all}}\left(B\right)$, and we denote by $\mathcal{C}_{\mathrm{prod}}\left(A:B\right)$ the set of all product channels. More generally, a channel is said to be separable, if it is a convex combination of completely positive channels, and we denote the set of separable channels by $\mathcal{C}_{\mathrm{sep}}\left(A:B\right)$. We denote the set of PPT channels by $\mathcal{C}_{\mathrm{ppt}}\left(A:B\right)$ [9], and finally we note that $\mathcal{C}_{\mathrm{prod}}\left(A:B\right) \subseteq \mathcal{C}_{\mathrm{sep}}\left(A:B\right) \subseteq \mathcal{C}_{\mathrm{ppt}}\left(A:B\right)$.

If we denote by $\mathcal{M}_{\mathrm{sep}}\left(A:B\right)$ and $\mathcal{M}_{\mathrm{ppt}}\left(A:B\right)$ the sets of all measurements, which can be realized as separable and PPT channels, respectively, we likewise have $\mathcal{M}_{\mathrm{sep}}\left(A:B\right) \subseteq \mathcal{M}_{\mathrm{ppt}}\left(A:B\right)$. A measurement $\Lambda \in \mathcal{M}_{\mathrm{all}}\left(AB\right)$ with POVM representation $\left\{M_i\right\}_{i\in I}$ is a separable or PPT measurement, if and only if M_i are separable or PPT operators for all $i \in I$.

Local Operations and Classical Communication

The spatial separation of two parties Alice and Bob limits the operations they can do on the joint quantum system AB. With no interaction, they can perform operations on their systems independently, and this is exactly the set of channels described by $\mathcal{C}_{\mathrm{prod}}\left(A:B\right)$. Furthermore, we denote that the set of all measurements Alice and Bob can perform with no interaction by $\mathcal{M}_{\mathrm{prod}}\left(A:B\right)$.

Let us now consider the set of channels Alice and Bob can perform when interacting through classical communication. If we only allow one-way communication from Alice to Bob, the most general channel is a quantum instrument applied to Alice's system in composition with a channel on Bob's system conditioned on the classical part of Alice's output system. More precisely, we define

$$\mathcal{C}_{A\to B}\left(A:B\rangle A':B'\right) := \bigcup_{M} \left[\mathrm{id}_{A'} \otimes \mathcal{C}_{\mathrm{all}}\left(MB\rangle B'\right)\right] \circ \left[\mathcal{I}_{\mathrm{all}}\left(A\rangle MA'\right) \otimes \mathrm{id}_B\right]$$

and

$$\mathcal{C}_{A\to B}\left(A:B\right) \coloneqq \bigcup_{A',B'} \mathcal{C}_{A\to B}\left(A:B\rangle A':B'\right).$$

Analogously, we define the set of channels, which can be realized by local operations and classical communication to Alice from Bob by

$$\mathcal{C}_{A\leftarrow B}\left(A:B\rangle A':B'\right) \coloneqq \bigcup_{M} \left[\mathcal{C}_{\mathrm{all}}\left(MA\rangle A'\right)\otimes \mathrm{id}_{B'}\right]\circ\left[\mathrm{id}_A \otimes \mathcal{I}_{\mathrm{all}}\left(B\rangle MB'\right)\right]$$

and

$$\mathcal{C}_{A\leftarrow B}\left(A:B\right) \coloneqq \bigcup_{A',B'} \mathcal{C}_{A\leftarrow B}\left(A:B\rangle A':B'\right).$$

To describe the set of channels Alice and Bob can perform when interacting through classical communication, we note that any such channel can be realized as a finite composition of channels realized by local operations and one-way classical communication. We define a single round of local operations and classical communication by

$$\mathcal{C}^{1}_{A\leftrightarrow B}\left(A:B\rangle A'':B''\right) \coloneqq \bigcup_{A',B'} \mathcal{C}_{A\leftarrow B}\left(A':B'\rangle A'':B''\right)\circ\mathcal{C}_{A\to B}\left(A:B\rangle A':B'\right),$$

and more generally we define n rounds of local operations and classical communication inductively by

$$\mathcal{C}^{n}_{A\leftrightarrow B}\left(A:B\rangle A'':B''\right) \coloneqq \bigcup_{A',B'} \mathcal{C}^{1}_{A\leftrightarrow B}\left(A':B'\rangle A'':B''\right)\circ\mathcal{C}^{n-1}_{A\leftrightarrow B}\left(A:B\rangle A':B'\right).$$

Finally, the set of all channels realizable by local operations and classical communication is given by

$$\mathcal{C}_{A\leftrightarrow B}\left(A:B\rangle A':B'\right) \coloneqq \bigcup_{n\in\mathbb{N}} \mathcal{C}^{n}_{A\leftrightarrow B}\left(A:B\rangle A':B'\right),$$

and

$$\mathcal{C}_{A\leftrightarrow B}\left(A:B\right) \coloneqq \bigcup_{A',B'} \mathcal{C}_{A\leftrightarrow B}\left(A:B\rangle A':B'\right).$$

We finish our discussion of channels on bipartite quantum systems by noting that

$$\mathcal{C}_{\mathrm{prod}}\left(A:B\right)\subseteq\mathcal{C}_{A\to B}\left(A:B\right)\subseteq\mathcal{C}_{A\leftrightarrow B}\left(A:B\right)\subseteq\mathcal{C}_{\mathrm{sep}}\left(A:B\right)\subseteq\mathcal{C}_{\mathrm{ppt}}\left(A:B\right),$$

and we use local channels as an umbrella term for all of the above sets of channels. We will use $\mathcal{C}_{\mathrm{loc}}\left(A:B\right)$ as a placeholder for any of the five sets of local measurements above. If we denote by $\mathcal{M}_{A\to B}\left(A:B\right)$, $\mathcal{M}_{A\leftrightarrow B}\left(A:B\right)$ the set of all measurements realizable by local operations and (one-way) classical communication, then we have

$$\mathcal{M}_{\mathrm{prod}}\left(A:B\right)\subseteq\mathcal{M}_{A\to B}\left(A:B\right)\subseteq\mathcal{M}_{A\leftrightarrow B}\left(A:B\right)\subseteq\mathcal{M}_{\mathrm{sep}}\left(A:B\right)\subseteq\mathcal{M}_{\mathrm{ppt}}\left(A:B\right),$$

and again we collectively refer to the sets of measurements above as local measurements, and we will use $\mathcal{M}_{\mathrm{loc}}\left(A:B\right)$ as a placeholder for any of the five sets of local measurements. An observation that will be particularly relevant later is that

$$\mathcal{M}_{A\to B}\left(A:B\rangle B'\right)=\bigcup_{A'}\mathcal{M}_{\mathrm{all}}\left(A'B\rangle B'\right)\circ\mathcal{M}_{\mathrm{all}}\left(A\rangle A'\right),$$

that is, any measurement implemented by local operations and one-way communication from Alice to Bob is some measurement of Alice's system followed by a measurement of Bob's system conditioned on the outcome of Alice's initial measurement.

1.3 Quantum Bits and Quantum Gates

A quantum bit is a 2-dimensional quantum system, which we shall refer to as a qubit. An m-qubit quantum system is the tensor product of m qubit systems. With this in mind, we introduce the notion of quantum gates as unitaries acting on one or more qubit(s) emphasizing the qubit states and quantum gates that will be of particular importance to our work. As the author found the structure of a similar introduction in [7] particularly amusing and elegant, we adopt this structure although the content is considerably different.

1.3.1 Single-Qubit Gates

Consider a qubit system A and denote by $\{|0\rangle, |1\rangle\}$ a choice of computational basis. The Pauli operators $X, Z \in \mathrm{U}(A)$ and the Hadamard gate $H \in \mathrm{U}(A)$ are given by

$$X := |1\rangle\langle 0| + |0\rangle\langle 1|, \qquad Z := |0\rangle\langle 0| - |1\rangle\langle 1|$$

and

$$H := \frac{1}{\sqrt{2}}\Big(|0\rangle\langle 0| + |0\rangle\langle 1| + |1\rangle\langle 0| - |1\rangle\langle 1| \Big).$$

Note that $XX = ZZ = HH = \mathbb{1}_A$. More generally, we may consider the 2^m-dimensional m-qubit system $A^m = A^{\otimes m}$, and for $\alpha \in \{0,1\}^m$ we define $X^\alpha, Z^\alpha, H^\alpha \in \mathrm{U}(A^{\otimes m})$ by

$$\begin{aligned} X^\alpha &:= \bigotimes_{i=1}^m X^{\alpha_i}, \qquad X^0 = \mathbb{1}_A, \quad X^1 = X \\ Z^\alpha &:= \bigotimes_{i=1}^m Z^{\alpha_i}, \qquad Z^0 = \mathbb{1}_A, \quad Z^1 = Z \\ H^\alpha &:= \bigotimes_{i=1}^m H^{\alpha_i}, \qquad H^0 = \mathbb{1}_A, \quad H^1 = H. \end{aligned}$$

We can equip $\{0,1\}^m$ with the structure of a vector space over the finite field with two elements, that is, we have $(\alpha+\beta)_i = \alpha_i + \beta_i$ for $\alpha, \beta \in \{0,1\}^m$ and $i \in \{1, \dots, m\}$, where the addition is carried out modulo 2. With this convention, it follows that

$$X^\alpha X^\beta = X^{\alpha+\beta}, \qquad Z^\alpha Z^\beta = Z^{\alpha+\beta}, \qquad H^\alpha H^\beta = H^{\alpha+\beta}$$

for all $\alpha, \beta \in \{0,1\}^m$. We define the corresponding channels $\mathcal{X}^\alpha, \mathcal{Z}^\alpha, \mathcal{H}^\alpha \in \mathcal{C}_{\text{all}}(A^m)$ by

$$\mathcal{X}^\alpha(\rho) := X^\alpha \rho X^{\alpha\dagger}, \qquad \mathcal{Z}^\alpha(\rho) := Z^\alpha \rho Z^{\alpha\dagger}, \qquad \mathcal{H}^\alpha(\rho) := H^\alpha \rho H^{\alpha\dagger},$$

and note that we analogously have

$$\mathcal{X}^\alpha \circ \mathcal{X}^\beta = \mathcal{X}^{\alpha+\beta}, \qquad \mathcal{Z}^\alpha \circ \mathcal{Z}^\beta = \mathcal{Z}^{\alpha+\beta}, \qquad \mathcal{H}^\alpha \circ \mathcal{H}^\beta = \mathcal{H}^{\alpha+\beta}$$

for all $\alpha, \beta \in \{0,1\}^m$. As a particular observation, we may note that if we measure the state ρ of an m-qubit system in the computational basis, then the resulting state is given by

$$\sum_{i \in \{0,1\}^m} \langle x| \rho |x\rangle \cdot |x\rangle\langle x| = \frac{1}{2^m} \sum_{\alpha \in \{0,1\}^m} \mathcal{Z}^\alpha(\rho).$$

We will make use of this observation at a later stage.

Finally, we define the rotation gate $Rz(\theta) \in \mathrm{U}(A)$ with parameter $\theta \in [0, 2\pi]$ given by

$$Rz(\theta) \coloneqq e^{-i\theta/2} |0\rangle\langle 0| + e^{i\theta/2} |1\rangle\langle 1|,$$

which we will use in Chapter 4.

1.3.2 Multiple-Qubit Gates

Two-Qubit Gates and Maximally Entangled States

Consider a two-qubit system AB, and denote by $\{|00\rangle, |01\rangle, |10\rangle, |11\rangle\}$ the computational basis. We refer to the two-qubit gate $CNOT \in \mathrm{U}(AB)$ given by

$$CNOT_{AB} \coloneqq |0\rangle\langle 0|_A \otimes \mathbb{1}_B + |1\rangle\langle 1|_A \otimes X_B$$

as the controlled-not gate, and we denote the corresponding channel by

$$\mathcal{CNOT}(\rho) = CNOT \rho CNOT^\dagger.$$

Now denote by $|\varphi\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$ with $|\varphi\rangle\langle\varphi| \in \mathcal{D}(AB)$ a maximally entangled qubit, and note that $|\varphi\rangle = CNOT(H \otimes \mathbb{1})|00\rangle$. We define

$$|\varphi_{ij}\rangle \coloneqq X^i Z^j |\varphi_{00}\rangle, \qquad i, j \in \{0, 1\},$$

and more generally, we denote by

$$|\varphi_{x\alpha}\rangle \coloneqq \bigotimes_{i=1}^{m} |\varphi_{x_i\alpha_i}\rangle, \qquad x, \alpha \in \{0, 1\}^m.$$

Let the swap gate $SWAP \in \mathrm{U}(AB)$ given by

$$SWAP \coloneqq |00\rangle\langle 00| + |01\rangle\langle 10| + |10\rangle\langle 01| + |11\rangle\langle 11|,$$

and note that $SWAP|\varphi_{ij}\rangle = (-1)^{i \cdot j}|\varphi_{ij}\rangle$. More generally, we have $SWAP^{\otimes m}|\varphi_{x\alpha}\rangle = (-1)^{\alpha \cdot x}|\varphi_{x\alpha}\rangle$, where $|\alpha| \coloneqq \sum_{i=1}^{m} \alpha_i$. We adapt the convention that $\varphi_{x\alpha} \in \mathcal{D}(AB)$ is the pure state $\varphi_{x\alpha} \coloneqq |\varphi_{x\alpha}\rangle\langle\varphi_{x\alpha}|$.

Three-Qubit Gates and Permutations of Computational Basis

Consider a three-qubit system ABC, and denote by $\{|x\rangle\}_{x \in \{0,1\}^3}$ the computational basis. We refer to the three-qubit gate $CCNOT \in \mathrm{U}(ABC)$ given by

$$CCNOT_{ABC} \coloneqq (\mathbb{1}_{AB} - |11\rangle\langle 11|_{AB}) \otimes \mathbb{1}_C + |11\rangle\langle 11|_{AB} \otimes X_C$$

as the Toffoli gate or the controlled-controlled-not gate, and we denote the corresponding channel by

$$\mathcal{CCNOT}(\rho) = CCNOT \rho CCNOT^\dagger.$$

Consider now an m-partite quantum system A^m given by the tensor product of m qubit systems, and denote by $\{|x\rangle\}_{x\in\{0,1\}^m}$ the computational basis. Let $\pi\colon\ \{0,1\}^m \to \{0,1\}^m$ be a permutation, and let the unitary $U_\pi \in \mathrm{U}\left(A^m\right)$ be defined by

$$U_\pi \left|x\right\rangle = \left|\pi\left(x\right)\right\rangle .$$

Let us now again equip $\{0,1\}^m$ with the structure of a vector space of the finite field with 2 elements. Let us now refer to the X-gate as the NOT-gate, and note that it has been shown [10] that

- if $\pi\left(x\right) = x + x_0$ for $x_0 \in \{0,1\}^m$, then U_π can be implemented by the NOT-gate.
- if π is affine, then U_π can be implemented by the NOT-gate and the $CNOT$-gate.
- any permutation π can be implemented using only the NOT-gate, the $CNOT$-gate and the $CCNOT$-gate.

Four-Qubit Gates and the Relative Phase of Bell States

Consider a four-qubit bipartite quantum system $A_1A_2 : B_1B_2$. We refer to the four-qubit gate $BNOT \in \mathrm{U}\left(A_1A_2B_1B_2\right)$ given by

$$BNOT_{A_1A_2B_1B_2} \coloneqq CNOT_{A_1A_2} \otimes CNOT_{B_1B_2}$$

as the bilateral $CNOT$-gate, and we denote the corresponding channel by

$$\mathcal{BNOT}\left(\rho\right) = BNOT \rho BNOT^\dagger.$$

For future reference, we state the following result.

Proposition 1. For all $x, i, y, j \in \{0,1\}$ we have

$$BNOT \left|\varphi_{xi}\right\rangle_{A_1B_1} \left|\varphi_{yj}\right\rangle_{A_2B_2} = \left|\varphi_{x(i+j)}\right\rangle_{A_1B_1} \left|\varphi_{(x+y)j}\right\rangle_{A_2B_2} .$$

Proof. This was shown in [1]. □

We may note that $\mathcal{BNOT}$ is a product channel with respect to the partition $A_1A_2 : B_1B_2$, which is the essential observation in the following result.

Corollary 2. Consider a $2m$-qubit system AB, where AB is given by the tensor product of two 2^m-dimensional Hilbert spaces. There exists a reversible channel $\mathcal{E}_{\text{Bell}} \in \mathcal{C}_{A\to B}\left(A : B\rangle A^*A : B^*B\right)$ such that

$$\mathcal{E}_{\text{Bell}}\left(|\varphi_{xi}\rangle\langle\varphi_{yj}|\right) = \frac{1}{2^m} \sum_{\alpha\in\{0,1\}^m} |\varphi_{x\alpha}\rangle\langle\varphi_{y\alpha}|^{A^*B^*} \otimes \left|\varphi_{x(i+\alpha)}\right\rangle\!\left\langle\varphi_{y(j+\alpha)}\right|^{AB} ,$$

where A^*, B^* are 2^m-dimensional quantum systems.

Proof. This was shown in [1]. □

Finally, we note that for $i, j \in \{0,1\}$ it holds that

$$\left(SWAP_{A_1A_2} \otimes SWAP_{B_1B_2}\right) BNOT \left|\varphi_{0i}\right\rangle \left|\varphi_{0j}\right\rangle = \left|\varphi_{0i}\right\rangle \left|\varphi_{0(i+j)}\right\rangle ,$$

which shows that a product channel implements a Z gate on the second Bell state conditional on the relative phase of the first Bell state. This transformation also has the more simple form of

$$CPH_2 \left|\varphi_{0i}\right\rangle \left|\varphi_{0j}\right\rangle \coloneqq \Big(\left|+\right\rangle\!\left\langle+\right| \otimes \mathbb{1} + \left|-\right\rangle\!\left\langle-\right| \otimes Z \Big)^{\otimes 2} \left|\varphi_{0i}\right\rangle \left|\varphi_{0j}\right\rangle = \left|\varphi_{0i}\right\rangle \left|\varphi_{0(i+j)}\right\rangle ,$$

which justifies thinking of this transformation as a controlled Z-gate on Bell states. This also invites generalizations, say, the three-qubit product gate $CPH_3 = U \otimes U$ with

$$U = \left|++\right\rangle\!\left\langle++\right| \otimes \mathbb{1} + \left|+-\right\rangle\!\left\langle+-\right| \otimes \begin{pmatrix} 1 & 0 \\ 0 & i \end{pmatrix} + \left|-+\right\rangle\!\left\langle-+\right| \otimes \begin{pmatrix} 1 & 0 \\ 0 & -i \end{pmatrix} + \left|--\right\rangle\!\left\langle--\right| \otimes Z,$$

which satisfies

$$CPH_3 \left|\varphi_{0^3ijk}\right\rangle = \left|\varphi_{0^2ij}\right\rangle \Big(\left|00\right\rangle + \left(-1\right)^{k+i\cdot j} i^{i-j} \left|11\right\rangle \Big),$$

and so resembles a Toffoli gate acting on the relative phase of Bell states. As we will not employ these gates later, we will not discuss them any further here.

1.4 Measures of Distance and State Discrimination

We begin this section by describing elementary measures of distance between linear operators and adapt the definitions to fit a setup with restricted sets of channels [11]. We proceed to discuss the distinguishability of states from two different perspectives, and we finish by showing an elementary relation between the two notions.

1.4.1 Trace norm, Fidelity and Purified Distance

For a finite dimensional Hilbert space A, recall the definition of the trace norm $\left\|\cdot\right\|_1 : \mathrm{L}\left(A\right) \to \mathbb{R}$ given by

$$\left\|\cdot\right\|_1 = \operatorname{Tr}\left|\cdot\right| .$$

This induces the trace distance $\delta\left(P, Q\right) \coloneqq \frac{1}{2}\left\|P - Q\right\|_1$ between operators $P, Q \in \mathrm{L}\left(A\right)$. We note that $\left\|\rho\right\|_1 = 1$ for all states $\rho \in \mathcal{D}\left(A\right)$, so it follows that

$$\delta\left(\rho, \rho'\right) = \frac{1}{2}\left\|\rho - \rho'\right\|_1 \leq \frac{1}{2}\Big(\left\|\rho\right\|_1 + \|\rho'\|_1 \Big) = 1$$

for all $\rho, \rho' \in \mathcal{D}\left(A\right)$. This upper bound is saturated by orthogonal states. For $\varepsilon > 0$ we denote by $\rho \approx_\varepsilon \rho'$ the statement that $\delta\left(\rho, \rho'\right) \leq \varepsilon$.

As a measure of overlap between positive operators, we denote the fidelity between two operators $P, Q \geq 0$ by

$$F\left(P, Q\right) \coloneqq \left\|\sqrt{P}\sqrt{Q}\right\|_1 ,$$

and we refer to F as the fidelity function. We may note that for all states $\rho \in \mathcal{D}\left(A\right)$ we have $F\left(\rho, \rho\right) = 1$, and for orthogonal operators, the fidelity function evaluates to zero. The relation between trace distance and the fidelity function is encaptured by the Fuchs-van de Graaf inequalities [12], namely,

$$1 - F(P, Q) \leq \delta(P, Q) \leq \sqrt{1 - F(P, Q)^2}, \qquad P, Q \geq 0.$$

Furthermore, we define the purified distance between subnormal states $\rho, \rho' \in \mathcal{D}_{\leq}(A)$ by

$$P(\rho, \rho') \coloneqq \sqrt{1 - \bar{F}(\rho, \rho')^2},$$

where $\bar{F}(\rho, \rho') \coloneqq F(\rho, \rho') + \sqrt{(1 - \operatorname{Tr}\rho)(1 - \operatorname{Tr}\rho')}$ denotes the generalized fidelity between subnormalized states $\rho, \rho' \in \mathcal{D}_{\leq}(A)$. It has been shown [13] that the purified distance is a metric and it is an upper bound on the trace distance, that is

$$\delta(\rho, \rho') \leq P(\rho, \rho')$$

for all $\rho, \rho' \in \mathcal{D}_{\leq}(A)$. For $\varepsilon \geq 0$ we denote by

$$\mathcal{B}^{\varepsilon}(\rho) \coloneqq \left\{\rho' \in \mathcal{D}_{\leq}(A) \,\middle|\, P(\rho, \rho') \leq \varepsilon\right\}, \qquad \rho \in \mathcal{D}_{\leq}(A),$$

that is, the closed ball of radius $\varepsilon \geq 0$ with respect to the purified distance around ρ in $\mathcal{D}_{\leq}(A)$.

1.4.2 Restricted Trace Norm

For any quantum system A we may consider a set of channels $\mathcal{C}(A)$, and denote by [11]

$$\|\cdot\|_{\mathcal{C}(A)} \coloneqq \sup_{\Lambda \in \mathcal{C}(A)} \|\Lambda(\cdot)\|_1,$$

and we note that due to the monotonicity of the trace norm, we have $\|\cdot\|_{\mathcal{C}(A)} \leq \|\cdot\|_1$ for any set of channels $\mathcal{C}(A)$. We make the trivial observation that whenever the identity channel id_A is an element of $\mathcal{C}(A)$, then it is the optimal choice of channel resulting in $\|\cdot\|_{\mathcal{C}(A)} = \|\cdot\|_1$. Note, however, that if we restrict ourselves to measurements, then we exclude the identity from the set of eligible channels.

We will be particularly concerned with bipartite quantum systems AB, and here we take note of the following hierarchy of restricted norms:

$$\begin{aligned}\|\cdot\|_{\mathcal{M}_{\mathrm{prod}}(A:B)} \leq \|\cdot\|_{\mathcal{M}_{A\to B}(A:B)} \leq \|\cdot\|_{\mathcal{M}_{A\leftrightarrow B}(A:B)} \leq \|\cdot\|_{\mathcal{M}_{\mathrm{sep}}(A:B)} \leq \|\cdot\|_{\mathcal{M}_{\mathrm{ppt}}(A:B)} \\ \leq \|\cdot\|_{\mathcal{M}_{\mathrm{all}}(AB)}.\end{aligned}$$

Furthermore, we note that the norm restricted to a partial measurement, namely, the measurement of system A, is an upper bound on the $\mathcal{M}_{A\to B}(A:B)$-restricted norm due to the monotonicity of the trace norm. More precisely, we have

$$\|\cdot\|_{\mathcal{M}_{A\to B}(A:B)} \leq \|\cdot\|_{\mathcal{M}(A)} \eqqcolon \|\cdot\|_A,$$

which is an observation we will make use of later.

1.4.3 State Discrimination

We introduce the relevant notation and terminology in order to discuss the discrimination of states with respect to certain sets of measurements. We begin by considering a quantum system E. Let Σ be an alphabet, that is, a finite set of letters, and consider the ensemble of state $\{p_x \rho_x\}_{x\in\Sigma} \subseteq \mathcal{D}_{\leq}(E)$, where $(p_x)_{x\in\Sigma}$ is a probability distribution. We define the corresponding classical-quantum state (cq-state) by

$$\rho^{XE} := \sum_{x\in\Sigma} p_x \left|x\right\rangle\!\left\langle x\right|^X \otimes \rho_x^E \in \mathcal{D}(XE),$$

where X is a quantum system with orthonormal basis $\{|x\rangle\}_{x\in\Sigma}$.

Definition 3. Let $\rho_{XE} \in \mathcal{D}(XE)$ be a cq-state given by

$$\rho = \sum_{x\in\Sigma} p_x \left|x\right\rangle\!\left\langle x\right| \otimes \rho_x, \qquad \rho_x \in \mathcal{D}(E)$$

where $\{|x\rangle\}_{x\in\Sigma}$ is an orthonormal basis of the quantum system X. We denote the optimal probability of correctly guessing the state of X based on a measurement of system E by

$$\Pr_{\text{guess}}(X|E)_\rho := \sup_{\{M_x\}_{x\in\Sigma}\text{ POVM}} \sum_{x\in\Sigma} p_x \operatorname{Tr} M_x \rho_x.$$

For a set of channels $\mathcal{C}(E\rangle E')$ we denote by

$$\Pr_{\text{guess}}^{\mathcal{C}(E\rangle E')}(X|E)_\rho := \sup_{\Lambda\in\mathcal{C}(E\rangle E')} \Pr_{\text{guess}}\left(X|E'\right)_{\Lambda(\rho)},$$

and, more generally, for an arbitrary set of channels $\mathcal{C}(E)$ we denote by

$$\Pr_{\text{guess}}^{\mathcal{C}(E)}(X|E)_\rho := \sup_{E'} \Pr_{\text{guess}}^{\mathcal{C}(E\rangle E')}(X|E)_\rho.$$

Analogous to our consideration of norms, we make the observation that if the identity channel is an element of $\mathcal{C}(E)$, then it is in fact the optimal choice of a channel. If we restrict ourselves to sets of measurements, however, we exclude the identity from the set of eligible channels as before. Identifying the optimal choice of channel is in general a non-trivial problem, however, when we consider the set of all measurements $\mathcal{M}_{\text{all}}(E)$, the pretty good measurement $\Lambda^{PG_\rho} \in \mathcal{C}(E)$ [14] with POVM representation $\left\{M_x^{PG_\rho}\right\}_{x\in\Sigma}$ given by

$$M_x^{PG_\rho} := \rho^{-1/2} p_x \rho_x \rho^{-1/2}$$

is indeed a pretty good choice of measurement as illustrated by the following statement: If we let $\rho_{XE} \in \mathcal{D}(XE)$ be a cq-state given by

$$\rho = \sum_{x\in\Sigma} p_x \left|x\right\rangle\!\left\langle x\right| \otimes \rho_x, \qquad \rho_x \in \mathcal{D}(E)$$

where $\{|x\rangle\}_{x\in\Sigma}$ is an orthonormal basis of the quantum system X, then

$$\Pr_{\text{guess}}(X|E)_\rho \leq \sqrt{\sum_{x\in\Sigma} p_x \operatorname{Tr} M_x^{PG_\rho} \rho_x}.$$

We will be particularly concerned with bipartite quantum systems AB, where we consider the restriction to local measurements. In this scenario, we note that all sets of local measurements

$$\mathcal{M}_{A\to B}\left(A:B\right), \quad \mathcal{M}_{A\leftrightarrow B}\left(A:B\right), \quad \mathcal{M}_{\mathrm{sep}}\left(A:B\right), \quad \mathcal{M}_{\mathrm{ppt}}\left(A:B\right)$$

are closed under post-composition with the set of all measurements, if we disregard whether the measurement output is at Alice's or Bob's. Thus, the quantity $\Pr_{\mathrm{guess}}^{\mathcal{M}_{\mathrm{loc}}(A:B)}\left(X|AB\right)_\rho$ describes the optimal probability of correctly guessing the state of X by a measurement from $\mathcal{M}_{\mathrm{loc}}\left(A:B\right)$ of system AB. Finally, we note for future reference that

$$\Pr_{\mathrm{guess}}^{\mathcal{M}_{\mathrm{all}}(A)}\left(X|AB\right)_\rho = \Pr_{\mathrm{guess}}^{\mathcal{M}_{A\to B}(A:B)}\left(X|AB\right)_\rho .$$

Let us now consider a different perspective on how to quantify the distinguishability of states of a quantum system E. For an ensemble $\left\{p_x\rho_x\right\}_{x\in\Sigma} \subseteq \mathcal{D}_{\leq}\left(E\right)$ we may note that the states are completely indistinguishable if $p_x = \frac{1}{|\Sigma|}$ and $\rho_x = \rho_{x'}$ for all $x, x' \in \Sigma$. The cq-state corresponding to this scenario is given by

$$\omega_X \otimes \rho_E = \frac{1}{|\Sigma|}\sum_{x\in\Sigma} |x\rangle\langle x|_X \otimes \rho_E \in \mathcal{D}\left(XE\right), \tag{1.2}$$

where we recall $\omega_X = \frac{1}{|\Sigma|}\mathbb{1}_X$. We quantify the distinguishability of the states in the ensemble in terms of proximity to the cq-state in (1.2). To this end, we introduce the distance to uniform given by

$$\Delta\left(X|E\right)_\rho := \frac{1}{2}\left\|\rho_{XE} - \omega_X \otimes \rho_E\right\|_1 ,$$

Analogous to previous considerations, we may consider the distinguishability of states when restricted to some set of channels $\mathcal{C}\left(E\right)$, so we define the $\mathcal{C}\left(E\right)$-restricted distance to uniform by

$$\Delta_{\mathcal{C}(E)}\left(X|E\right)_\rho := \frac{1}{2}\left\|\rho_{XE} - \omega_X \otimes \rho_E\right\|_{\mathcal{C}(E)} .$$

With this in mind, we show the following result relating the two approaches to quantifying distinguishability.

Proposition 4. Let $\rho_{XE} \in \mathcal{D}\left(XE\right)$ be a cq-state given by

$$\rho = \sum_{x\in\Sigma} p_x\, |x\rangle\langle x| \otimes \rho_x, \qquad \rho_x \in \mathcal{D}\left(E\right)$$

where $\left\{|x\rangle\right\}_{x\in\Sigma}$ is an orthonormal basis of quantum system X. Then

$$\Pr_{\mathrm{guess}}^{\mathcal{C}(E)}\left(X|E\right)_\rho - \frac{1}{|\Sigma|} \leq \Delta_{\mathcal{C}(E)}\left(X|E\right)_\rho$$

with equality whenever $|\Sigma| = 2$.

Proof. We begin by proving the result for $\mathcal{C}(E) = \mathcal{C}_{\text{all}}(E)$. Let $\Lambda \in \mathcal{M}(E)$ be a measurement with POVM representation $\{M_x\}_{x\in\Sigma}$ satisfying

$$\Pr_{\text{guess}}(X|E)_\rho = \sum_{x\in\Sigma} p_x \operatorname{Tr} M_x \rho_x .$$

Then

$$\begin{aligned}
\Delta(X|E)_\rho &\geq \Delta\left(X\middle|X'\right)_{\Lambda(\rho)} \\
&= \frac{1}{2}\left\| \sum_{x,x'\in\Sigma} p_x \operatorname{Tr}\left(M_{x'}\rho_x\right)|xx'\rangle\langle xx'| - \sum_{x,x'\in\Sigma} \frac{1}{|\Sigma|}\operatorname{Tr}\left(M_{x'}\rho\right)|xx'\rangle\langle xx'| \right\|_1 \\
&= \frac{1}{2}\sum_{x,x'\in\Sigma}\left| p_x \operatorname{Tr} M_{x'}\rho_x - \frac{1}{|\Sigma|}\operatorname{Tr} M_{x'}\rho \right| \\
&\geq \frac{1}{2}\sum_{x\in\Sigma}\left(p_x \operatorname{Tr} M_x\rho_x - \frac{1}{|\Sigma|}\operatorname{Tr} M_x\rho \right) + \frac{1}{2}\sum_{x,x'\in\Sigma, x\neq x'}\left(\frac{1}{|\Sigma|}\operatorname{Tr} M_{x'}\rho - p_x \operatorname{Tr} M_{x'}\rho_x \right) \\
&= \sum_{x\in\Sigma}\left(p_x \operatorname{Tr} M_x\rho_x - \frac{1}{|\Sigma|}\operatorname{Tr} M_x\rho \right) \\
&= \Pr_{\text{guess}}(X|E)_\rho - \frac{1}{|\Sigma|}.
\end{aligned}$$

This proves the desired inequality when $\mathcal{C}(E) = \mathcal{C}_{\text{all}}(E)$. It follows from the Holevo-Helstrom Theorem [15] that we have equality whenever $|\Sigma| = 2$.

From the argument above it follows that for an arbitrary $\Lambda \in \mathcal{C}(E\rangle E')$, we have

$$\Pr_{\text{guess}}\left(X\middle|E'\right)_{\Lambda(\rho)} - \frac{1}{|\Sigma|} \leq \Delta\left(X\middle|E'\right)_{\Lambda(\rho)}$$

with equality whenever $|\Sigma| = 2$. Taking supremum over all $\Lambda \in \mathcal{C}(E)$ yields the desired result. □

The statement of Proposition 4 gives a relation between the two notions of distinguishability introduced here. We will apply this in Chapter 2. As a final remark, we note that in the scenario of a bipartite quantum system AB, we have

$$\Pr_{\text{guess}}^{\mathcal{M}_{A\to B}(A:B)}(X|AB)_\rho - \frac{1}{|\Sigma|} = \Pr_{\text{guess}}^{\mathcal{M}_{\text{all}}(A)}(X|AB)_\rho - \frac{1}{|\Sigma|} \leq \Delta_{\mathcal{M}_{\text{all}}(A)}(X|AB)_\rho ,$$

and we will use this observation later.

1.5 On the Topic of Entropy

In the following, we define several elementary entropic quantities and proceed to adapt the definitions to a setup with restricted sets of channels analogous to the previous section. Furthermore, we state the quantum asymptotic equipartition property [16] and the leftover hash lemma [17] in versions adapted to our future needs.

1.5.1 Entropic Quantities

Let us begin by letting $\eta(x) \coloneqq -x \log x$ for $x \geq 0$ with the convention $\eta(0) = 0$. Here, we denote by log the base 2 logarithm. This allows us to concisely express the binary entropy as

$$h(x) \coloneqq \eta(x) + \eta(1-x), \qquad x \in [0,1].$$

Let A be a quantum system, and consider a state $\rho \in \mathcal{D}(A)$ with spectral decomposition $\rho = \sum_i p_i |\psi_i\rangle\langle\psi_i|$. Then we may define the von Neumann entropy of ρ on system A by

$$\mathrm{H}(A)_\rho \coloneqq \sum_i \eta(p_i) = -\sum_i p_i \log p_i,$$

and note that it is non-negative and upper bounded by $\log d_A$. In particular, the von Neumann entropy of a pure state is zero, and the upper bound is attained by the maximally mixed state. Furthermore, we note that the von Neumann entropy is additive on tensor products of states and invariant under isometric operators.

A key tool in showing many elementary properties of the von Neumann entropy is the relative entropy $D(\cdot\|\cdot)$. It is defined for all positive operators $P, Q \geq 0$ and given by

$$D(P\|Q) \coloneqq \begin{cases} \operatorname{Tr} P \log Q - \operatorname{Tr} P \log Q, & \operatorname{supp} P \subseteq \operatorname{supp} Q \\ \infty, & \text{otherwise}, \end{cases}$$

where $\operatorname{supp} P$ denotes the support of P. The relative entropy satisfies a data processing inequality, namely,

$$D(\Lambda(\rho)\|\Lambda(\sigma)) \leq D(\rho\|\sigma), \qquad \rho, \sigma \in \mathcal{D}(E)$$

for all $\Lambda \in \mathcal{C}_{\text{all}}(E)$, which is an essential tool in proving the more delicate data processing inequalities we shall see below.

With the introduction of the von Neumann entropy and the relative entropy in place, we may introduce the following entropic quantities for a bipartite state $\rho \in \mathcal{D}(AB)$:

- The conditional entropy of A given B is $\mathrm{H}(A|B)_\rho \coloneqq \mathrm{H}(AB)_\rho - \mathrm{H}(B)_\rho$.
- The mutual information of A and B is $\mathrm{I}(A:B)_\rho \coloneqq \mathrm{H}(A)_\rho - \mathrm{H}(A|B)_\rho$.

For a channel $\Lambda \in \mathcal{C}_{\text{all}}(B\rangle B')$ we have that

$$\mathrm{H}(A|B)_\rho \leq \mathrm{H}\left(A|B'\right)_{\Lambda(\rho)}.$$

which in turn implies $\mathrm{I}(A:B)_\rho \geq \mathrm{I}(A:B')_{\Lambda(\rho)}$. Due to symmetry in the definition of the mutual information of systems A and B an analogous result holds for the mutual information when considering channels $\Lambda \in \mathcal{C}_{\text{all}}(A)$.

Consider now the special case of $\rho \in \mathcal{D}(XE)$ being a cq-state given by

$$\rho = \sum_{x \in \Sigma} p_x |x\rangle\langle x| \otimes \rho_x,$$

where $\{|x\rangle\}_{x\in\Sigma}$ is an orthonormal basis of X. Let $f \colon \Sigma \to \Sigma'$ be a function for some finite set Σ', and denote by $\Lambda_f \in \mathcal{C}_{\text{all}}(X\rangle X')$ the measurement with POVM representation

$M_{x'} = \sum_{x \in f^{-1}(x)} |x\rangle\langle x|$, where $\{|x'\rangle\}_{x' \in \Sigma'}$ is an orthonormal basis of quantum system X'. To ease the notation later, we denote by

$$f(\rho) \coloneqq \Lambda_f(\rho) = \sum_{x \in \Sigma} |f(x)\rangle\langle f(x)|^{X'} \otimes \rho_x^E.$$

It follows from operator concavity of η [18] that

$$\mathrm{H}\left(X' \middle| E\right)_{f(\rho)} \leq \mathrm{H}(X|E)_\rho.$$

Previously, we phrased the distinguishability of states belonging to an ensemble of states $\{p_x \rho_x\}_{x \in \Sigma} \subseteq \mathcal{D}(E)$ in terms of the distance between the cq-state ρ_{XE} and the uniform distribution on X uncorrelated with E, that is, $\omega_X \otimes \rho_E$. The asymptotic continuity of entropic quantities allow us to discuss distinguishability of states in terms of entropies, as we shall see now.

Let us consider two states $\rho, \rho' \in \mathcal{D}(A)$ and let $\varepsilon \geq \frac{1}{2}\|\rho - \rho'\|_1$. A short argument outlined in [7] applied to the Fannes-Audenart inequality [19] upper bounds the difference in von Neumann entropies by

$$\left|\mathrm{H}(A)_\rho - \mathrm{H}(A)_{\rho'}\right| \leq \varepsilon \log d_A + h(\varepsilon).$$

Building upon this, we have for states $\rho, \rho' \in \mathcal{D}(AB)$ with $\varepsilon \geq \frac{1}{2}\|\rho - \rho'\|_1$ an upper bound on the difference of conditional entropies [20, 21] given by

$$\left|\mathrm{H}(A|B)_\rho - \mathrm{H}(A|B)_{\rho'}\right| \leq 2\varepsilon \log d_A + g(\varepsilon),$$

where $g(\varepsilon) = (1+\varepsilon)\, h\left(\frac{\varepsilon}{1+\varepsilon}\right)$. Combining the upper bounds on the von Neumann entropy and the conditional entropy above, we may obtain an upper bound on the difference between mutual information, namely,

$$\left|\mathrm{I}(A:B)_\rho - \mathrm{I}(A:B)_{\rho'}\right| \leq 3\varepsilon \log d_A + g(\varepsilon) + h(\varepsilon),$$

which notably is only dependent on the dimension of one among systems A, B. Due to the symmetry of systems A, B in the definition of the mutual information, we may choose the minimal dimension $d = \min(d_A, d_B)$.

Now, to see how entropic quantities can quantify the distinguishability of states in an ensemble, let $\varepsilon > 0$ and consider a cq-state $\rho \in \mathcal{D}(XE)$ with $\Delta(X|E)_\rho \leq \varepsilon$. Applying the upper bound on the difference above yields

$$\mathrm{I}(X:E)_\rho \leq 3\varepsilon \log d_X + g(\varepsilon) + h(\varepsilon),$$

which shows that the mutual information of a state approximating a product state is small.

The conditional min-entropy

As a next step towards unifying the previously introduced notions of state discrimination with entropic quantities, we note that for a cq-state $\rho \in \mathcal{D}(XE)$ the conditional min-entropy [22] is given by

$$\mathrm{H}_{\min}(X|E)_\rho \coloneqq -\log\left(\mathrm{Pr}_{\text{guess}}(X|E)_\rho\right).$$

Also, we define the conditional collision entropy [17] by

$$\mathrm{H}_{\mathrm{col}}\left(X|E\right)_{\rho} := -\log \operatorname{Tr}\left(\rho_{XE}\left(\mathbb{1}_X \otimes \rho_E^{-1/2}\right)\right)^2 = -\log\left(\sum_{x\in\Sigma} p_x \operatorname{Tr} M_x^{PG_\rho}\rho_x\right),$$

where we note that the collision entropy of a cq-state is actually given by the negative logarithm of the probability of correctly guessing the state of system X by applying the pretty good measurement. More generally, we have the following definition [23] of the conditional min-entropy of a generic state, which reduces to the definition above for cq-states [22].

Definition 5. Let $\rho_{AB} \in \mathcal{D}\left(AB\right)$. The *min-entropy of A conditioned on B* is given by

$$\mathrm{H}_{\min}\left(A|B\right)_{\rho} := \sup\left\{\lambda \in \mathbb{R} \,\middle|\, \rho_{AB} \leq 2^{-\lambda}\mathbb{1}_A \otimes \rho_B\right\}.$$

Furthermore, we define the *ε-smoothed conditional min-entropy* as

$$\mathrm{H}^{\varepsilon}_{\min}\left(A|B\right)_{\rho} := \sup_{\rho'\in\mathcal{B}^{\varepsilon}(\rho)} \mathrm{H}_{\min}\left(A|B\right)_{\rho'},$$

where we recall that $\mathcal{B}^{\varepsilon}\left(\rho\right)$ denotes the set of states $\rho' \in \mathcal{D}_{\leq}\left(AB\right)$ satisfying $P\left(\rho,\rho'\right) \leq \varepsilon$.

The conditional collision entropy is a special case of the notion of conditional Rényi-entropies, however, for the concerns of our work we will only consider the following notion of Rényi-like entropies [16], which are referred to as α-entropies.

Definition 6. Let $\rho_{AB} \in \mathcal{D}\left(AB\right)$ and $\alpha \geq 0$, $\alpha \neq 1$. The *conditional Rényi-like entropy of order α*, or simply the *conditional α-entropy*, is given by

$$\mathrm{H}_{\alpha}\left(A|B\right)_{\rho} := -\frac{1}{\alpha-1}\log \operatorname{Tr}\left(\rho_{AB}^{\alpha}\left(\mathbb{1}_A \otimes \rho_B\right)^{1-\alpha}\right)$$

The α-entropies are monotonically decreasing in α, additive on tensor products, and reduce to the von Neumann entropy in the limit $\alpha \to 0$. Furthermore, they play a central role in a proof of the quantum asymptotic equipartition property [16], which states that in the limit of a large number of repetitions of a state ρ, the ε-smoothed min-entropy approximates the conditional entropy of n copies of ρ. More precisely, we have

$$\lim_{\varepsilon\to 0}\lim_{n\to\infty}\frac{1}{n}\mathrm{H}^{\varepsilon}_{\min}\left(A^n|B^n\right)_{\rho^{\otimes n}} = \mathrm{H}\left(A|B\right)_{\rho}. \tag{1.3}$$

The statement in (1.3) is referred to as the quantum asymptotic equipartition property (QAEP).

1.5.2 Restricted Entropic Quantities

Consider a quantum system AB and let $\mathcal{C}\left(B\right)$ be a set of channels on system B. For a state $\rho \in \mathcal{D}\left(AB\right)$ we define the $\mathcal{C}\left(B\rangle B'\right)$-restricted conditional entropy of A given B as

$$\mathrm{H}_{\mathcal{C}(B\rangle B')}\left(A|B\right)_{\rho} := \inf_{\Lambda\in\mathcal{C}(B\rangle B')} \mathrm{H}\left(A|B'\right)_{\Lambda(\rho)},$$

and, more generally, the $\mathcal{C}\left(B\right)$-restricted conditional entropy of A given B is defined as

$$\mathrm{H}_{\mathcal{C}(B)}\left(A|B\right)_{\rho} \coloneqq \inf_{B'} \mathrm{H}_{\mathcal{C}(B\rangle B')}\left(A|B\right)_{\rho},$$

where we recall that $\mathcal{C}\left(B\rangle B'\right) = \mathcal{C}\left(B\right) \cap \mathcal{C}_{\mathrm{all}}\left(B\rangle B'\right)$. Due to the monotonicity of conditional entropy, we have $\mathrm{H}_{\mathcal{C}(B)}\left(A|B\right)_{\rho} \geq \mathrm{H}\left(A|B\right)_{\rho}$ for all sets of channels $\mathcal{C}\left(B\right)$ with equality whenever the identity channel is an element of $\mathcal{C}\left(B\right)$. Analogously, we define the $\mathcal{C}\left(B\rangle B'\right)$-restricted conditional α-entropy of A given B by

$$\mathrm{H}_{\alpha,\mathcal{C}(B\rangle B')}\left(A|B\right)_{\rho} \coloneqq \inf_{\Lambda \in \mathcal{C}(B\rangle B')} \mathrm{H}_{\alpha}\left(A\middle|B'\right)_{\Lambda(\rho)},$$

and, more generally, the $\mathcal{C}\left(B\right)$-restricted conditional α-entropy of A given B is defined as

$$\mathrm{H}_{\alpha,\mathcal{C}(B)}\left(A|B\right)_{\rho} \coloneqq \inf_{B'} \mathrm{H}_{\alpha,\mathcal{C}(B\rangle B')}\left(A|B\right)_{\rho}.$$

In an effort to build towards a generalization of the quantum asymptotic equipartition property stated in (1.3), we consider a sequence of sets of channels $\mathcal{C} = \left(\mathcal{C}_n\left(B^n\right)\right)_{n\in\mathbb{N}}$, and denote by

$$\begin{aligned}
\mathrm{H}^{\infty}_{\mathcal{C}}\left(A|B\right)_{\rho} &\coloneqq \inf_{n\in\mathbb{N}} \frac{1}{n} \mathrm{H}_{\mathcal{C}_n(B^n)}\left(A^n|B^n\right)_{\rho^{\otimes n}},\\
\mathrm{H}^{\infty}_{\alpha}\left(A|B\right)_{\rho} &\coloneqq \inf_{n\in\mathbb{N}} \frac{1}{n} \mathrm{H}_{\alpha,\mathcal{C}_n(B^n)}\left(A^n|B^n\right)_{\rho^{\otimes n}}.
\end{aligned}$$

To justify this choice of notation we now prove the following result.

Lemma 7. Consider a state $\rho \in \mathcal{D}\left(AB\right)$. Suppose $\mathcal{C} = \left(\mathcal{C}_n\left(B^n\right)\right)_{n\in\mathbb{N}}$ is a sequence of sets of quantum channels satisfying

$$\mathcal{C}_m\left(B^m\right) \otimes \mathcal{C}_n\left(B^n\right) \subseteq \mathcal{C}_{m+n}\left(B^{m+n}\right).$$

Then

$$\begin{aligned}
\mathrm{H}^{\infty}_{\mathcal{C}}\left(A|B\right)_{\rho} &= \lim_{n\to\infty} \frac{1}{n} \mathrm{H}_{\mathcal{C}_n(B^n)}\left(A^n|B^n\right)_{\rho^{\otimes n}}\\
\mathrm{H}^{\infty}_{\alpha,\mathcal{C}}\left(A|B\right)_{\rho} &= \lim_{n\to\infty} \frac{1}{n} \mathrm{H}_{\alpha,\mathcal{C}_n(B^n)}\left(A^n|B^n\right)_{\rho^{\otimes n}}
\end{aligned}$$

Proof. Let $m, n \in \mathbb{N}$ and note

$$\begin{aligned}
\overbrace{\mathrm{H}_{\mathcal{C}_{m+n}(B^{m+n})}\left(A^{m+n}\middle|B^{m+n}\right)_{\rho^{\otimes(m+n)}}}^{s_{m+n}=} &\leq \mathrm{H}_{\mathcal{C}_m(B^m)\otimes\mathcal{C}_n(B^n)}\left(A^{m+n}\middle|B^{m+n}\right)_{\rho^{\otimes(m+n)}}\\
&= \underbrace{\mathrm{H}_{\mathcal{C}_m(B^m)}\left(A^m|B^m\right)_{\rho^{\otimes m}}}_{s_m=} + \underbrace{\mathrm{H}_{\mathcal{C}_n(B^n)}\left(A^n|B^n\right)_{\rho^{\otimes n}}}_{s_n=},
\end{aligned}$$

where we have used the additivity of the von Neumann entropy. This shows that the sequence $\left(s_n\right)_{n\in\mathbb{N}}$ is subadditive, and it follows from Fekete's subadditivity lemma [24] that the sequence $\frac{1}{n}s_n$ converges with the desired limit. As the conditional α-entropy is additive as well, the second statement follows from a similar argument. □

As a special case of interest, we note that the sequence of sets of all measurements satisfies the condition of Lemma 7. As this will be an example of interest, we introduce the slightly less cumbersome notation

$$\mathrm{H}_B^\infty (A|B)_\rho \coloneqq \mathrm{H}_{\mathcal{M}_{\text{all}}^B}^\infty (A|B)_\rho ,$$

where $\mathcal{M}_{\text{all}}^B$ denotes the sequence of sets of all measurements $(\mathcal{M}_{\text{all}}(B^n))_{n\in\mathbb{N}}$.

Furthermore, if we consider a tripartite system XAB, we may note that all sets of local measurements on system AB satisfy the condition of Lemma 7. Again to reduce the complexity of the notation, we will for instance denote by

$$\mathrm{H}_A^\infty (X|AB)_\rho \coloneqq \mathrm{H}_{\mathcal{M}_{\text{all}}^A}^\infty (X|AB)_\rho , \qquad \mathrm{H}_{A\to B}^\infty (X|AB)_\rho \coloneqq \mathrm{H}_{\mathcal{M}^{A\to B}}^\infty (X|AB)_\rho ,$$

where we denote by $\mathcal{M}^{A\to B}$ the sequence of sets $(\mathcal{M}_{A\to B}(A^n : B^n))_{n\in\mathbb{N}}$. We define the restricted conditional entropies with respect to the remaining local measurements analogously.

The restricted conditional min-entropy

Consider a bipartite quantum system AB and a set of channels $\mathcal{C}(B)$. Analogously to previous definitions, we denote the $\mathcal{C}(B\rangle B')$-restricted conditional min-entropy of $\rho \in \mathcal{D}(AB)$ by

$$\mathrm{H}_{\min,\mathcal{C}(B\rangle B')} (A|B)_\rho \coloneqq \inf_{\Lambda\in\mathcal{C}(B\rangle B')} \mathrm{H}_{\min} \left(A|B'\right)_{\Lambda(\rho)} ,$$

and so the $\mathcal{C}(B)$-restricted min-entropy is given by

$$\mathrm{H}_{\min,\mathcal{C}(B)} (A|B)_\rho \coloneqq \inf_{B'} \mathrm{H}_{\min,\mathcal{C}(B\rangle B')} (A|B)_\rho .$$

For $\varepsilon \geq 0$, we define the ε-smoothed $\mathcal{C}(B)$-restricted min-entropy completely analogously, that is,

$$\mathrm{H}^\varepsilon_{\min,\mathcal{C}(B\rangle B')} (A|B)_\rho \coloneqq \inf_{\Lambda\in\mathcal{C}(B\rangle B')} \mathrm{H}^\varepsilon_{\min} \left(A|B'\right)_{\Lambda(\rho)} ,$$

with the ε-smoothed $\mathcal{C}(B)$-restricted min-entropy given by

$$\mathrm{H}^\varepsilon_{\min,\mathcal{C}(B)} (A|B)_\rho \coloneqq \inf_{B'} \mathrm{H}^\varepsilon_{\min,\mathcal{C}(B\rangle B')} (A|B)_\rho .$$

Let us now restrict our attention to cq-states $\rho \in \mathcal{D}(XE)$ given by

$$\rho = \sum_{x\in\Sigma} p_x \, |x\rangle\langle x| \otimes \rho_x, \qquad \rho_x \in \mathcal{D}(E) ,$$

where $\{|x\rangle\}_{x\in\Sigma}$ is an orthonormal basis of X. As before, we may introduce a regularized version of the restricted min-entropy, so consider a sequence of sets of channels $\mathcal{C} = (\mathcal{C}_n(E^n))_{n\in\mathbb{N}}$, and denote by

$$\mathrm{H}^\infty_{\min,\mathcal{C}} (X|E)_\rho \coloneqq \inf_{n\in\mathbb{N}} \frac{1}{n} \mathrm{H}_{\min,\mathcal{C}_n(E^n)} (X^n|E^n)_{\rho^{\otimes n}} .$$

As before we justify this notation by the following result.

Lemma 8. Consider a cq-state $\rho \in \mathcal{D}(XE)$. Suppose $\mathcal{C} = \left(\mathcal{C}_n\left(E^n\right)\right)_{n\in\mathbb{N}}$ is a sequence of sets of quantum channels satisfying

$$\mathcal{C}_m\left(E^m\right) \otimes \mathcal{C}_n\left(E^n\right) \subseteq \mathcal{C}_{m+n}\left(E^{m+n}\right).$$

Then

$$\mathrm{H}^{\infty}_{\min,\mathcal{C}}\left(X|E\right)_\rho = \lim_{n\to\infty} \frac{1}{n} \mathrm{H}_{\min,\mathcal{C}_n(E^n)}\left(X^n|E^n\right)_{\rho^{\otimes n}}.$$

Proof. Let $m, n \in \mathbb{N}$ and note

$$\begin{aligned}\overbrace{\mathrm{H}_{\min,\mathcal{C}_{m+n}(E^{m+n})}\left(X^{m+n}\middle|E^{m+n}\right)_{\rho^{\otimes(m+n)}}}^{s_{m+n}=} &\leq \mathrm{H}_{\min,\mathcal{C}_m(E^m)\otimes\mathcal{C}_n(E^n)}\left(X^{m+n}\middle|E^{m+n}\right)_{\rho^{\otimes(m+n)}} \\ &\leq \underbrace{\mathrm{H}_{\min,\mathcal{C}_m(E^m)}\left(X^m|E^m\right)_{\rho^{\otimes m}}}_{s_m=} + \underbrace{\mathrm{H}_{\mathcal{C}_n(E^n)}\left(X^n|E^n\right)_{\rho^{\otimes n}}}_{s_n=},\end{aligned}$$

where we have used the operational interpretation of the conditional min-entropy to infer subadditivity on cq-states. This shows that the sequence $\left(s_n\right)_{n\in\mathbb{N}}$ is subadditive, and it follows from Fekete's subadditivity lemma that the sequence $\frac{1}{n}s_n$ converges with the desired limit. □

As we aim towards a generalization of the QAEP, it is natural to consider the regularization of the ε-smoothed restricted conditional min-entropy, which we define for a sequence of sets of channels $\mathcal{C} = \left(\mathcal{C}_n\left(B^n\right)\right)_{n\in\mathbb{N}}$ by

$$\mathrm{H}^{\varepsilon,\infty}_{\min,\mathcal{C}}\left(A|B\right)_\rho \coloneqq \liminf_{n\to\infty} \frac{1}{n} \mathrm{H}^{\varepsilon}_{\min,\mathcal{C}_n(B^n)}\left(A^n|B^n\right)_{\rho^{\otimes n}}.$$

We have included the limit infimum outright in the definition as we do not have convergence results analogous to Lemma 8 for the ε-smoothed restricted conditional min-entropy. Furthermore, we denote by

$$\widetilde{\mathrm{H}}^{\infty}_{\min,\mathcal{C}}\left(A|B\right)_\rho \coloneqq \lim_{\varepsilon\to 0} \mathrm{H}^{\varepsilon,\infty}_{\min,\mathcal{C}}\left(A|B\right)_\rho,$$

and with this in place, we have the following statement inspired by the QAEP and the proof thereof.

Lemma 9. Let $\rho \in \mathcal{D}(AB)$, and suppose $\mathcal{C} = \left(\mathcal{C}_n\left(B^n\right)\right)_{n\in\mathbb{N}}$ is a sequence of sets of quantum channels. Then

$$\mathrm{H}^{\infty}_{\alpha,\mathcal{C}}\left(A|B\right)_\rho \leq \widetilde{\mathrm{H}}^{\infty}_{\min,\mathcal{C}}\left(A|B\right)_\rho \leq \mathrm{H}^{\infty}_{\mathcal{C}}\left(A|B\right)_\rho$$

for all $\alpha \in (1, 2]$.

Proof. Let $\varepsilon > 0$, $n \in \mathbb{N}$ and $\Lambda \in \mathcal{C}\left(B^n\rangle B_n\right)$. Let $\xi \in \mathcal{B}^{\varepsilon}\left(\Lambda\left(\rho^{\otimes n}\right)\right) \subseteq \mathcal{D}_{\leq}\left(A^n B_n\right)$ and suppose

$$\mathrm{H}^{\varepsilon}_{\min}\left(A^n|B_n\right)_{\Lambda(\rho^{\otimes n})} = \mathrm{H}_{\min}\left(A^n|B_n\right)_\xi.$$

Due to the refined Alicki-Fannes inequality [20, 21], it follows that

$$\begin{aligned}\mathrm{H}^{\varepsilon}_{\min}\left(A^n|B_n\right)_{\Lambda(\rho^{\otimes n})} &= \mathrm{H}_{\min}\left(A^n|B_n\right)_\xi \\ &\leq \mathrm{H}\left(A^n|B_n\right)_\xi \\ &\leq \mathrm{H}\left(A^n|B_n\right)_{\Lambda(\rho^{\otimes n})} + 2\varepsilon n \log d_A + g\left(\varepsilon\right).\end{aligned}$$

Dividing both sides of the inequality by n and taking infimum of both sides over all $\Lambda \in \mathcal{C}_n(B^n)$ yields

$$\frac{1}{n}\,\mathrm{H}^{\varepsilon}_{\min,\mathcal{C}_n(B^n)}\left(A^n|B^n\right)_{\rho^{\otimes n}} \leq \frac{1}{n}\,\mathrm{H}_{\mathcal{C}_n(B^n)}\left(A^n|B^n\right)_{\rho^{\otimes n}} + 2\varepsilon\log d_A + \frac{1}{n}g\left(\varepsilon\right),$$

which allows us to infer $\widetilde{\mathrm{H}}^{\infty}_{\min,\mathcal{C}}\left(A|B\right)_{\rho} \leq \mathrm{H}^{\infty}_{\mathcal{C}}\left(A|B\right)_{\rho}$.

For the converse inequality, note that for all $\alpha \in (1,2]$ we have

$$\mathrm{H}^{\varepsilon}_{\min}\left(A^n|B_n\right)_{\Lambda(\rho^{\otimes n})} \geq \mathrm{H}_{\alpha}\left(A^n|B_n\right)_{\Lambda(\rho^{\otimes n})} - \frac{1}{\alpha-1}\log\frac{2}{\varepsilon^2},$$

which is shown in [16] (see Theorem 81 in Appendix A). Dividing by n and taking infimum on both sides over all $\Lambda \in \mathcal{C}_n(B^n)$ yields

$$\frac{1}{n}\,\mathrm{H}^{\varepsilon}_{\min,\mathcal{C}_n(B^n)}\left(A^n|B^n\right)_{\rho^{\otimes n}} \geq \frac{1}{n}\,\mathrm{H}_{\alpha,\mathcal{C}_n(B^n)}\left(A^n|B^n\right)_{\rho^{\otimes n}} - \frac{1}{n\left(\alpha-1\right)}\log\frac{2}{\varepsilon^2},$$

and so letting $n \to \infty$ we may infer that $\mathrm{H}^{\infty}_{\alpha,\mathcal{C}}\left(A|B\right)_{\rho} \leq \widetilde{\mathrm{H}}^{\infty}_{\min,\mathcal{C}}\left(A|B\right)_{\rho}$ for all $\alpha \in (1,2]$. □

We note that for all states $\rho \in \mathcal{D}(AB)$ we have

$$\mathrm{H}^{\infty}_{\min,\mathcal{C}}\left(A|B\right)_{\rho} \leq \widetilde{\mathrm{H}}^{\infty}_{\min,\mathcal{C}}\left(A|B\right)_{\rho},$$

which serves as a useful lower bound when computing lower bounds in concrete examples as we shall see later. Finally, we conjecture that the upper bound in Lemma 9 is also a lower bound; this gives the following statement.

Conjecture 10. Let $\rho \in \mathcal{D}(AB)$, and suppose $\mathcal{C}(B) = \left(\mathcal{C}_n(B^n)\right)_{n\in\mathbb{N}}$ is a sequence of sets of quantum channels satisfying

$$\mathcal{C}_m(B^m) \otimes \mathcal{C}_n(B^n) \subseteq \mathcal{C}_{m+n}\left(B^{m+n}\right).$$

Then

$$\widetilde{\mathrm{H}}^{\infty}_{\min,\mathcal{C}}\left(A|B\right)_{\rho} = \mathrm{H}^{\infty}_{\mathcal{C}}\left(A|B\right)_{\rho}.$$

The restricted conditional min-entropy of a cq-state

Let us now restrict our attention to cq-state states. We begin by showing an elementary inequality concerning the ε-smoothed restricted conditional min-entropy. If we consider a cq-state $\rho \in \mathcal{D}(XE)$, it has been shown [17] that for any $\varepsilon \geq 0$ there exists a cq-state $\xi \in \mathcal{B}^{\varepsilon}(\rho)$ satisfying

$$\mathrm{H}^{\varepsilon}_{\min}\left(X|E\right)_{\rho} = \mathrm{H}_{\min}\left(X|E\right)_{\xi}.$$

We will use this observation repeatedly, in fact already in the following result.

Lemma 11. Let $\rho \in \mathcal{D}(XE)$ be a cq-state and consider

$$\rho^{XET} = \sum_{x\in\Sigma} p_x \left|x\right\rangle\!\left\langle x\right|^X \otimes \rho^E_x \otimes \left|s(x)\right\rangle\!\left\langle s(x)\right|^T,$$

where $s\colon \Sigma \to \mathcal{T}$ is a function with $\mathcal{T}$ a finite set, and the quantum system T has orthonormal basis $\{|t\rangle\}_{t\in\mathcal{T}}$. For any set of channels $\mathcal{C}(E)$ we have

$$\mathrm{H}^{\varepsilon}_{\min,\mathcal{C}(E)}\left(X|ET\right)_{\rho} \geq \mathrm{H}^{\varepsilon}_{\min,\mathcal{C}(E)}\left(X|E\right)_{\rho} - \log|\mathcal{T}|.$$

Proof. Let $\varepsilon > 0$ and $\Lambda \in \mathcal{C}(E\rangle E')$. Let $\xi^{XA'} \in \mathcal{B}^{\varepsilon}(\Lambda(\rho)) \subseteq \mathcal{D}_{\leq}(XE')$ be a cq-state satisfying

$$\mathrm{H}^{\varepsilon}_{\min}\left(X\middle|E'\right)_{\Lambda(\rho)} = \mathrm{H}_{\min}\left(X\middle|E'\right)_{\xi}.$$

Let $V \in \mathrm{L}(X, XT)$ denote the isometric operator given by $V = \sum_{x\in\Sigma} |x\rangle\, |s(x)\rangle\, \langle x|$, and note that $\rho^{XET} = V\rho^{XE}V^{\dagger}$. As $\Lambda \in \mathcal{C}(E)$ and V acts on distinct systems, it follows that $\Lambda\left(\rho^{XET}\right) = V\Lambda\left(\rho^{XE}\right)V^{\dagger}$. Finally, we may denote by $\xi^{XE'T} = V\xi^{XE'}V^{\dagger}$, and note that $\xi^{XE'T}$ is an extension of $\xi^{XE'}$, since $\xi^{XE'}$ is a cq-state. It follows that

$$\begin{aligned}\mathrm{H}^{\varepsilon}_{\min}\left(X\middle|E'\right)_{\Lambda(\rho)} - \log|\mathcal{T}| = \mathrm{H}_{\min}\left(X\middle|E'\right)_{\xi} - \log|\mathcal{T}| &\leq \mathrm{H}_{\min}\left(X\middle|E'T\right)_{\xi}\\ &\leq \mathrm{H}^{\varepsilon}_{\min}\left(X\middle|E'T\right)_{\Lambda(\rho)}\end{aligned}$$

where we have used the chain rule of the min-entropy and isometric invariance of the purified distance, respectively, to prove the last two inequalities. The desired statement thus follows from taking infimum on both sides of the inequality over all channels in $\mathcal{C}(E)$. □

Before we proceed to consider a generalized version of the leftover hash lemma [17], we give the following definition. Consider two finite sets Σ, Σ', and let $\mathcal{F}$ be a set of functions $f\colon \Sigma \to \Sigma'$ with an associated probability distribution $(p_f)_{f\in\mathcal{F}}$. If the probability that two distinct elements are mapped to the same value by f is less than $\frac{1}{|\Sigma'|}$, that is,

$$\sum_{f\in\mathcal{F}} p_f \delta_{f(x),f(y)} \leq \frac{1}{|\Sigma'|}$$

for all distinct $x, y \in \Sigma$, then we say $\mathcal{F}$ is a two-universal family of functions. The most natural example of a two-universal family of functions is the family of all functions equipped with a uniform distribution.

For a cq-state $\rho \in \mathcal{D}(XE)$ given by

$$\rho^{XE} = \sum_{x\in\Sigma} p_x\, |x\rangle\langle x|^{X} \otimes \rho_x^{E}, \qquad \rho_x \in \mathcal{D}(E),$$

and a family $\mathcal{F}$ of functions $f\colon \Sigma \to \Sigma'$ with corresponding probability distribution $(p_f)_{f\in\mathcal{F}}$ we denote by

$$f(\rho) = \sum_{x\in\Sigma} p_x\, |f(x)\rangle\langle f(x)|^{X'} \otimes \rho_x = \sum_{x'\in\Sigma'} |x'\rangle\langle x'|^{X'} \otimes \sum_{x\in f^{-1}(\{x'\})} p_x\rho_x \in \mathcal{D}\left(X'E\right),$$

and furthermore

$$\mathcal{F}(\rho) \coloneqq \sum_{f\in\mathcal{F}} p_f f(\rho) \otimes |f\rangle\langle f|^{F} \in \mathcal{D}\left(X'EF\right),$$

where $\{|f\rangle\}_{f\in\mathcal{F}}$ is an orthonormal basis of system F. With all of the above in place, we are now in a position to state a restricted version of the leftover hash lemma [17].

Lemma 12. Consider a cq-state $\rho^{XE} \in \mathcal{D}(XE)$, and let $\mathcal{C}(E)$ be a set of channels. Let Σ' be a finite set and let $\mathcal{F}$ be a two-universal family of functions $f\colon \Sigma \to \Sigma'$. For all $\varepsilon \geq 0$ we have

$$\Delta_{\mathcal{C}(E)}\left(X'\middle|EF\right)_{\mathcal{F}(\rho)} \leq 2\varepsilon + \sqrt{2^{\log|\mathcal{X}'| - \mathrm{H}^{\varepsilon}_{\min,\mathcal{C}(E)}(X|E)_{\rho}}}.$$

Proof. Let $\Lambda \in \mathcal{C}\left(E\rangle E'\right)$. It follows from the leftover hash lemma [17] (see Lemma 83 in Appendix) that

$$\Delta\left(X'\middle|EF\right)_{\mathcal{F}(\Lambda(\rho))} \leq 2\varepsilon + \sqrt{2^{\log|\mathcal{X}'|-\mathrm{H}^{\varepsilon}_{\min}(X|E')_{\Lambda(\rho)}}},$$

and so taking supremum over all $\Lambda \in \mathcal{C}\left(E\right)$ yields the desired result. □

In the statement of Lemma 12 the choice of channel $\Lambda \in \mathcal{C}\left(E\right)$ is independent of the function $f \in \mathcal{F}$ applied to the state of register X. We pose the question of whether this implicit assumption is necessary, which is presented as a conjecture below.

Conjecture 13. Consider a cq-state $\rho^{XE} \in \mathcal{D}\left(XE\right)$, and let $\mathcal{C}\left(E\right)$ be a set of channels. Let Σ' be a finite set and let $\mathcal{F}$ be a two-universal family of functions $f\colon \Sigma \to \Sigma'$ with corresponding probability distribution $\left(p_f\right)_{f\in\mathcal{F}}$. For all $\varepsilon \geq 0$ we have

$$\sum_{f\in\mathcal{F}} p_f \Delta_{\mathcal{C}(E)}\left(X'\middle|E\right)_{f(\rho)} \leq 2\varepsilon + \sqrt{2^{\log|\mathcal{X}'|-\mathrm{H}^{\varepsilon}_{\min,\mathcal{C}(E)}(X|E)_{\rho}}}.$$

Chapter 2

Secure States and Hiding States as a Resource

In this chapter, we will consider the fundamental problem of extracting secure data from only partially secure data, where an eavesdropper may have information encoded into the state of a quantum system. More precisely, we consider a two-party scenario involving Xavier and Eve, which is described by them sharing a cq-state $\rho \in \mathcal{D}(XE)$ given by

$$\rho^{XE} = \sum_{x \in \Sigma} p_x \, |x\rangle\langle x|^X \otimes \rho_x^E, \qquad \rho_x \in \mathcal{D}(E) \tag{2.1}$$

where Σ is an alphabet and $(p_x)_{x \in \Sigma}$ is a probability distribution on Σ. Note that Eve may try to guess the state of X by, say, measuring her system; if the states ρ_x for $x \in \Sigma$ are orthogonal, she may even infer the state of system X with certainty! In general, however, Eve is only able to infer partial information about the state of system X, and in this chapter we consider to what extent Xavier is able to obtain a cq-state, where access to system E only yields negligible information on the state of system X. In particular, we will discuss the scenario where Eve's quantum memory is imperfect, and to what extent Xavier is able to exploit this fact in order to obtain a larger amount of secure data.

We will quantify the amount of secure data in a cq-state given by (2.1) within a resource theoretic framework. To describe the set of free operations, we first allow Xavier to choose an alphabet Σ' and a function $f \colon \Sigma \to \Sigma'$, and hence apply it to the state of system X. This yields the state

$$f(\rho) \coloneqq \sum_{x \in \Sigma} |f(x)\rangle\langle f(x)|^{X'} \otimes p_x \rho_x^E \in \mathcal{D}\left(X'E\right), \tag{2.2}$$

where X' is spanned by an orthonormal basis $\{|x'\rangle\}_{x' \in \Sigma'}$. Furthermore, we allow Xavier to act probabilistically, that is, he may sample f from a family of functions $\mathcal{F}$ with an associated probability distribution $(p_f)_{f \in \mathcal{F}}$, as long as the choice of $f \in \mathcal{F}$ is publicly announced to Eve. The resulting state is given by

$$\mathcal{F}(\rho) = \sum_{f \in \mathcal{F}} p_f f(\rho) \otimes |f\rangle\langle f| \in \mathcal{D}\left(X'EF\right), \tag{2.3}$$

where F is spanned by an orthonormal basis $\{|f\rangle\}_{f \in \mathcal{F}}$. As mentioned above, the desired cq-state is one where access to systems EF yields only negligible information on the state of system X', which we will define rigorously below.

In the first section, we discuss the definition of a secure state, namely, one where access to Eve's quantum side information only yields negligible information about Xavier's classical data. Next, we rephrase the work on strong randomness extraction [25] in this terminology, and we proceed to show that the rate at which secure bits can be distilled from a generic cq-state is exactly the conditional von Neumann entropy $\mathrm{H}\left(X|E\right)_{\rho}$.

In Section 2.2 we first consider the task of distilling secure bits with respect to an eavesdropper Eve with imperfect quantum memory. We model this scenario by assuming Eve has to perform some operation from a set of quantum channels $\mathcal{C}\left(E\right)$ on her system prior to knowing $f \in \mathcal{F}$. Furthermore, we provide bounds on the rate at which secure bits can be distilled from a generic cq-state with respect to an eavesdropper with imperfect memory. Additionally, we briefly discuss the scenario, where Eve's choice of quantum channel $\Lambda \in \mathcal{C}\left(E\right)$ is allowed to depend on Xavier's choice of $f \in \mathcal{F}$. This corresponds to the scenario of Eve having perfect quantum memory, but she is restricted to some set of operations $\mathcal{C}\left(E\right)$; this allows her to retain the state of system E until Xavier has chosen $f \in \mathcal{F}$, and so her choice of operation may depend on Xavier's choice of function.

Finally, in the last section, we add to the communication setup that Xavier may pass some information on the state of system X to Eve. Then we pose the task of hiding data from an eavesdropper with imperfect memory, that is, achieving a cq-state where an eavesdropper with perfect quantum memory would be able to infer the state of system X' with high probability, while an eavesdropper with imperfect quantum memory only has negligible information about the state of X. We will refer to such a state as a hiding state.

2.1 Secure States with respect to an Eavesdropper

In this section, we are concerned with a two-party setting involving Xavier and Eve. Let Σ be an alphabet and let $\left(p_x\right)_{x\in\Sigma}$ be a probability distribution. Let X be a quantum system with orthonormal basis $\left\{|x\rangle\right\}_{x\in\Sigma}$, and suppose Xavier has encoded $x \in \Sigma$ into the state of system X. Meanwhile, Eve has received some information about the state of Xavier's system X, which is encoded in the state of an additional quantum system E. The scenario is described by the cq-state given in (2.1). In the following, we discuss how to define the security of the data encoded into the state of system X when an eavesdropper is given a quantum system E with information encoded into the state of E.

We begin this section by considering two natural approaches to quantifying security. First, we will consider an operational approach, where we quantify the security of the data encoded in system X in terms of an eavesdropper Eve's ability to guess the state of system X given the outcome of a measurement on system E. Below, we describe security in terms of Eve's bias when trying to infer the state of system X from such a measurement.

Definition 14. Let Σ be an alphabet and consider a cq-state $\rho^{XE} \in \mathcal{D}\left(XE\right)$ given by

$$\rho^{XE} = \sum_{x\in\Sigma} p_x \left|x\right\rangle\!\left\langle x\right|^{X} \otimes \rho_x^{E}, \qquad \rho_x \in \mathcal{D}\left(E\right),$$

For $\varepsilon \geq 0$, we say that ρ is an $(\log|\Sigma|, \varepsilon)$-*secure state with respect to* E, if

$$\Pr_{\text{guess}}(X|E)_\rho - \frac{1}{|\Sigma|} < \varepsilon.$$

If $\Pr_{\text{guess}}(X|E)_\rho = \frac{1}{|\Sigma|}$, we say that ρ is a *secure state with respect to* E.

Remark. When it is clear from the context, we simply refer to ρ as an $(\log|\Sigma|, \varepsilon)$-secure state without mentioning system E.

It is easy to see that a cq-state $\rho \in \mathcal{D}(XE)$ is a secure state, if and only if

$$\rho^{XE} = \sum_{x\in\Sigma} \frac{1}{|\Sigma|} |x\rangle\langle x|^X \otimes \rho^E.$$

Our second approach to defining the security of data encoded into the state of system X is given in terms of proximity to a secure state. We use the trace norm as a measure of distance because of its natural interpretation in terms of the distinguishability of states due to the Holevo-Helstrom Theorem [15, 26].

Definition 15. Let Σ be an alphabet and consider a cq-state $\rho^{XE} \in \mathcal{D}(XE)$ given by

$$\rho^{XE} = \sum_{x\in\Sigma} p_x |x\rangle\langle x|^X \otimes \rho_x^E.$$

For $\varepsilon > 0$, we say that ρ is a $(\log|\Sigma|, \varepsilon)$-*approximate secure state with respect to* E, if

$$\Delta(X|E)_\rho = \frac{1}{2}\|\rho_{XE} - \omega_X \otimes \rho_E\|_1 < \varepsilon.$$

Remark. Again, when it is clear from the context, we simply refer to ρ as an $(\log|\Sigma|, \varepsilon)$-approximate secure state without mentioning system E.

With two notions of secure data with respect to an eavesdropper, it is natural to pose the question of their (in)equivalence. It is a direct consequence of Proposition 4 that defining the security of data with respect to an eavesdropper in terms of proximity to a secure state is at least as restrictive as the definition arising from an operational approach. We will thus focus our attention on this more restrictive notion of security.

In the following, we consider an operationally relevant task of Xavier, namely, to obtain secure states with respect to an eavesdropper Eve when given a cq-state $\rho^{XE} \in \mathcal{D}(XE)$ as a resource. We allow Xavier access to public randomness and he is allowed to perform classical processing of the data encoded in system X. More precisely, for any alphabet Σ' Xavier may sample a function $f\colon \Sigma \to \Sigma'$ from a family of functions $\mathcal{F}$ with associated probability distribution $(p_f)_{f\in\mathcal{F}}$, and apply it to his data $x \in \Sigma$ with the choice of $f \in \mathcal{F}$ being available to the eavesdropper. The resulting cq-state is denoted by $\mathcal{F}(\rho)$, which is defined in (2.3). In general, Xavier is not able to obtain exact secure states using this protocol, so we will consider the asymptotic setting, where he has $n \in \mathbb{N}$ copies of a state ρ, and aims towards obtaining (m, ε)-secure states for $m \in \mathbb{N}$ at fixed rates $\frac{m}{n}$ for any $\varepsilon > 0$.

Definition 16. Let $\rho \in \mathcal{D}\left(XE\right)$ be a cq-state given by

$$\rho = \sum_{x \in \Sigma} p_x \left|x\right\rangle\!\left\langle x\right|_X \otimes \rho_E, \qquad \rho_x \in \mathcal{D}\left(E\right),$$

and let $r \geq 0$. We say that r is an *achievable rate of secure state distillation*, if $r = 0$ or the following condition holds: For sufficiently large $n \in \mathbb{N}$ and $m = \lfloor rn \rfloor$ there exists a family $\mathcal{F}$ of functions $f \colon \Sigma^n \to \Sigma'$, where Σ' is of size $\left|\Sigma'\right| = 2^m$, with associated probability distribution $\left(p_f\right)_{f \in \mathcal{F}}$, such that

$$\mathcal{F}\left(\rho^{\otimes n}\right) = \sum_{x' \in \Sigma'} \left|x'\right\rangle\!\left\langle x'\right|^{X'} \otimes \sum_{f \in \mathcal{F}} p_f \sum_{x \in f^{-1}(\{x'\})} p_x \rho_x^{E^n} \otimes \left|f\right\rangle\!\left\langle f\right|^F,$$

is an (m, ε)-approximate secure state with respect to $E^n F$. Here, we denote by $\rho_x = \otimes_{i=1}^n \rho_{x_i}$ for $x = (x_1, \ldots, x_n) \in \Sigma^n$.

The *rate of secure state distillation* $S_D\left(\rho\right)$ is the supremum over all achievable rates of secure state distillation.

Proposition 17. Let Σ be an alphabet and consider a cq-state $\rho^{XE} \in \mathcal{D}\left(XE\right)$. Then

$$S_D\left(\rho\right) = \mathrm{H}\left(X|E\right)_\rho .$$

Remark. The result corresponds to the single-shot version stated without proof in [22].

Proof. Let $\varepsilon > 0$. Let $n \in \mathbb{N}$, $\delta > 0$, and consider $m \in \mathbb{N}_0$ given by $m = \left\lfloor n \left(\mathrm{H}\left(X|E\right)_\rho - \delta \right) \right\rfloor$. Let $\mathcal{F}$ be a two-universal family of functions $f \colon \Sigma^n \to \{0,1\}^m$, and note that it follows from the leftover hash lemma [17] (see Lemma 83 in Appendix A) that

$$\begin{aligned}
\Delta\left(X' \middle| E^n F\right)_{\mathcal{F}(\rho^{\otimes n})} &\leq 2\varepsilon + \sqrt{2^{m - \mathrm{H}_{\min}^{\varepsilon}(X^n|E^n)_{\rho^{\otimes n}}}} \\
&= 2\varepsilon + \sqrt{2^{n\left(\frac{m}{n} - \frac{1}{n}\mathrm{H}_{\min}^{\varepsilon}(X^n|E^n)_{\rho^{\otimes n}}\right)}}. \\
&\leq 2\varepsilon + \sqrt{2^{n\left(\mathrm{H}(X|E)_\rho - \frac{1}{n}\mathrm{H}_{\min}^{\varepsilon}(X^n|E^n)_{\rho^{\otimes n}} - \delta\right)}}, \qquad (2.4)
\end{aligned}$$

where the last inequality is due to our choice of $m \in \mathbb{N}_0$. Furthermore, it is a direct consequence of a quantum asymptotic equipartition property [16] (see Theorem 82 in Appendix A) that

$$\frac{1}{n}\mathrm{H}_{\min}^{\varepsilon}\left(X^n|E^n\right)_{\rho^{\otimes n}} > \mathrm{H}\left(X|E\right)_\rho - \delta$$

for sufficiently large $n \in \mathbb{N}$, which in turn implies that the second term in (2.4) tends to 0 as n tends to infinity. This proves $\mathrm{H}\left(X|E\right)_\rho - \delta$ is an achievable rate of secure state distillation for all $\delta > 0$, and this implies the achievability of the desired result.

For the converse inequality, let $\varepsilon > 0$, $m, n \in \mathbb{N}$ and suppose there exists a family $\mathcal{F}$ of functions $f \colon \Sigma^n \to \Sigma'$, where Σ' is of size $\left|\Sigma'\right| = 2^m$, with associated probability distribution $\left(p_f\right)_{f \in \mathcal{F}}$, such that

$$\Delta\left(X' \middle| EF\right)_{\mathcal{F}(\rho^{\otimes n})} < \varepsilon.$$

It follows that

$$\Delta\left(X'|E^nF\right)_{\mathcal{F}(\rho^{\otimes n})} = \sum_{f\in\mathcal{F}} p_f \Delta\left(X'|E^n\right)_{f(\rho^{\otimes n})} < \varepsilon,$$

which implies the existence of $f \in \mathcal{F}$ satisfying $\Delta\left(X'|E^n\right)_{f(\rho^{\otimes n})} < \varepsilon$. Finally, this yields

$$\begin{aligned} m = \mathrm{H}\left(X'|E^n\right)_{\omega^{X'}\otimes\rho^{\otimes n}} &\leq \mathrm{H}\left(X'|E^n\right)_{f(\rho^{\otimes n})} + 2\varepsilon\log\left|\Sigma'\right| + g\left(\varepsilon\right) \\ &\leq n\,\mathrm{H}\left(X|E\right)_\rho + \varepsilon n\log|\Sigma| + g\left(\varepsilon\right) \end{aligned}$$

where the first inequality is a consequence of Winter's refinement [21] of the Fannes-Audenart inequality [20] (see Theorem 85 and Theorem 86 in Appendix A), and the second inequality follows from a data processing inequality of the conditional von Neumann entropy discussed in Subsection 1.5.1, and the fact that we may assume without loss of generality $|\Sigma'| \leq |\Sigma|^n$. If we divide by $n \in \mathbb{N}$ at both sides of the inequality, we obtain

$$\frac{m}{n} \leq \mathrm{H}\left(X|E\right)_\rho + 2\varepsilon\log|\Sigma| + \frac{1}{n}g\left(\varepsilon\right),$$

and as this holds for arbitrary $\varepsilon > 0$ it follows that the rate of secure state distillation cannot exceed $\mathrm{H}\left(X|E\right)_\rho$. This proves the desired statement. □

Considering the intuitive interpretation of the conditional von Neumann entropy $\mathrm{H}\left(X|E\right)_\rho$ as a measure of uncertainty concerning the state of system X given the state of system E, the statement of Proposition 17 is not surprising. In fact, the result supports this interpretation of the conditional von Neumann entropy; the number of extractable bits with respect to which Eve is negligible should indeed correspond to a measure of Eve's uncertainty about the original letter $x \in \Sigma$ encoded in the state of system X.

2.2 Secure States with respect to a Restricted Eavesdropper

Again in this section, we will be concerned with a two-party setting involving Xavier and Eve described by a cq-state as given in (2.1). This time, however, we consider the situation of Eve having imperfect quantum memory, that is, upon obtaining information about the state of Xavier's system X encoded in the state of an additional quantum system E, then she has to apply some operation, say, a measurement if she has no quantum memory at all. We describe this by letting $\mathcal{C}\left(E\right)$ denote a set of channels on Eve's system E, and the scenario is thus described by

$$\Lambda\left(\rho^{XE}\right) = \sum_{x\in\Sigma} p_x \left|x\right\rangle\!\left\langle x\right|^X \otimes \Lambda\left(\rho_x^E\right), \qquad \rho_x \in \mathcal{D}\left(E\right), \quad \Lambda \in \mathcal{C}\left(E\right), \tag{2.5}$$

We will now discuss to what extent Xavier is able to exploit Eve's imperfect quantum memory in order to extract more secure data. More precisely, we suppose that the set of channels $\mathcal{C}\left(E\right)$ is known to Xavier, while Eve's choice of $\Lambda \in \mathcal{C}\left(E\right)$ remains unknown to him.

We describe the security in a cq-state $\rho \in \mathcal{D}\left(XE\right)$ with respect to a $\mathcal{C}\left(E\right)$-restricted eavesdropper in terms of proximity to a secure state.

Definition 18. Let Σ be an alphabet and consider a cq-state $\rho^{XE} \in \mathcal{D}(XE)$ given by

$$\rho^{XE} = \sum_{x \in \Sigma} p_x \left|x\right\rangle\!\left\langle x\right|^X \otimes \rho_x^E.$$

Let $\mathcal{C}(E)$ be a set of quantum channels. For $\varepsilon > 0$, we say that ρ is an *ε-approximate secure state with respect to an $\mathcal{C}(E)$-restricted eavesdropper on E*, if

$$\Delta_{\mathcal{C}(E)}\left(X|E\right)_\rho < \varepsilon.$$

Any approximate secure state is also an approximate secure state with respect to a restricted eavesdropper due to the monotonicity of the trace norm. Below, we exhibit an example showing that even orthogonal states may give rise to an ε-approximate secure state with respect to a restricted eavesdropper for small values of $\varepsilon > 0$. Here, we consider a bipartite setup, where we think of the eavesdropper as two spatially separated parties Alice and Bob. In this setup, a natural limitation is local operations on their individual systems and classical communication.

Example 19. Let $\sigma \in \mathcal{D}(XAB)$ be a cqq-state given by

$$\sigma = \frac{1}{2}\left|0\right\rangle\!\left\langle 0\right|^X \otimes \sigma_0^{AB} + \frac{1}{2}\left|1\right\rangle\!\left\langle 1\right|^X \otimes \sigma_1^{AB},$$

where σ_0, σ_1 denote the normalized projections onto the symmetric and antisymmetric subspace of AB, which is of local dimension $d \in \mathbb{N}$, respectively. These are the extremal Werner states introduced in [27], and we discuss them further in Subsection 2.2.1. As σ_0, σ_1 are orthogonal states, we have

$$\mathrm{Pr}_{\text{guess}}\left(X|AB\right)_\sigma = 1.$$

To see that σ is an ε-approximate secure state with respect to an eavesdropper restricted to PPT measurements, we simply note

$$\Delta_{\mathcal{M}_{\mathrm{PPT}}(A:B)}\left(X|AB\right)_\sigma = \mathrm{Pr}_{\text{guess}}^{\mathcal{M}_{\mathrm{PPT}}(A:B)}\left(X|AB\right)_\sigma - \frac{1}{2} = \frac{1}{d+1},$$

where the first equality follows from Proposition 4, and the second equality is due to [28]. For $\varepsilon = \frac{1}{d+1}$, this proves σ is an ε-approximate secure state.

We now consider the operationally relevant task of Xavier to obtain secure states with respect to an eavesdropper Eve with imperfect quantum memory modeled by a set of channels $\mathcal{C}(E)$. As in the previous section, we allow Xavier to access public randomness and he is furthermore allowed to perform classical processing of the data encoded in system X. This time, however, Xavier can additionally exploit Eve's imperfect memory. The choice of $f \in \mathcal{F}$ is then made known to the eavesdropper Eve, but as she has already applied some quantum channel $\Lambda \in \mathcal{C}(E)$ to her system independently of the choice of f, the resulting cq-state is of the form

$$\mathcal{F}\left(\Lambda\left(\rho\right)\right) = \sum_{f \in \mathcal{F}} p_f f\left(\Lambda\left(\rho\right)\right) \otimes \left|f\right\rangle\!\left\langle f\right|^F,$$

where $f\left(\Lambda\left(\rho\right)\right) = \sum_{x \in \Sigma} p_x \left|f\left(x\right)\right\rangle\!\left\langle f\left(x\right)\right| \otimes \Lambda\left(\rho_x\right)$.

Again, Xavier is not necessarily able to obtain secure states with respect to an eavesdropper with imperfect quantum memory using this approach, so we will consider the asymptotic setting, where he has $n \in \mathbb{N}$ copies of a state ρ, and aims towards obtaining (m, ε)-secure states with respect to an eavesdropper with imperfect memory for $m \in \mathbb{N}$ at fixed rates $\frac{m}{n}$ for any $\varepsilon > 0$.

Definition 20. Let $\rho \in \mathcal{D}(XE)$ be a cq-state, and let $\mathcal{C} = \left(\mathcal{C}_n\left(E^n\right)\right)_{n\in\mathbb{N}}$ be a sequence of sets of channels. Let $r \geq 0$. We say that r is an *achievable rate of secure state distillation with respect to a $\mathcal{C}$-restricted eavesdropper*, if $r = 0$ or the following condition holds: For sufficiently large $n \in \mathbb{N}$ and $m = \lfloor rn \rfloor$ there exists a family $\mathcal{F}$ of functions $f\colon \Sigma^n \to \Sigma'$, where Σ' is of size $|\Sigma'| = 2^m$, with associated probability distribution $\left(p_f\right)_{f\in\mathcal{F}}$, such that

$$\mathcal{F}\left(\rho^{\otimes n}\right) = \sum_{x'\in\Sigma'} |x'\rangle\langle x'|^{X'} \otimes \sum_{f\in\mathcal{F}} p_f \sum_{x\in f^{-1}(\{x'\})} p_x \rho_x^{E^n} \otimes |f\rangle\langle f|^F$$

is an ε-approximate secure state with respect to a $\mathcal{C}_n\left(E^n\right)$-restricted eavesdropper on $E^n F$.

The *rate of secure state distillation with respect to a $\mathcal{C}$-restricted eavesdropper $S_{D,\mathcal{C}}(\rho)$* is the supremum of all achievable rates of secure state distillation with respect to a $\mathcal{C}$-restricted eavesdropper.

Remark. Whenever $\mathrm{id} \in \mathcal{C}_n\left(E^n\right)$ for all $n \in \mathbb{N}$, we shall refer to the rate above as the *rate of secure state distillation with respect to an unrestricted eavesdropper*. Due to the monotonicity of the trace norm, this is equivalent to the previously introduced notion of secure state distillation.

For certain sequences of sets of channels $\mathcal{C}$ we introduce a less verbose notation. For a bipartite quantum system AB we may consider the sequence of sets of all measurements on Alice's system, that is, $\mathcal{M}_{\text{all}}^A = \left(\mathcal{M}_{\text{all}}\left(A^n\right)\right)_{n\in\mathbb{N}}$, and then we write

$$S_{D,A}(\rho) \coloneqq S_{D,\mathcal{M}(A)}(\rho), \qquad \rho \in \mathcal{D}(XAB).$$

This notation specializes to a single quantum system E by simply considering a trivial system as a placeholder for Bob's system B.

Theorem 21. Consider a cq-state $\rho^{XE} \in \mathcal{D}(XE)$. Let $\mathcal{C} = \left(\mathcal{C}_n\left(E^n\right)\right)_{n\in\mathbb{N}}$ be a sequence of sets channels satisfying

$$\mathcal{C}_m\left(E^m\right) \otimes \mathcal{C}_n\left(E^n\right) \subseteq \mathcal{C}_{m+n}\left(E^{m+n}\right)$$

for all $m, n \in \mathbb{N}$. Then

$$\widetilde{\mathrm{H}}_{\min,\mathcal{C}}^{\infty}(X|E)_\rho \leq S_{D,\mathcal{C}}(\rho) \leq \mathrm{H}_{\mathcal{C}}^{\infty}(X|E)_\rho. \tag{2.6}$$

Proof. Let $\varepsilon > 0$. Let $n \in \mathbb{N}$, $\delta > 0$, and consider $m \in \mathbb{N}_0$ given by $m = \left\lfloor n\left(\widetilde{\mathrm{H}}_{\min,\mathcal{C}}^{\infty}(X|E)_\rho - \delta\right)\right\rfloor$. Let $\mathcal{F}$ be a two-universal family of functions $f\colon \Sigma^n \to \{0,1\}^m$, and note that it follows from our adaptation of the leftover hash lemma in Lemma 12 that

$$\begin{aligned}
\Delta_{\mathcal{C}_n(E^n)}\left(X'\middle|E^n F\right)_{\mathcal{F}(\rho^{\otimes n})} &\leq 2\varepsilon + \sqrt{2^{m-\mathrm{H}_{\min,\mathcal{C}_n(E^n)}^{\varepsilon}(X^n|E^n)_{\rho^{\otimes n}}}} \\
&= 2\varepsilon + \sqrt{2^{n\left(\frac{m}{n}-\frac{1}{n}\mathrm{H}_{\min,\mathcal{C}_n(E^n)}^{\varepsilon}(X^n|E^n)_{\rho^{\otimes n}}\right)}}. \\
&\leq 2\varepsilon + \sqrt{2^{n\left(\widetilde{\mathrm{H}}_{\min,\mathcal{C}}^{\infty}(X|E)_\rho-\frac{1}{n}\mathrm{H}_{\min,\mathcal{C}_n(E^n)}^{\varepsilon}(X^n|E^n)_{\rho^{\otimes n}}-\delta\right)}},
\end{aligned} \tag{2.7}$$

where the last inequality is due to our choice of $m \in \mathbb{N}_0$. Furthermore, by definition we have

$$\frac{1}{n}\,\mathrm{H}^{\varepsilon}_{\min,\mathcal{C}_n(E^n)}\left(X^n|E^n\right)_{\rho^{\otimes n}} > \widetilde{\mathrm{H}}^{\infty}_{\min,\mathcal{C}}\left(X|E\right)_{\rho} - \delta$$

for sufficiently large $n \in \mathbb{N}$. This implies the second term in (2.7) tends to 0 as n tends to infinity, which proves $\widetilde{\mathrm{H}}^{\infty}_{\min,\mathcal{C}}\left(X|E\right)_{\rho} - \delta$ is an achievable rate of secure state distillation for all $\delta > 0$. Finally, this proves the achievability of the lower bound in the statement.

For the converse inequality, let $\varepsilon > 0$, $m, n \in \mathbb{N}$ and suppose there exists a family $\mathcal{F}$ of functions $f\colon \Sigma^n \to \Sigma'$, where Σ' is of size$|\Sigma'| = 2^m$, with associated probability distribution $(p_f)_{f\in\mathcal{F}}$, such that

$$\Delta_{\mathcal{C}_n(E^n)}\left(X'\middle|E^n F\right)_{\mathcal{F}(\rho^{\otimes n})} < \varepsilon.$$

Consider an arbitrary $\Lambda \in \mathcal{C}\left(E^n\rangle E_n\right)$ and note

$$\Delta\left(X'\middle|E^n F\right)_{\mathcal{F}(\Lambda(\rho^{\otimes n}))} = \sum_{f\in\mathcal{F}} p_f \Delta\left(X'\middle|E_n\right)_{\Lambda(f(\rho^{\otimes n}))} < \varepsilon.$$

Let $\mathcal{F}_\varepsilon$ denote the set of all $f \in \mathcal{F}$ satisfying

$$\Delta\left(X'\middle|E_n\right)_{\Lambda(f(\rho^{\otimes n}))} < \sqrt{\varepsilon}.$$

It follows from Winter's refined version of the Fannes-Audenart inequality (see Theorem 86 in Appendix A) that for $f \in \mathcal{F}_\varepsilon$ we have

$$\mathrm{H}\left(X'\middle|E_n\right)_{\omega^{X'}\otimes\Lambda(f(\rho^{\otimes n}))} \le \mathrm{H}\left(X'\middle|E_n\right)_{\Lambda(f(\rho^{\otimes n}))} + 2\sqrt{\varepsilon}\log d_{X'} + g\left(\sqrt{\varepsilon}\right). \tag{2.8}$$

Furthermore, note that

$$\varepsilon > \sum_{f\in\mathcal{F}\backslash\mathcal{F}_\varepsilon} p_f \Delta\left(X'\middle|E_n\right)_{\Lambda(f(\rho^{\otimes n}))} \ge \sqrt{\varepsilon} \sum_{f\in\mathcal{F}\backslash\mathcal{F}_\varepsilon} p_f,$$

which in turn implies $\sum_{f\in\mathcal{F}\backslash\mathcal{F}_\varepsilon} p_f \le \sqrt{\varepsilon}$. This allows us to deduce

$$\sum_{f\in\mathcal{F}\backslash\mathcal{F}_\varepsilon} p_f\,\mathrm{H}\left(X'\middle|E_n\right)_{\omega^{X'}\otimes\Lambda(f(\rho^{\otimes n}))} \le \sqrt{\varepsilon}\log d_{X'}. \tag{2.9}$$

If we combine these initial observations, we may note that

$$\begin{aligned} m &= \sum_{f\in\mathcal{F}} p_f\,\mathrm{H}\left(X'\middle|E_n\right)_{\omega^{X'}\otimes\Lambda(f(\rho^{\otimes n}))} \\ &\le \sum_{f\in\mathcal{F}} p_f\,\mathrm{H}\left(X'\middle|E_n\right)_{\Lambda(f(\rho^{\otimes n}))} + 3\sqrt{\varepsilon}\log d_{X'} + g\left(\sqrt{\varepsilon}\right) \\ &\le \mathrm{H}\left(X|E_n\right)_{\Lambda(\rho^{\otimes n})} + 3n\sqrt{\varepsilon}\log d_X + g\left(\sqrt{\varepsilon}\right) \end{aligned}$$

where the first inequality is due to (2.8) and (2.9). The second inequality is due to a data processing inequality discussed in Subsection 1.5.1 of Chapter 1, and the fact that we may assume $d_{X'} \leq d_{X^n}$ without loss of generality. If we divide both sides of the inequality by n, we have

$$\frac{m}{n} \leq \frac{1}{n}\,\mathrm{H}\left(X|E_n\right)_{\Lambda(\rho^{\otimes n})} + 3\sqrt{\varepsilon}\log d_X + \frac{1}{n} g\left(\sqrt{\varepsilon}\right)$$

for all $\Lambda \in \mathcal{C}\left(E^n\rangle E_n\right)$. This proves the rate of secure state distillation with respect to a $\mathcal{C}$-restricted eavesdropper E cannot exceed $\mathrm{H}_{\mathcal{C}}^{\infty}\left(X|E\right)_{\rho}$, and due to Lemma 9 this is a well-defined upper bound. □

There are a few observations to make concerning the statement of Theorem 21. First note that the lower and upper bound on the rate of secure state distillation with respect to a restricted eavesdropper coincide if we assume the statement of Conjecture 10 is true. Secondly, it is not hard to see that whenever the eavesdropper acts independently in each round, that is,

$$\mathcal{C}_n\left(E^n\right) = \mathcal{C}_{n-1}\left(E^{n-1}\right) \otimes \mathcal{C}'_n\left(E\right), \qquad \mathcal{C}'_n\left(E\right) \subseteq \mathcal{C}_{\text{all}}\left(E\right)$$

for all $n \in \mathbb{N}$, then the lower and upper bound coincide.

We finish this section by remarking upon two situations, where the upper bound on the rate of secure state distillation with respect to a restricted eavesdropper does not need to be regularized.

Corollary 22. Let Σ be an alphabet and consider a cq-state $\rho_{XAB} \in \mathcal{D}\left(XAB\right)$. Suppose $\rho_x \in \mathcal{D}_{\text{sep}}\left(AB\right)$ for all $x \in \Sigma$. Then

$$\widetilde{\mathrm{H}}^{\infty}_{\min,A}\left(X|AB\right)_{\rho} \leq S_{D,A}\left(\rho\right) \leq \mathrm{H}_A\left(X|AB\right)_{\rho}.$$

Proof. The additivity of locally accessible information for separable states is shown in Section VII.A of [29]. Together with Theorem 21 this yields the desired statement. □

Corollary 23. Consider a cq-state $\rho_{XE} \in \mathcal{D}\left(XE\right)$. Then

$$\widetilde{\mathrm{H}}^{\infty}_{\min,E}\left(X|E\right)_{\rho} \leq S_{D,E}\left(\rho\right) \leq \mathrm{H}_E\left(X|E\right)_{\rho}.$$

Proof. The additivity of locally accessible information for separable states is shown in Section VII.A of [29], and this generalizes to accessible information. Together with Theorem 21 this yields the desired statement. □

2.2.1 Example: Secure State with Extremal Werner States

We are now in a position to discuss the rate of secure state distillation with respect to a restricted eavesdropper in comparison to the rate of secure state distillation introduced in Section 2.1. We saw in Example 19 that information about the state of a system X encoded into orthogonal states of a bipartite system AB can give rise to an ε-secure state with respect to a restricted eavesdropper for arbitrary $\varepsilon > 0$. In the following, we will again consider the cq-state $\sigma \in \mathcal{D}\left(XAB\right)$ from Example 19 given by

$$\sigma^{XAB} = \frac{1}{2}\left|0\right\rangle\!\left\langle 0\right|^X \otimes \sigma_0^{AB} + \frac{1}{2}\left|0\right\rangle\!\left\langle 0\right|^X \otimes \sigma_1^{AB}, \tag{2.10}$$

where σ_0, σ_1 respectively denote the normalized projections onto the symmetric and antisymmetric subspace of AB, which is of local dimension $d \in \mathbb{N}$. We begin by introducing a slightly more general notation.

Definition 24. Let AB be a bipartite quantum system of local dimension $d \in \mathbb{N}$. The projections onto the symmetric and antisymmetric subspace of AB are given by

$$\Pi_0 := \frac{1}{2}\left(\mathbb{1} + \mathbb{F}\right), \qquad \Pi_1 := \frac{1}{2}\left(\mathbb{1} - \mathbb{F}\right),$$

where $\mathbb{F} = \sum_{i,j=1}^{d} |ij\rangle\langle ji|$. The *extremal Werner states* $\sigma_0, \sigma_1 \in \mathcal{D}\left(AB\right)$ are given by

$$\sigma_0 = \frac{2}{d\left(d+1\right)}\Pi_0, \qquad \sigma_1 = \frac{2}{d\left(d-1\right)}\Pi_1,$$

and more generally we define the *Werner state* $\sigma_p \in \mathcal{D}\left(AB\right)$ by

$$\sigma_p = \left(1-p\right)\sigma_0 + p\sigma_1, \qquad p \in \left[0,1\right].$$

For $n \in \mathbb{N}$ and $\alpha, x \in \left\{0,1\right\}^n$ we denote by

$$\Pi_\alpha = \bigotimes_{i=1}^{n} \Pi_{\alpha_i}, \qquad \sigma_x = \bigotimes_{j=1}^{n} \sigma_{x_j}.$$

We now set out to derive a lower bound on the rate of secure state distillation of σ given by (2.10) with respect to a $\mathcal{M}_{\mathrm{ppt}}\left(A:B\right)$-restricted eavesdropper. Our method of proof is an adaptation of the techniques introduced in [28], where the key technical observation is the invariance of Werner states under certain unitaries, more precisely that $\rho \in \mathcal{D}\left(AB\right)$ is a Werner state, if and only if ρ is invariant under all bi-unitary transformations [27], that is,

$$\left(U \otimes U\right)\rho\left(U \otimes U\right)^\dagger = \rho$$

for all unitaries $U \in \mathrm{U}\left(A\right)$, $U \in \mathrm{U}\left(B\right)$.

The defining property of Werner states outlined above plays a central role in our efforts to lower bound the rate of secure state distillation of the state given by (2.10). In the result below, we show how the property of bi-unitary invariance of Werner states can be transferred to the POVM representation of measurements applied to systems AB.

Lemma 25. Let AB be a bipartite quantum system of local dimension $d \in \mathbb{N}$. Let $\sigma \in \mathcal{D}\left(XAB\right)$ be given by

$$\sigma^{XAB} = \frac{1}{2}\left|0\right\rangle\!\left\langle 0\right|^X \otimes \sigma_0^{AB} + \frac{1}{2}\left|0\right\rangle\!\left\langle 0\right|^X \otimes \sigma_1^{AB}.$$

There exists a measurement with POVM representation $\left\{M_x\right\}_{x \in \{0,1\}^n}$ satisfying

$$\Pr\nolimits_{\mathrm{guess}}^{\mathcal{M}_{\mathrm{ppt}}(A:B)}\left(X^n | A^n B^n\right)_{\sigma^{\otimes n}} = \frac{1}{2^n}\sum_{x \in \{0,1\}^n} \operatorname{Tr} M_x \sigma_x,$$

where $M_x^\Gamma \geq 0$ and M_x is invariant under all $U \in \mathrm{U}\left(A^n B^n\right)$, which is bi-unitary on each copy of AB.

Proof. This was shown in [28] using an observation from [30]. □

Lemma 26. Let AB be a bipartite quantum system of local dimension $d \in \mathbb{N}$. Let $\sigma \in \mathcal{D}(XAB)$ be given by

$$\sigma^{XAB} = \frac{1}{2} |0\rangle\langle 0|^{X} \otimes \sigma_0^{AB} + \frac{1}{2} |0\rangle\langle 0|^{X} \otimes \sigma_1^{AB}.$$

For $n \in \mathbb{N}$ there exists a measurement with POVM representation $\{M_x\}_{x\in\{0,1\}^n}$ satisfying

$$\mathrm{Pr}_{\text{guess}}^{\mathcal{M}_{\text{ppt}}(A:B)} (X^n|A^nB^n)_{\sigma^{\otimes n}} = \frac{1}{2^n} \sum_{x\in\{0,1\}^n} \operatorname{Tr} M_x\sigma_x,$$

where $M_x^\Gamma \geq 0$ and

$$M_x = \sum_{\alpha\in\{0,1\}^n} \lambda_{x,\alpha}\Pi_\alpha, \qquad \lambda \in \mathbb{R}^{2^{2n}}.$$

Remark. This was mentioned without proof in [28].

Proof. First note that Lemma 25 allows us to infer the existence of an optimal measurement with POVM representation $\{M_x\}_{x\in\{0,1\}^n} \subseteq \mathrm{L}(A^nB^n)$ satisfying $M_x^\Gamma \geq 0$, which is invariant under bi-unitary transformations of all subsystems AB. If we enumerate the subsystems, that is, $A^n = A_1 \dots A_n$, $B^n = B_1 \dots B_n$, then it follows that $M_x^{A_iB_i}$ is invariant under local unitaries, and so we may infer that

$$M_x^{A_iB_i} = \lambda_{x,0}\Pi_0 + \lambda_{x,1}\Pi_1.$$

Furthermore, as the partial transpose is taken with respect to a choice of computational basis, it follows that $M_x^{A_iB_i}$ is PPT.

Now define $N_x = \bigotimes_{i=1}^n M_x^{A_iB_i}$, and note that $N_x^\Gamma \geq 0$. As a basis of each subsystem A_iB_i we may choose the symmetric and anti-symmetric vectors, that is,

$$\frac{1}{\sqrt{2}} (|ij\rangle + |ji\rangle), \qquad \frac{1}{\sqrt{2}} (|ij\rangle - |ji\rangle).$$

In this basis, the difference between M_x and N_x is off-diagonal, and since σ_x is diagonal in this basis, it follows that

$$\operatorname{Tr} M_x\sigma_x = \operatorname{Tr} N_x\sigma_x.$$

To see that the operators N_x for $x \in \{0,1\}^n$ sum to identity, we first note that $\{M_x\}_{x\in\{0,1\}^n}$ is a POVM representation of a measurement. This implies that the off-diagonal entries of the operators M_x sum to zero. As the difference $M_x - N_x$ is zero on the diagonal in the symmetric and anti-symmetric basis, and the off-diagonal entries sum to zero, this implies

$$\sum_{x\in\{0,1\}^n} N_x = \sum_{x\in\{0,1\}^n} M_x = \mathbb{1}_{A^nB^n},$$

which proves $\{N_x\}_{x\in\{0,1\}^n}$ is a POVM representation of a measurement. □

Lemma 27. Let AB be a bipartite quantum system of local dimension $d \in \mathbb{N}$. Let $\sigma \in \mathcal{D}(XAB)$ be given by

$$\sigma^{XAB} = \frac{1}{2} |0\rangle\langle 0|^{X} \otimes \sigma_0^{AB} + \frac{1}{2} |0\rangle\langle 0|^{X} \otimes \sigma_1^{AB}.$$

There exists a measurement with POVM representation $\{M_x\}_{x\in\{0,1\}^n}$ satisfying

$$\mathrm{Pr}_{\mathrm{guess}}^{\mathcal{M}_{\mathrm{ppt}}(A:B)}\left(X|AB\right)_\sigma = \sum_{x\in\{0,1\}^n} p_x \operatorname{Tr} M_x \sigma_x,$$

where $M_x^\Gamma \geq 0$ and $M_x = \sum_{\alpha\in\{0,1\}^n} \lambda_{x,\alpha}\Pi_\alpha$ for some $\lambda \in \mathbb{R}^{2^{2n}}$ satisfying

$$0 \leq \lambda, \qquad \sum_{x\in\{0,1\}^n} \lambda_{x,\alpha} \leq 1, \qquad \text{and} \qquad \frac{1}{2^n}\sum_{\alpha\in\{0,1\}^n} \lambda_{x,\alpha}\left(1+d\right)^{|\beta|-\alpha\cdot\beta}\left(1-d\right)^{\alpha\cdot\beta} \geq 0$$

for all $\alpha, \beta \in \{0,1\}^n$.

Proof. First note that Lemma 25 allows us to infer the existence of an optimal measurement with POVM representation $\{M_x\}_{x\in\{0,1\}^n}$ satisfying $M_x^\Gamma \geq 0$, which is invariant under bi-unitary transformations of all subsystems AB. It follows from Lemma 26 that we may assume that

$$M_x = \sum_{\alpha\in\{0,1\}^n} \lambda_{x,\alpha}\Pi_\alpha$$

for some $\lambda \in \mathbb{R}^{2^{2n}}$. As $\{M_x\}_{x\in\{0,1\}^n}$ constitutes the POVM representation of a measurement, we have

$$0 \leq \lambda, \qquad \sum_{x\in\{0,1\}^n} \lambda_{x,\alpha} \leq 1$$

for all $\alpha \in \{0,1\}^n$. If we denote by $\varphi_d = |\varphi_d\rangle\langle\varphi_d|$, where $|\varphi_d\rangle = \frac{1}{\sqrt{d}}\sum_{i=1}^d |ii\rangle$, then we may note that

$$\Pi_0^\Gamma = \frac{1}{2}\left(\mathbb{1} + d\varphi_d\right) = \frac{1}{2}\left(\left(\mathbb{1}-\varphi_d\right) + \left(1+d\right)\varphi_d\right)$$
$$\Pi_1^\Gamma = \frac{1}{2}\left(\mathbb{1} - d\varphi_d\right) = \frac{1}{2}\left(\left(\mathbb{1}-\varphi_d\right) + \left(1-d\right)\varphi_d\right)$$

and so

$$\Pi_\alpha^\Gamma = \frac{1}{2^n}\bigotimes_{i=1}^n \left(\left(\mathbb{1}-\varphi_d\right) + \left(1+(-1)^{\alpha_i}d\right)\varphi_d\right) = \frac{1}{2^n}\sum_{\beta\in\{0,1\}^n}\left(1+d\right)^{|\beta|-\alpha\cdot\beta}\left(1-d\right)^{\alpha\cdot\beta} B_\beta,$$

where $B_\beta = B_{\beta_1}\otimes\ldots\otimes B_{\beta_n}$ with $B_0 = \mathbb{1}-\varphi_d$ and $B_1 = \varphi_d$. The fact $0 \leq M_x^\Gamma$ translates to

$$\frac{1}{2^n}\sum_{\alpha,\beta\in\{0,1\}^n} \lambda_{x,\alpha}\left(1+d\right)^{|\beta|-\alpha\cdot\beta}\left(1-d\right)^{\alpha\cdot\beta} B_\beta \geq 0,$$

which is equivalent to

$$\frac{1}{2^n}\sum_{\alpha\in\{0,1\}^n} \lambda_{x,\alpha}\left(1+d\right)^{|\beta|-\alpha\cdot\beta}\left(1-d\right)^{\alpha\cdot\beta} \geq 0$$

for all $\beta \in \{0,1\}^n$ due to orthogonality of B_β, $B_{\beta'}$ for distinct $\beta \neq \beta'$. This proves the desired statement. □

Lemma 28. Let AB be a bipartite quantum system of local dimension $d \in \mathbb{N}$. Let $\sigma \in \mathcal{D}(XAB)$ be given by

$$\sigma^{XAB} = \frac{1}{2} |0\rangle\langle 0|^X \otimes \sigma_0^{AB} + \frac{1}{2} |0\rangle\langle 0|^X \otimes \sigma_1^{AB}.$$

Define $Q \in \mathbb{R}^{2^n \times 2^n}$ by $Q_{\beta,\alpha} = \frac{1}{2^n} (1+d)^{|\beta| - \alpha\cdot\beta} (1-d)^{\alpha\cdot\beta}$, and let $b \in \mathbb{R}^{2^{2n}+2^n}$, $c \in \mathbb{R}^{2^{2n}}$ be given by

$$b_i = \begin{cases} 0, & i = 1, \dots, 2^{2n}, \\ -1, & i = 2^{2n} + 2^n + 1, \dots, 2^{2n} + 2^{n+1} \end{cases}, \qquad c_{x,\alpha} = (1 - \delta_{x=0^n}) (\delta_{\alpha=0^n} - \delta_{x=\alpha}).$$

Then the optimal probability of correctly guessing X^n based on a measurement with of the systems $A^n B^n$, where the POVM representation consists of $\mathcal{M}_{\mathrm{ppt}}(A:B)$ operators, is given by the following semi-definite program:

$$\frac{1}{2^n} - \frac{1}{2^n} \min_{\lambda \in \mathbb{R}^{2^{2n}}} c^T \cdot \lambda, \qquad \text{where} \begin{pmatrix} & \mathbb{1}_{2^n} \otimes Q & \\ -\mathbb{1} & \dots & -\mathbb{1} \end{pmatrix} \lambda \geq b, \quad \lambda \geq 0.$$

Proof. First note that by Lemma 27 there exists an optimal PPT measurement for guessing X^n based on a measurement of systems $A^n B^n$ with POVM representation $\{M_x\}_{x \in \{0,1\}^n}$ given by

$$M_x = \sum_{\alpha \in \{0,1\}^n} \lambda_{x,\alpha} \Pi_\alpha,$$

where $\lambda \in \mathbb{R}^{2^{2n}}$ satisfies

$$0 \leq \lambda, \qquad \sum_{x \in \{0,1\}^n} \lambda_{x,\alpha} \leq 1, \qquad \text{and} \qquad \frac{1}{2^n} \sum_{\alpha \in \{0,1\}^n} \lambda_{x,\alpha} (1+d)^{|\beta| - \alpha\cdot\beta} (1-d)^{\alpha\cdot\beta} \geq 0$$

for all $\alpha, \beta \in \{0,1\}^n$. From this, we may express the probability of correctly guessing X^n based on a PPT measurement of systems $A^n B^n$ as a maximization of the expression

$$\frac{1}{2^n} \sum_{x \in \{0,1\}^n} \operatorname{Tr} M_x \sigma_x = \frac{1}{2^n} \sum_{x \in \{0,1\}^n} \lambda_{x,x} = \frac{1}{2^n} - \frac{1}{2^n} \sum_{x \in \{0,1\}^n, x \neq 0^n} (\lambda_{x,0} - \lambda_{x,x}).$$

With $Q \in \mathbb{R}^{2^n \times 2^n}$ and $b \in \mathbb{R}^{2^{2n}+2^n}$, $c \in \mathbb{R}^{2^{2n}}$ given as in the statement of the result, the maximization above is exactly the desired statement. □

Proposition 29. Let AB be a bipartite quantum system of local dimension $d \in \mathbb{N}$. Let $\sigma \in \mathcal{D}(XAB)$ be given by

$$\sigma^{XAB} = \frac{1}{2} |0\rangle\langle 0|^X \otimes \sigma_0^{AB} + \frac{1}{2} |0\rangle\langle 0|^X \otimes \sigma_1^{AB}.$$

Then

$$\left(\frac{1}{2} + \frac{1}{d+1}\right)^n \leq \Pr_{\mathrm{guess}}^{\mathcal{M}_{\mathrm{ppt}}(A:B)} (X^n | A^n B^n)_{\sigma^{\otimes n}} \leq \left(\frac{1}{2} + \frac{1}{d-1} + \frac{1}{2(d+1)}\right)^n.$$

Proof. Applying the PPT measurement with POVM representation $\{M_x\}_{x\in\{0,1\}^n}$ given by

$$M_x = \bigotimes_{i=1}^{n} M_{x_i}, \qquad \text{where } M_0 = \frac{2}{d+1}\Pi_0, \quad M_1 = \frac{d-1}{d+1}\Pi_0 + \Pi_1$$

yields a lower bound given by

$$\begin{aligned}\frac{1}{2^n}\sum_{x\in\{0,1\}^n} \operatorname{Tr} M_x\sigma_x &= \frac{1}{2^n}\sum_{x\in\{0,1\}^n}\left(\frac{2}{d+1}\right)^{n-|x|}\\ &= \sum_{k=0}^{n}\binom{n}{n-k}\left(\frac{1}{d+1}\right)^{n-k}\frac{1}{2^k}\\ &= \left(\frac{1}{2}+\frac{1}{d+1}\right)^n.\end{aligned}$$

To find an upper bound, we consider the minimization problem in Lemma 28 and note that the dual program is given by

$$\max_{\mu} b^T\cdot\mu, \qquad \text{where } \begin{pmatrix} & \mathbb{1} & -\mathbb{1}\\ \mathbb{1}_{2^n}\otimes Q^T & \vdots & \vdots\\ & \mathbb{1} & -\mathbb{1}\end{pmatrix}\mu \le c \quad \mu\ge 0.$$

Denote the entries of $\mu\in\mathbb{R}^{2^{2n}+2^n}$ by

$$\mu = \left(\mu_{0^n,0^n},\dots,\mu_{0^n,1^n},\dots,\mu_{1^n,0^n},\dots,\mu_{1^n,1^n},\nu_{0^n},\dots,\nu_{1^n}\right)^T,$$

and let $\mu_x = (\mu_{x,0^n},\dots,\mu_{x,1^n})^T$, $Q_\beta = (Q_{0^n,\beta},\dots,Q_{1^n,\beta})^T$ and $\nu = (\nu_{0^n},\dots,\nu_\beta,\dots,\nu_{1^n})$. Then the dual program can be rewritten as

$$\max_{\mu}\sum_{\beta\in\{0,1\}^n} -\nu_\beta, \qquad \text{where } \mu\ge 0, \qquad Q_\beta^T\mu_x - \nu_\beta \le c_{x,\beta}$$

for all $x,\beta\in\{0,1\}^n$, or equivalently

$$\max_{\mu}\sum_{\beta\in\{0,1\}^n} -\nu_\beta, \qquad \text{where } \mu\ge 0, \qquad Q_\beta^T\mu_x - c_{x,\beta} \le \nu_\beta.$$

Any feasible point of the dual program is a lower bound on the primal program, so define

$$\mu^*_{x,\alpha} = \underbrace{\frac{(d+1)^{|x|}}{d^{|x|}\,(d+1)^{|\alpha|}}\cdot\left(\frac{d+1}{d-1}\right)^{\alpha\cdot x}}_{\mu^*_{x,\alpha}{}^{+}} - \underbrace{\frac{(d+1)^{|x|}}{d^{|x|}\,(d+1)^{|\alpha|}}\cdot(-1)^{\alpha\cdot x}}_{\mu^*_{x,\alpha}{}^{-}} \ge 0.$$

First, in an effort to calculate $Q_\beta^T\mu_x^{*+}$, we note

$$\begin{aligned}Q_\beta^T\mu_x^{*+} &= \frac{(d+1)^{|x|}}{2^n d^{|x|}}\sum_{\alpha\in\{0,1\}^n}\left(\frac{1-d}{1+d}\right)^{\alpha\cdot\beta}\left(\frac{d+1}{d-1}\right)^{\alpha\cdot x}\\ &= \frac{(d+1)^{|x|}}{2^n d^{|x|}}\sum_{\alpha\in\{0,1\}^n}(-1)^{\alpha\cdot\beta}\left(\frac{d+1}{d-1}\right)^{\alpha\cdot x-\alpha\cdot\beta}.\end{aligned}$$

We now consider two cases. Suppose $\beta \cdot x \neq 0$, that is, there exists $i \in \{0, \ldots, n\}$ such that $\beta_i, x_i = 1$. The terms corresponding to $\alpha_1 \ldots \alpha_i \ldots \alpha_n$ and $\alpha_1 \ldots \overline{\alpha}_i \ldots \alpha_n$ cancel in the expression above, which implies the sum equals 0. Thus we restrict ourselves to the situation where $\beta \cdot x = 0$. Then

$$\begin{aligned}
Q_\beta^T \mu_x^{*+} &= \frac{(d+1)^{|x|}}{2^n d^{|x|}} 2^{n-|\beta|-|x|} \sum_{k=0}^{|\beta|} \binom{|\beta|}{k} (-1)^k \left(\frac{d-1}{d+1}\right)^k \sum_{l=0}^{|x|} \binom{|x|}{l} \left(\frac{d+1}{d-1}\right)^l \\
&= \frac{(d+1)^{|x|}}{2^n d^{|x|}} 2^{n-|\beta|-|x|} \left(1 - \frac{d-1}{d+1}\right)^{|\beta|} \left(1 + \frac{d+1}{d-1}\right)^{|x|} \\
&= \frac{1}{(d+1)^{|\beta|}} \left(\frac{d+1}{d-1}\right)^{|x|}
\end{aligned}$$

Next, in an effort to calculate $Q_\beta^T \mu_x^-$, we note

$$Q_\beta^T \mu_x^{*-} = \frac{(d+1)^{|x|}}{2^n d^{|x|}} \frac{1}{2^n} \sum_{\alpha \in \{0,1\}^n} (-1)^{\alpha \cdot x} \left(\frac{1-d}{1+d}\right)^{\alpha \cdot \beta}.$$

Similarly to the argument above, we consider two cases. Suppose $x \not\leq \beta$, namely, there exists some $i \in \{0, \ldots, n\}$ such that $\beta_i = 0$, $x_i = 1$. Then the terms corresponding to $\alpha_1 \ldots \alpha_i \ldots \alpha_n$ and $\alpha_1 \ldots \overline{\alpha}_i \ldots \alpha_n$ cancel in the expression above, which implies the sum equals 0. Thus we restrict ourselves to the situation where $x \leq \beta$. Then

$$\begin{aligned}
Q_\beta^T \mu_x^{*-} &= \frac{(d+1)^{|x|}}{2^n d^{|x|}} 2^{n-|\beta|} \sum_{k=0}^{|x|} (-1)^k \binom{|x|}{k} \left(\frac{1-d}{1+d}\right)^k \sum_{l=0}^{|\beta|-|x|} \binom{|\beta|-|x|}{l} \left(\frac{1-d}{1+d}\right)^l \\
&= \frac{(d+1)^{|x|}}{2^n d^{|x|}} 2^{n-|\beta|} \left(1 - \frac{1-d}{1+d}\right)^{|x|} \left(1 + \frac{1-d}{1+d}\right)^{|\beta|-|x|} \\
&= \frac{(d+1)^{|x|}}{(d+1)^{|\beta|}}.
\end{aligned}$$

Evidently, the optimal choice of ν_β for $\beta \in \{0,1\}^n$, which we denote by ν_β^*, is given by

$$\nu_\beta^* = \max_{x \in \{0,1\}^n} \left(Q_\beta^T \mu_x^* - c_{x,\beta}, 0\right).$$

To simplify the expression on the right-hand side, we consider the various cases of $\beta \in \{0,1\}^n$.

Case 1: $\beta = 0^n$. For any $x \in \{0,1\}^n$ we have

$$\begin{aligned}
Q_{0^n}^T \mu_x^* - c_{x,0^n} &= Q_{0^n}^T \mu_x^{*+} - Q_{0^n}^T \mu_x^{*-} - c_{x,0^n} \\
&= \left(\frac{d+1}{d-1}\right)^{|x|} - \delta_{x,0^n} - (1 - \delta_{x,0^n})(1 - \delta_{x,0^n}) \\
&= \left(\frac{d+1}{d-1}\right)^{|x|} - \delta_{x,0^n}.
\end{aligned}$$

Case 2: $\beta \neq 0^n$. First, note that for $x = 0^n$ we have

$$Q_{0^n}^T \mu_x^* - c_{0^n,\beta} = \frac{1}{(d+1)^{|\beta|}} - \frac{1}{(d+1)^{|\beta|}} = 0,$$

so consider $x \in \{0,1\}^n$, $x \neq 0^n$.

Case 2.1: $x \cdot \beta = 0$. As $x \neq 0^n$, we must have $x \not\leq \beta$, which implies

$$Q_\beta^T \mu_x^* - c_{x,\beta} = \frac{1}{(d+1)^{|\beta|}} \left(\frac{d+1}{d-1}\right)^{|x|}.$$

Case 2.2: $x \leq \beta$. As $x \neq 0^n$, we must have $x \cdot \beta \neq 0$, which implies

$$Q_\beta^T \mu_x^* - c_{x,\beta} = -\frac{(d+1)^{|x|}}{(d+1)^{|\beta|}} + \delta_{x,\beta} \leq 0.$$

Case 2.3: $x \cdot \beta \neq 0$ and $x \not\leq \beta$. Then we have

$$Q_\beta^T \mu_x^* - c_{x,\beta} = 0.$$

From the observations above, we deduce

$$\nu_\beta^* = \frac{1}{(d+1)^{|\beta|}} \left(\frac{d+1}{d-1}\right)^{n-|\beta|} - \delta_{\beta,0^n}.$$

for all $\beta \in \{0,1\}^n$. Thus we see a lower bound of the maximization problem at hand is given by

$$\begin{aligned}
\max_\mu b^T \cdot \mu &\geq -\sum_{\beta \in \{0,1\}^n} \nu_\beta^* \\
&= 1 - \sum_{\beta \in \{0,1\}^n} \frac{1}{(d+1)^{|\beta|}} \left(\frac{d+1}{d-1}\right)^{n-|\beta|} \\
&= 1 - \sum_{k=0}^n \binom{n}{k} \frac{1}{(d+1)^k} \left(\frac{d+1}{d-1}\right)^{n-k} \\
&= 1 - \left(\frac{1}{d+1} + \frac{d+1}{d-1}\right)^n.
\end{aligned}$$

Using this as a lower bound for the primal program yields the desired lower bound on the probability of correctly guessing X^n based on a measurement of the systems $A^n B^n$ given $\sigma^{\otimes n}$. □

Corollary 30. Let $\sigma \in \mathcal{D}(XAB)$ be given by

$$\sigma = \frac{1}{2} |0\rangle\langle 0| \otimes \sigma_0 + \frac{1}{2} |1\rangle\langle 1| \otimes \sigma_1.$$

Then

$$1 - \frac{3}{d-1} \log e \leq S_{D,\mathcal{M}_{\mathrm{ppt}(A:B)}}(\sigma)$$

Proof. First, note that from Theorem 21 and the fact that

$$\mathrm{H}^{\infty}_{\min,\mathcal{M}_{\mathrm{ppt}}(A:B)}\left(X|AB\right)_{\sigma} \leq \widetilde{\mathrm{H}}^{\infty}_{\min,\mathcal{M}_{\mathrm{ppt}}(A:B)}\left(X|AB\right)_{\sigma},$$

we may deduce

$$S_{D,\mathcal{M}_{\mathrm{ppt}}(A:B)}\left(\sigma\right) \geq \mathrm{H}^{\infty}_{\min,\mathcal{M}_{\mathrm{ppt}}(A:B)}\left(X|AB\right)_{\sigma}.$$

Due to the operational interpretation of the conditional min-entropy [16], it follows from Proposition 29 that we have the lower bound

$$S_{D,\mathcal{M}_{\mathrm{ppt}}(A:B)}\left(\sigma\right) \geq -\log\left(\frac{1}{2}+\frac{1}{d-1}+\frac{1}{2\left(d+1\right)}\right) \geq 1-\frac{3}{d-1}\log e,$$

which proves the desired inequality. □

Our efforts in the lengthy calculations above yield a lower bound on the rate of secure state distillation from σ which is essentially optimal for large values of $d \in \mathbb{N}$, namely,

$$S_{D,\mathcal{M}_{\mathrm{ppt}}(A:B)}\left(\sigma\right) \geq 1-O\left(\frac{1}{d}\right).$$

Considering the result from [28], it is tempting to conjecture that the lower bound in Proposition 29 is tight and this method proof is sufficient to prove this result with a suitable choice of feasible solution.

This example shows that there can be a large gap between the rate of secure state distillation with respect to a restricted and unrestricted eavesdropper, as we note

$$S_{D,\mathcal{M}_{\mathrm{ppt}}(A:B)}\left(\sigma\right) \geq 1-O\left(\frac{1}{d}\right) \gg 0 = S_{D}\left(\sigma\right),$$

which is of interest in itself.

2.2.2 Deterministic Secure State Distillation

We saw in the proof of Theorem 21 that Xavier's ability to choose a probabilistic strategy for obtaining a secure state is key to achieving the desired rate of secure state distillation. In the following, we revisit the notion of secure state distillation and introduce a new variation, namely, deterministic secure state distillation. We still consider a two-party setting involving Xavier and Eve described by the cq-state in (2.1), however, this time we consider the situation of Xavier being restricted to a deterministic protocol in order to distill a secure state. We describe this by letting $\mathcal{C}\left(E\right)$ denote a set of channels on Eve's system E, and we allow Eve to choose her channel $\Lambda \in \mathcal{C}\left(E\right)$ depending on Xavier's strategy to extract secure information. In the following, we discuss to what extent Xavier is able to exploit Eve's limitation to a certain set of channels in order to extract secure information when his strategy is known to Eve.

Definition 31. Let $\rho \in \mathcal{D}\left(XE\right)$ be a cq-state, and let $\mathcal{C} = \left(\mathcal{C}_n\left(E^n\right)\right)_{n\in\mathbb{N}}$ be a sequence of sets of channels. Let $r \geq 0$. We say that r is an *achievable rate of deterministic secure state distillation with respect to a $\mathcal{C}$-restricted eavesdropper*, if $r = 0$ or the following condition holds: For sufficiently large $n \in \mathbb{N}$ and $m = \lfloor rn \rfloor$ there exists a function $f\colon \Sigma^n \to \Sigma'$, where $\Sigma' = \{0,1\}^m$, such that

$$f\left(\rho^{\otimes n}\right) = \sum_{x \in \Sigma^n} p_x \left|f(x)\right\rangle\!\left\langle f(x)\right|^{X'} \otimes \rho_x^{E^n} = \sum_{x' \in \Sigma'} \left|x'\right\rangle\!\left\langle x'\right|^{X'} \otimes \sum_{x \in f^{-1}(\{x'\})} p_x \rho_x^{E^n}$$

is an ε-approximate secure state with respect to a $\mathcal{C}_n(E^n)$-restricted eavesdropper on E^n.

The *rate of deterministic secure state distillation with respect to a $\mathcal{C}$-restricted eavesdropper* $S^{\text{det}}_{D,\mathcal{C}}(\rho)$ is the supremum of all achievable rates of deterministic secure state distillation with respect to a $\mathcal{C}$-restricted eavesdropper.

Remark. Naturally, restricting Xavier to a deterministic strategy cannot increase the rate of secure state distillation, that is, $S^{\text{det}}_{D,\mathcal{C}}(\rho) \leq S_{D,\mathcal{C}}(\rho)$.

Let us now consider the natural question of the (in)equivalence of deterministic and probabilistic secure state distillation. We start out by considering particular cq-states $\rho \in \mathcal{D}(XE)$ and families of functions $\mathcal{F}$ with associated probability distributions, where we may obtain an ε-approximate secure state with respect to a $\mathcal{M}(E)$-restricted eavesdropper, while none of the functions $f \in \mathcal{F}$ are useful for deterministic secure state distillation!

Example 32. For $n \in \mathbb{N}$ denote by $\Sigma = \{0,1\}^n$, and let $\rho \in \mathcal{D}(XE)$ be a cq-state given by

$$\rho = \frac{1}{2^n} \sum_{x \in \Sigma} |x\rangle\!\langle x|^X \otimes |\rho_x\rangle\!\langle \rho_x|^E, \qquad \text{where } |\rho_x\rangle = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} (-1)^{x_i} |i\rangle .$$

For $\alpha \in \left(0, \frac{1}{4}\right]$, let $\Sigma' = \{0,1\}^{\alpha n}$. We say that a function $f\colon \Sigma \to \Sigma'$ is an *α-matching* function, if

$$f(x) = \left(x_{i_1} \oplus x_{j_1}\right) \ldots \left(x_{i_{\alpha n}} \oplus x_{j_{\alpha n}}\right),$$

for distinct indices $1 \leq i_k, j_l \leq n$ for all $k, l = 1, \ldots, \alpha n$. Let $\mathcal{F}$ denote the set of all α-matching functions, and suppose $\mathcal{F}$ is equipped with the uniform distribution. For more details on this construction, we refer the reader to Chapter 4.

Let $\varepsilon > 0$, and let $\mathcal{Y}$ be an alphabet of size $|\mathcal{Y}| = n$, and let Y be a system spanned by an orthonormal basis $\{|y\rangle\}_{y \in \mathcal{Y}}$. It has previously been shown [2] that there exists a constant $\gamma > 0$, such that if $\log n \leq \gamma \varepsilon \sqrt{n/\alpha}$, then

$$\Delta_{\mathcal{M}(E\rangle Y)}\left(X' \middle| EF\right)_{\mathcal{F}(\rho)} < \varepsilon.$$

This implies that $\mathcal{F}(\rho)$ is an ε-approximate secure state with respect to a $\mathcal{M}(E\rangle Y)$-restricted eavesdropper Eve.

Conversely, it is not hard to see that for any given $f \in \mathcal{F}$ there exists a measurement $\Lambda_f \in \mathcal{M}(E\rangle Y)$, which reveals one bit of $x' = f(x)$ with probability 2α [2]. This observation provides a lower bound on the $\mathcal{M}(E\rangle Y)$-restricted distance to uniform of $f(\rho)$, more precisely, we have

$$\Delta_{\mathcal{M}(E\rangle Y)}(X'|E)_{f(\rho)} \geq 2\alpha.$$

In turn, this shows that none of the functions $f \in \mathcal{F}$ provide security against a $\mathcal{M}(E\rangle Y)$-restricted eavesdropper. In other words, this shows that none of the functions $f \in \mathcal{F}$ can be used for deterministic secure state distillation with respect to a $\mathcal{M}(E\rangle Y)$-restricted eavesdropper.

The example above shows that the family $\mathcal{F}$ of α-matching functions $f \in \mathcal{F}$ allows Xavier to distill an ε-approximate secure state with respect to a $\mathcal{M}\left(E\rangle Y\right)$-restricted eavesdropper, while none of the functions $f \in \mathcal{F}$ are a suitable choice for the deterministic protocol. However, this does by no means prove the inequivalence of secure state distillation and deterministic secure state distillation as we are not considering an asymptotic scenario, nor have we considered all choices of families of functions $\mathcal{F}$. We pose the equivalence of secure state distillation and deterministic secure state distillation as a conjecture below.

Conjecture 33. Let $\rho \in \mathcal{D}\left(XE\right)$ be a cq-state, and let $\mathcal{C} = \left(\mathcal{C}_n\left(E^n\right)\right)_{n\in\mathbb{N}}$ be a sequence of sets of channels. Then

$$S_{D,\mathcal{C}}^{\mathrm{det}}\left(\rho\right) = S_{D,\mathcal{C}}\left(\rho\right).$$

Remark. In support of the conjecture, we note that the rate of deterministic secure state distillation with respect to an unrestricted eavesdropper is in fact equal to $S_D\left(\rho\right)$. This is not hard to see, as we note

$$\Delta\left(X'|EF\right)_{\mathcal{F}(\rho)} = \sum_{f\in\mathcal{F}} p_f \Delta\left(X'|E\right)_{f(\rho)},$$

and this implies the existence of $f \in \mathcal{F}$ such that $\Delta\left(X'|E\right)_{f(\rho)} \leq \Delta\left(X'|EF\right)_{\mathcal{F}(\rho)}$. This is, however, not surprising, as Eve's optimal strategy does not depend on the choice of f. An eavesdropper with perfect quantum memory can simply choose to store the state of her system without acting upon it, and this is the optimal strategy due to the monotonicity of the trace norm.

A lower bound on the rate of deterministic secure state distillation of PPT states

In this section, we exhibit a general approach to identifying examples of PPT states exhibiting a gap between the rate of deterministic secure state distillation with respect to a locally restricted eavesdropper and the secure state distillation with respect to an unrestricted eavesdropper. The key observation is the following result.

Lemma 34. Let $\rho \in \mathcal{D}\left(XAB\right)$ be a cqq-state. Then

$$\Delta_{\mathcal{C}_{A\leftrightarrow B}}\left(X|AB\right)_{\rho} = \Delta_{\mathcal{C}_{A\leftrightarrow B}}\left(X|AB\right)_{\rho^{\Gamma}}.$$

Proof. This was shown in [31]. □

Proposition 35. Let $\rho \in \mathcal{D}\left(XAB\right)$ be a cqq-state and suppose $\rho^{\Gamma} \geq 0$. Then

$$S_{D,\mathcal{C}_{A\leftrightarrow B}}\left(\rho\right) = S_{D,\mathcal{C}_{A\leftrightarrow B}}\left(\rho^{\Gamma}\right), \qquad S_{D,\mathcal{C}_{A\leftrightarrow B}}^{\mathrm{det}}\left(\rho\right) = S_{D,\mathcal{C}_{A\leftrightarrow B}}^{\mathrm{det}}\left(\rho^{\Gamma}\right).$$

Proof. This is a straightforward application of Lemma 34. □

Corollary 36. Let $\rho \in \mathcal{D}\left(XAB\right)$ be a cqq-state and suppose $\rho^{\Gamma} \geq 0$. Then

$$S_{D,\mathcal{C}_{A\leftrightarrow B}}\left(\rho\right) \geq S_{D,\mathcal{C}_{A\leftrightarrow B}}^{\mathrm{det}}\left(\rho\right) \geq S_D\left(\rho^{\Gamma}\right).$$

Proof. The first inequality is trivial, and the second inequality follows as we note

$$S_{D,\mathcal{C}_{A\leftrightarrow B}}^{\mathrm{det}}\left(\rho\right) = S_{D,\mathcal{C}_{A\leftrightarrow B}}^{\mathrm{det}}\left(\rho^{\Gamma}\right) \geq S_D^{\mathrm{det}}\left(\rho^{\Gamma}\right) = S_D\left(\rho^{\Gamma}\right),$$

where the first equality is due to Proposition 35. □

With these initial observations in place, we proceed to identify an example of a cqq-state exhibiting another gap between the rate of deterministic secure state distillation with respect to an $\mathcal{C}_{A\leftrightarrow B}$-restricted eavesdropper and an unrestricted eavesdropper.

Example 37. Let $\sigma \in \mathcal{D}\left(XAB\right)$ be the cqq-state given by

$$\sigma = \frac{1}{2}\left|0\right\rangle\!\left\langle 0\right| \otimes \sigma_0 + \frac{1}{2}\left|1\right\rangle\!\left\langle 1\right| \otimes \frac{1}{2}\left(\sigma_0 + \sigma_1\right),$$

where σ_0, σ_1 denote the extremal Werner states with local dimension $d \in \mathbb{N}$. A straightforward calculation shows

$$\mathrm{H}\left(X|AB\right)_\sigma = -\frac{1}{2}\log\frac{1}{2} - \frac{1}{4}\log\frac{1}{4} + \left(\frac{1}{2}+\frac{1}{4}\right)\log\left(\frac{1}{2}+\frac{1}{4}\right) = 1 + \frac{3}{4}\log\frac{3}{4} \approx 0.69.$$

If we denote by $\left|\varphi_d\right\rangle = \frac{1}{\sqrt{d}}\sum_{i=1}^{d}\left|ii\right\rangle$, then

$$\begin{aligned}\frac{1}{2}\left(\sigma_0^\Gamma + \sigma_1^\Gamma\right) &= \frac{1}{2d\left(d+1\right)}\left(\mathbb{1} + d\varphi_d\right) + \frac{1}{2d\left(d-1\right)}\left(\mathbb{1} - d\varphi_d\right)\\ &= \frac{1}{\left(d-1\right)\left(d+1\right)}\left(\mathbb{1} - \varphi_d\right)\\ &\geq 0,\end{aligned}$$

which shows $\sigma^\Gamma \geq 0$. Furthermore, we have

$$\begin{aligned}\frac{1}{2}\left\|\frac{1}{2}\sigma_0^\Gamma - \frac{1}{4}\left(\sigma_0^\Gamma + \sigma_1^\Gamma\right)\right\|_1 &= \frac{1}{4}\left\|\left(\frac{1}{d\left(d+1\right)} - \frac{1}{\left(d-1\right)\left(d+1\right)}\right)\left(\mathbb{1} - \varphi_d\right) + \frac{1}{d}\varphi_d\right\|_1\\ &= \frac{1}{2d},\end{aligned}$$

so it follows from the operational interpretation of the conditional min-entropy [22] and the Holevo-Helstrom Theorem [15, 26] that

$$2^{-\mathrm{H}_{\min}\left(X|AB\right)_{\sigma^\Gamma}} = \frac{1}{2} + \frac{1}{2}\left\|\frac{1}{2}\sigma_0^\Gamma - \frac{1}{4}\left(\sigma_0^\Gamma + \sigma_1^\Gamma\right)\right\|_1 = \frac{1}{2}\left(1 + \frac{1}{d}\right).$$

This lower bounds the rate of deterministic secure state distillation from σ, since

$$S_{D,\mathcal{C}_{A\leftrightarrow B}}^{\mathrm{det}}\left(\sigma\right) \geq S_D\left(\sigma^\Gamma\right) = \mathrm{H}\left(X|AB\right)_{\sigma^\Gamma} \geq \mathrm{H}_{\min}\left(X|AB\right)_{\sigma^\Gamma} = 1 - \log\left(1 + \frac{1}{d}\right).$$

For sufficiently large $d \in \mathbb{N}$, this constitutes an example, where the rate of deterministic secure state distillation with respect to an $\mathcal{C}_{A\leftrightarrow B}$-restricted eavesdropper is notably larger than the rate of secure state distillation.

2.3 Hiding States and the Rate of Hiding State Distillation

Along the lines of the preceding sections, we consider a two-party setting involving Xavier and Eve described by a cq-state $\rho^{XE} \in \mathcal{D}\left(XE\right)$. In the following, we consider a slightly more involved task of Xavier, namely, that of hiding data in the state of Eve's system E.

Here, Xavier has to ensure that the state of his system is retrievable by measurement of Eve's system, if she has perfect quantum memory, while an eavesdropper with imperfect quantum memory has little to no knowledge about the state of Xavier's system. In the following, we discuss to what extent Xavier is able to exploit Eve's limitation to a certain set of channels in order to extract hidden data.

Definition 38. Consider a cq-state

$$\rho^{XE} = \sum_{x\in\Sigma} p_x \left|x\right\rangle\!\left\langle x\right|^X \otimes \rho_x^E.$$

Let $\varepsilon > 0$. We say ρ is *ε-correct*, if

$$\Pr_{\text{guess}}\left(X|E\right)_\rho \geq 1-\varepsilon, \tag{2.11}$$

and we say that ρ is *ε-secure with respect to $\mathcal{C}\left(E\right)$*, if

$$\Delta_{\mathcal{C}(E)}\left(X|E\right)_\rho < \varepsilon. \tag{2.12}$$

Finally, if ρ is ε-correct and ε-secure with respect to $\mathcal{C}\left(E\right)$, then ρ is a $\left(\log\left|\Sigma\right|, \varepsilon\right)$*-hiding state with respect to $\mathcal{C}\left(E\right)$*.

Definition 39. Let $\rho \in \mathcal{D}\left(XE\right)$ be a cq-state, and let $\mathcal{C} = \left(\mathcal{C}_n\left(E^n\right)\right)_{n\in\mathbb{N}}$ be a sequence of sets of channels. Let $r \geq 0$. We say that r is an *achievable rate of hiding state distillation with respect to a $\mathcal{C}$-restricted eavesdropper*, if $r = 0$ or the following condition holds: For sufficiently large $n \in \mathbb{N}$ and $m = \lfloor rn \rfloor$ there exists a function $s\colon \Sigma \to \mathcal{T}$ for some alphabet $\mathcal{T}$ and a family $\mathcal{F}$ of functions $f\colon \Sigma^n \to \Sigma'$, where Σ' is of size $\left|\Sigma'\right| = 2^m$, with associated probability distribution $\left(p_f\right)_{f\in\mathcal{F}}$, such that if we denote by

$$\rho_s^{\otimes n} = \sum_{x\in\Sigma^n} p_x \left|x\right\rangle\!\left\langle x\right| \otimes \rho_x^E \otimes \left|s\left(x\right)\right\rangle\!\left\langle s\left(x\right)\right|^T,$$

then

$$\mathcal{F}\left(\rho_s^{\otimes n}\right) = \sum_{x'\in\Sigma'} \left|x'\right\rangle\!\left\langle x'\right|^{X'} \otimes \sum_{f\in\mathcal{F}} p_f \left|f\right\rangle\!\left\langle f\right|^F \otimes \sum_{x\in f^{-1}(\{x'\})} p_x \rho_x^{E^n} \otimes \left|s\left(x\right)\right\rangle\!\left\langle s\left(x\right)\right|^T$$

is a $\left(\log\left|\Sigma'\right|, \varepsilon\right)$-hiding state with respect to a $\mathcal{C}_n\left(E^n\right)$-restricted eavesdropper on E^nFT.

The *the hiding rate with respect to a $\mathcal{C}$-restricted eavesdropper $H_{D,\mathcal{C}}\left(\rho\right)$* is the supremum of all achievable rates of hiding state distillation with respect to a $\mathcal{C}$-restricted eavesdropper.

With the definition of hiding rate in place, we proceed to prove a statement analogous to Theorem 21, namely, that under general assumptions on Eve's quantum memory, we can place bounds on the hiding rate in terms of entropic quantities.

Theorem 40. Let Σ be an alphabet and consider a cq-state $\rho^{XE} \in \mathcal{D}\left(XE\right)$. Let $\mathcal{C} = \mathcal{C}_n\left(E^n\right)$ be a sequence of sets of channels satisfying

$$\mathcal{C}_m\left(E^m\right) \otimes \mathcal{C}_n\left(E^n\right) \subseteq \mathcal{C}_{m+n}\left(E^{m+n}\right)$$

for all $m, n \in \mathbb{N}$. Then

$$\widetilde{\mathrm{H}}_{\min,\mathcal{C}}^{\infty}\left(X|E\right)_\rho - \mathrm{H}\left(X|E\right)_\rho \leq H_{D,\mathcal{C}}\left(\rho\right) \leq \mathrm{H}_{\mathcal{C}}^{\infty}\left(X|E\right)_\rho - \mathrm{H}\left(X|E\right)_\rho$$

Proof. Let $\varepsilon > 0$, consider $n \in \mathbb{N}$ and let $k \in \mathbb{N}$ be given by $k = \left\lceil n \operatorname{H}(X|E)_\rho \right\rceil$. For all but finitely many n there exists a function $s\colon \Sigma^n \to \{0,1\}^k$ [32] such that

$$\Pr_{\mathrm{guess}}\left(X^n|E^nT\right)_{\rho_s^{\otimes n}} \geq 1-\varepsilon,$$

which we will use again later. Now, let $\delta > 0$ and consider $m \in \mathbb{N}$ given by $m = \left\lfloor n\left(\widetilde{\operatorname{H}}^{\infty}_{\min,\mathcal{C}}(X|E)_\rho - \operatorname{H}(X|E)_\rho - \delta\right)\right\rfloor$. For any two-universal family $\mathcal{F}$ of functions $f\colon \Sigma^n \to \Sigma'$, where $\Sigma' = \{0,1\}^m$, it follows from the adapted leftover hash lemma (see Lemma 12) that

$$\begin{aligned}\Delta_{\mathcal{C}_n(E^n)}\left(X'\middle|E^nFT\right)_{\mathcal{F}\left(\rho_s^{\otimes n}\right)} &\leq 2\varepsilon + \sqrt{2^{m-\operatorname{H}^{\varepsilon}_{\min,\mathcal{C}_n(E^n)}(X^n|E^nT)_{\rho_s^{\otimes n}}}} \\ &\leq 2\varepsilon + \sqrt{2^{m-\operatorname{H}^{\varepsilon}_{\min,\mathcal{C}_n(E^n)}(X^n|E^n)_{\rho^{\otimes n}}+k}} \\ &\leq 2\varepsilon + \sqrt{2^{n\left(\widetilde{\operatorname{H}}^{\infty}_{\min,\mathcal{C}}(X|E)_\rho - \frac{1}{n}\operatorname{H}^{\varepsilon}_{\min,\mathcal{C}_n(E^n)}(X^n|E^n)_{\rho^{\otimes n}} - \delta\right)+1}}, \end{aligned} \tag{2.13}$$

where the second inequality is due to the adapted chain rule (see Lemma 11), and the last inequality follows from our choice of $k, m \in \mathbb{N}$. Furthermore, we have by definition that

$$\frac{1}{n}\operatorname{H}^{\varepsilon}_{\min,\mathcal{C}_n(E^n)}\left(X^n|E^n\right)_{\rho^{\otimes n}} \geq \widetilde{\operatorname{H}}^{\infty}_{\min,\mathcal{C}}(X|E)_\rho - \delta$$

for sufficiently large $n \in \mathbb{N}$. This implies the second term in (2.13) tends to 0 as n tends to infinity. Finally, we also have

$$\Pr_{\mathrm{guess}}\left(X'\middle|E^nFT\right)_{\mathcal{F}\left(\rho_s^{\otimes n}\right)} \geq \Pr_{\mathrm{guess}}\left(X^n|E^nT\right)_{\rho_s^{\otimes n}} \geq 1-\varepsilon,$$

which proves $\widetilde{\operatorname{H}}^{\infty}_{\min,\mathcal{C}}(X|E)_\rho - \operatorname{H}(X|E)_\rho - \delta$ is an achievable rate of hiding state distillation with respect to a $\mathcal{C}$-restrict eavesdropper for all $\delta > 0$, and thus achievability of the lower bound in the statement.

For the converse inequality, let $\varepsilon > 0$ and $n \in \mathbb{N}$, and denote by Σ', $\mathcal{T}$ two alphabets. Suppose there exists a family $\mathcal{F}$ of functions $f\colon \Sigma \to \Sigma'$ with corresponding probability distribution $(p_f)_{f\in\mathcal{F}}$ and a function $s\colon \Sigma^n \to \mathcal{T}$, such that

$$\Pr_{\mathrm{guess}}\left(X'\middle|E^nFT\right)_{\mathcal{F}\left(\rho_s^{\otimes n}\right)} \geq 1-\varepsilon,$$

and

$$\Delta_{\mathcal{C}_n(E^n)}\left(X'\middle|E^nFT\right)_{\mathcal{F}\left(\rho_s^{\otimes n}\right)} < \varepsilon.$$

Let $\Lambda \in \mathcal{C}_n\left(E^n\rangle E_n\right)$ and note that by an argument identical to the one given in the proof of Theorem 21, we have

$$m \leq \sum_{f\in\mathcal{F}} p_f \operatorname{H}\left(X'\middle|E_nT\right)_{\Lambda\left(f\left(\rho_s^{\otimes n}\right)\right)} + 3\sqrt{\varepsilon}\log d_{X'} + g\left(\sqrt{\varepsilon}\right).$$

To obtain the desired result, we upper bound the terms in the sum in the expression above, that is

$$\begin{aligned}\mathrm{H}\left(X'\middle|E_nT\right)_{\Lambda\left(f\left(\rho_s^{\otimes n}\right)\right)} &= \mathrm{H}\left(X'T\middle|E_n\right)_{\Lambda\left(f\left(\rho_s^{\otimes n}\right)\right)} - \mathrm{H}\left(T|E_n\right)_{\Lambda\left(f\left(\rho_s^{\otimes n}\right)\right)}\\ &\leq \mathrm{H}\left(X'T\middle|E_n\right)_{\Lambda\left(f\left(\rho_s^{\otimes n}\right)\right)} - \mathrm{H}\left(T|E^n\right)_{f\left(\rho_s^{\otimes n}\right)}\\ &= \mathrm{H}\left(X'T\middle|E_n\right)_{\Lambda\left(f\left(\rho_s^{\otimes n}\right)\right)} - \mathrm{H}\left(X'T\middle|E^n\right)_{f\left(\rho_s^{\otimes n}\right)}\\ &\quad + \mathrm{H}\left(X'\middle|E^nT\right)_{f\left(\rho_s^{\otimes n}\right)} \end{aligned} \tag{2.14}$$

The expected value of the last term in (2.14) when sampling $f \in \mathcal{F}$ according to the associated probability distribution $\left(p_f\right)_{f\in\mathcal{F}}$ is upper bounded as follows

$$\sum_{f\in\mathcal{F}} p_f\, \mathrm{H}\left(X'\middle|E^nT\right)_{f\left(\rho_s^{\otimes n}\right)} = \mathrm{H}\left(X'\middle|E^nTF\right)_{\mathcal{F}\left(\rho_s^{\otimes n}\right)} \leq \varepsilon \log|\Sigma'| + h\left(\varepsilon\right),$$

where the inequality follows from Fano's inequality (see Lemma 84 in Appendix A). Next, consider the first two terms in (2.14), and note

$$\begin{aligned}\mathrm{H}\left(X'T\middle|E_n\right) - \mathrm{H}\left(X'T\middle|E^n\right) &= \mathrm{H}\left(X^n|E_n\right) - \mathrm{H}\left(X^n\middle|X'E_nT\right)\\ &\quad - \mathrm{H}\left(X^n|E^n\right) + \mathrm{H}\left(X^n\middle|X'E^nT\right)\\ &\leq \mathrm{H}\left(X^n|E_n\right) - \mathrm{H}\left(X^n|E^n\right),\end{aligned}$$

where the equality follows from the observation that $\mathrm{H}\left(X^nX'T\right) = \mathrm{H}\left(X^n\right)$. Finally, if we consider the expected value in (2.14) when sampling $f \in \mathcal{F}$ according to the associated probability distribution $\left(p_f\right)_{f\in\mathcal{F}}$, we obtain

$$\begin{aligned}\sum_{f\in\mathcal{F}} p_f\, \mathrm{H}\left(X'\middle|E_nT\right)_{\Lambda\left(f\left(\rho_s^{\otimes n}\right)\right)} &\leq \mathrm{H}\left(X^n|E_n\right)_{\Lambda\left(\rho^{\otimes n}\right)} - \mathrm{H}\left(X^n|E^n\right)_{\rho^{\otimes n}}\\ &\quad + \left(\varepsilon + 3\sqrt{\varepsilon}\right)\log\left|\Sigma'\right| + g\left(\sqrt{\varepsilon}\right) + h\left(\varepsilon\right).\end{aligned}$$

If we take infimum of the expression over over all $\Lambda \in \mathcal{C}_n\left(E^n\right)$, it follows that

$$m \leq \mathrm{H}_{\mathcal{C}_n\left(E^n\right)}\left(X^n|E^n\right)_{\rho^{\otimes n}} - \mathrm{H}\left(X^n|E^n\right)_{\rho^{\otimes n}} + \left(\varepsilon + 3\sqrt{\varepsilon}\right)\log\left|\Sigma'\right| + g\left(\sqrt{\varepsilon}\right) + h\left(\varepsilon\right),$$

so dividing both sides by n and taking the limit as $n \to \infty$, it follows that we cannot achieve a rate of hiding state distillation with respect to a $\mathcal{C}$-restricted eavesdropper exceeding $\mathrm{H}_{\mathcal{C}}^{\infty}\left(X|E\right)_\rho - \mathrm{H}\left(X|E\right)_\rho$. □

The identification of a significant gap between the rate of secure state distillation with respect to a restricted and an unrestricted eavesdropper in Subsection 2.2.1 is key to emphasize the relevance of the concept of a hiding rate: This shows that it is indeed possible to achieve non-zero hiding rates! From this, it is natural to conjecture a relation between the rate of secure state distillation and the hiding rate as follows.

Conjecture 41. Consider a cq-state $\rho^{XE} \in \mathcal{D}\left(XE\right)$, and let $\mathcal{C} = \mathcal{C}_n\left(E^n\right)$ be a sequence of sets of channels. Then

$$H_{D,\mathcal{C}}\left(\rho\right) = S_{D,\mathcal{C}}\left(\rho\right) - S_D\left(\rho\right).$$

Remark. Note that this is a direct consequence of Theorem 21 and Theorem 40 if we assume the statement in Conjecture 10 to be true and add the assumption that the sequence of sets of channels satisfy

$$\mathcal{C}_m\left(E^m\right) \otimes \mathcal{C}_n\left(E^n\right) \subseteq \mathcal{C}_{m+n}\left(E^{m+n}\right)$$

for all $m, n \in \mathbb{N}$.

Chapter 3

On Privacy, Entanglement and Data Hiding

In this chapter, we consider a bipartite scenario consisting of two parties Alice and Bob with a joint quantum system AB in some state $\rho \in \mathcal{D}(AB)$. We will primarily be concerned with a communication setup allowing one-way LOCC operations, that is, we allow Alice and Bob to perform arbitrary local operations on their individual systems, and furthermore Alice is allowed to communicate classically to Bob as depicted below.

$$A \xrightarrow[\text{cl. comm.}]{\rho} B$$

Having restricted Alice and Bob to one-way LOCC protocols and provided them with a shared resource state $\rho \in \mathcal{D}(AB)$, we will consider three communication tasks, namely, the distillation of 1) maximally entangled states, 2) private states and 3) phase hiding states. Maximally entangled states are special cases of private states, so it follows that the task of entanglement distillation is at least as hard as private state distillation. In fact, it was shown [6] that there are states ρ from which no entanglement can be distilled, while private state distillation remains possible at a non-zero rate. We introduce the notion of phase hiding states in an effort to address the gap between private state distillation and entanglement distillation.

In the first section, we introduce terminology and notation particularly relevant to our discussion of private state and entanglement distillation. Furthermore, we prove a lower bound on the rate of private state distillation analogous to a previously shown lower bound on the distillable entanglement [1]. These bounds provide an indication of a connection from the notion of hiding states discussed in Chapter 2 to the gap between distillable private key and distillable entanglement. Finally, we employ our lower bound on the distillable private key in the context of a quantum key repeater to show an elementary lower bound, which gives rise to an interesting example.

Taking a step in the direction of generality, we introduce a notion of locally private states and prove various statements analogous to results regarding private states [33]. We are able to connect this notion of local privacy to the task of encoding data in the state of a quantum system by applying Z-gates. With this in mind, we introduce the notion of hiding states in an effort to describe a desired resource, which can be distilled whenever a gap between the distillable key and distillable entanglement is present. Finally, we provide elementary bounds on the rate at which hiding states can be distilled.

3.1 Private States and Key-Correlated States

In this section, we introduce notation and terminology particularly relevant to the study of correlation, privacy, and entanglement. Many of the core concepts are from [1]. Consider two spatially separated parties Alice and Bob each with a 2-dimensional quantum system A_k and B_k, respectively, and suppose that the state of the joint system is a maximally entangled bit $\varphi_{xi} \in \mathcal{D}(A_k B_k)$ for $x, i \in \{0,1\}$, that is,

$$|\varphi\rangle = \frac{1}{\sqrt{2}}\left(|00\rangle + |11\rangle\right), \qquad |\varphi_{xi}\rangle = X^x_{B_k} Z^i_{B_k} |\varphi\rangle$$

with corresponding density operators $\varphi_{xi} = |\varphi_{xi}\rangle\langle\varphi_{xi}| \in \mathcal{D}(A_k B_k)$. We refer to the systems A_k and B_k as the *key systems*. If Alice and Bob measure φ_{xi} in the computational basis, they will produce two perfectly correlated or anti-correlated, unbiased bits, and we shall refer to the post-measurement state in this scenario as the *key attacked state* $\hat{\varphi}_i$ given by

$$\begin{aligned}\hat{\varphi}_0 &\coloneqq \frac{1}{2}\left(|00\rangle\langle 00| + |11\rangle\langle 11|\right) = \frac{1}{2}\left(\varphi_{00} + \varphi_{01}\right),\\ \hat{\varphi}_1 &\coloneqq \frac{1}{2}\left(|01\rangle\langle 01| + |10\rangle\langle 10|\right) = \frac{1}{2}\left(\varphi_{10} + \varphi_{11}\right)\end{aligned}$$

where the support of $\hat{\varphi}_0$ is referred to as the *maximally correlated subspace*. More generally, we define for $m \in \mathbb{N}$ and $x, i \in \{0,1\}^m$ an m-bit maximally entangled state by

$$|\varphi_{xi}\rangle = |\varphi_{x_1 i_1}\rangle \otimes |\varphi_{x_2 i_2}\rangle \otimes \ldots \otimes |\varphi_{x_n i_n}\rangle$$

with corresponding density operators $\varphi_{xi} = |\varphi_{xi}\rangle\langle\varphi_{xi}|$. When the dimension is clear from the context we shall write $\varphi = \varphi_{0^m 0^m}$.

We define secrecy with respect to an eavesdropper Eve as follows [33]. Suppose Alice and Bob share a bipartite quantum state $\rho \in \mathcal{D}(A_k B_k A_s B_s)$, where A_s and B_s are referred to as *shield systems*. Furthermore, suppose Eve holds a purification of ρ. If Eve's system is uncorrelated with the key systems A_k and B_k subsequent to the measurement in the computational basis, we say Alice and Bob have obtained secret key. As an example, we may note that if Alice and Bob share the maximally entangled state φ and measure it in the computational basis, they will produce a perfectly secret key with respect to an eavesdropper as φ is a pure state.

A 1*-bit private state* $\gamma \in \mathcal{D}(A_k A_s B_k B_s)$ [6] is a state with the property that upon measurement of the key systems A_k and B_k in the computational basis, Alice and Bob achieve a shared secret bit. Trivially, this notion generalizes the notion of a maximally entangled state as remarked upon in the example above. It was shown [6] that γ is an *m-bit private state*, if and only if it can be written as

$$\gamma = \gamma_{0^m 0^m} \coloneqq U\left(\varphi_{0^m 0^m} \otimes \sigma\right) U^\dagger, \qquad U = \sum_{i=\{0,1\}^m} |ii\rangle\langle ii|^{A_k B_k} \otimes U_i^{A_s B_s} \tag{3.1}$$

where $\sigma \in \mathcal{D}(A_sB_s)$ and $U_i \in \mathrm{U}(A_sB_s)$ for $i = \{0,1\}^m$ are unitary transformations. Furthermore, we denote by γ_{xi} for $x, i \in \{0,1\}^m$ the state

$$\gamma_{xi} := \left(\mathcal{X}_{B_k}^x \circ \mathcal{Z}_{B_k}^i\right)(\gamma_{0^m0^m}) = U\varphi_{xi} \otimes \sigma U^\dagger.$$

We refer to γ_{xi} as an x-bit and i-phase flipped private state. For ease of notation, we will on occasion refer to Alice's and Bob's systems by the single letter abbreviations $A = A_kA_s$, $B = B_kB_s$.

Let $\rho \in \mathcal{D}(AB)$ be an arbitrary state. We may consider ρ as a resource in terms of Alice and Bob trying to achieve various communication tasks, say, distilling maximally entangled states or distilling private states. When Alice and Bob are spatially separated and restricted to classical communication, it is not necessarily possible to achieve exact copies of maximally entangled states, nor is it necessarily possible to achieve perfectly secret, shared bits. However, in the asymptotic setting, where Alice and Bob share $\rho^{\otimes n}$ for $n \in \mathbb{N}$, they may achieve ε-approximate m_e-bit entangled states or m_p-bit private states at fixed rates. To this end, we formally define

$$E_D(\rho) := \lim_{\varepsilon\to 0} \limsup_{n\to\infty} \sup_{m\in\mathbb{N}_0} \left\{ \frac{m}{n} \,\middle|\, \exists \Lambda \in \mathcal{C}_{A\leftrightarrow B}(A^n : B^n) : \Lambda\left(\rho^{\otimes n}\right) \approx_\varepsilon \varphi_{0^m0^m} \right\},$$
$$K_D(\rho) := \lim_{\varepsilon\to 0} \limsup_{n\to\infty} \sup_{m\in\mathbb{N}_0} \left\{ \frac{m}{n} \,\middle|\, \exists \gamma_{0^m0^m}, \Lambda \in \mathcal{C}_{A\leftrightarrow B}(A^n : B^n) : \Lambda\left(\rho^{\otimes n}\right) \approx_\varepsilon \gamma_{0^m0^m} \right\},$$

and we define the corresponding quantities $E_D^{\rightarrow}(\rho)$, $K_D^{\rightarrow}(\rho)$, where Alice and Bob are restricted to one-way communication from Alice to Bob, analogously. Trivially, we have

$$K_D(\rho) \geq E_D(\rho), \qquad K_D(\rho) \geq K_D^{\rightarrow}(\rho), \qquad \text{and} \qquad E_D(\rho) \geq E_D^{\rightarrow}(\rho).$$

We will use the terms rate of distillable private key and rate of private state distillation interchangeably. Furthermore, as a measure of entanglement, we will additionally consider the relative entropy of entanglement E_R given by

$$E_R(\rho) := \inf_{\sigma\in\mathcal{D}_{\mathrm{sep}}(A:B)} D(\rho\|\sigma), \qquad \rho \in \mathcal{D}(A:B),$$

with a regularized relative entropy of entanglement given by $E_R^\infty(\rho) := \lim_{n\to\infty} \frac{1}{n} E_R(\rho^{\otimes n})$. It has been shown to be an upper bound on the rate of distillable key [6], that is, $E_R(\rho) \geq E_R^\infty(\rho) \geq K_D(\rho)$.

As we will be particularly concerned with the phenomenon of $K_D(\rho)$ being strictly larger than $E_D(\rho)$, we take note of the following example from [6].

Example 42. Let $\gamma \in \mathcal{D}(A_kB_kA_sB_s)$ be given by

$$\gamma^{A_kB_kA_sB_s} = \frac{1}{2}\varphi_{00}^{A_kB_k} \otimes \left(1 + \frac{1}{d}\right)\sigma_0^{A_sB_s} + \frac{1}{2}\varphi_{01}^{A_kB_k} \otimes \left(1 - \frac{1}{d}\right)\sigma_1^{A_sB_s},$$

where σ_0, σ_1 denote the extremal Werner states of local dimension $d \in \mathbb{N}$. The orthogonality of the extremal Werner states ensures that γ is a private state, and applying the log-negativity as an upper bound the distillable entanglement yields

$$K_D(\rho) \geq K_D^{\rightarrow}(\rho) \geq 1 \gg \log\left(1 + \frac{1}{d}\right) \geq E_D(\rho) \geq E_D^{\rightarrow}(\rho).$$

This is shown in more detail in [6].

Inspired by the example above, we shall refer to states $\gamma \in \mathcal{D}\left(A_k B_k A_s B_s\right)$ given by

$$\gamma = \sum_{\alpha \in \{0,1\}^m} p_\alpha \varphi_{0^m \alpha} \otimes \rho_\alpha, \qquad 0 \leq p_\alpha, \qquad \sum_{\alpha \in \{0,1\}^m} p_\alpha = 1$$

where $\rho_\alpha \perp \rho_\beta$ for all distinct $\alpha, \beta \in \{0,1\}^m$ as *m-bit Bell private states* [1], a description justified by the observation that γ is a private state, if $\rho_\alpha \perp \rho_\beta$ for all distinct $\alpha, \beta \in \{0,1\}^m$ [1].

Taking another step in the direction of generality, we may consider all states with key systems supported on the maximally correlated subspace. For $m \in \mathbb{N}$ we note that $\{|\varphi_{0^m\alpha}\rangle\langle\varphi_{0^m\beta}|\}_{\alpha,\beta \in \{0,1\}^m}$ constitutes a basis of the 2^m-levelled maximally correlated subspace of $A_k B_k$, and so we may define *m-bit key-correlated states* $\rho \in \mathcal{D}\left(A_k B_k A_s B_s\right)$ by

$$\rho = \sum_{\alpha,\beta \in \{0,1\}^m} |\varphi_{0^m\alpha}\rangle\langle\varphi_{0^m\beta}| \otimes \sigma_{\alpha\beta}, \qquad \rho_{\alpha\beta} \in \mathrm{L}\left(A_s B_s\right)$$

and whenever $\sigma_{\alpha\beta} = 0$ for all distinct $\alpha, \beta \in \{0,1\}^m$ we refer to ρ as a *Bell key-correlated state*. Finally, for any key-correlated state, we shall consider the associated *Bell key-correlated state* given by

$$\rho_{qq} := \frac{1}{2^m} \sum_{\beta \in \{0,1\}^m} \varphi_{0^m\beta} \otimes \rho^{0^m\beta}, \qquad \rho^{0^m\beta} = \mathcal{Z}^{\beta}_{B_k}\left(\rho\right).$$

As we saw in Lemma 2 there exists a reversible quantum channel $\mathcal{E}_{\text{Bell}} \in \mathcal{C}_{A \to B}\left(A_k : B_k\right)$ mapping key-correlated states to Bell key-correlated states, more precisely we have

$$\rho_{qq} = \mathcal{E}_{\text{Bell}}\left(\rho\right).$$

We finish the introduction of relevant notation and terminology by the following definition, which generalizes previously introduced notation.

Definition 43. Let $\rho \in \mathcal{D}\left(A_k B_k A_s B_s\right)$, suppose $d_{A_k}, d_{B_k} = 2^m$ for $m \in \mathbb{N}$ and write out ρ as

$$\rho = \sum_{x,y \in \{0,1\}^m} \underbrace{\sum_{i,j \in \{0,1\}^m} |\varphi_{xi}\rangle\langle\varphi_{yj}| \otimes \rho_{xiyj}}_{\rho_{xy}}, \qquad \rho_{xiyj} \in \mathrm{L}\left(A_s B_s\right).$$

If we denote by $\rho^{\alpha\beta} := \left(\mathcal{Z}^{\alpha}_{A_k} \otimes \mathcal{Z}^{\beta}_{B_k}\right)(\rho)$, then we may define

$$\rho_{qq} := \mathcal{E}_{\text{Bell}}\left(\rho\right) \in \mathcal{D}\left(A_k^* B_k^* A_k B_k A_s B_s\right)$$

where A_k^* and B_k^* are 2^m-dimensional quantum systems. Furthermore, we define

$$\rho_{cq} := \frac{1}{2^m} \sum_{\alpha \in \{0,1\}^m} |\alpha\rangle\langle\alpha|^X \otimes \overline{\rho}^{\alpha 0^m} = \frac{1}{2^m} \sum_{\beta \in \{0,1\}^m} |\beta\rangle\langle\beta|^X \otimes \overline{\rho}^{0^m\beta} \in \mathcal{D}\left(X A_k B_k A_s B_s\right),$$

where $\overline{\rho} \in \mathcal{D}\left(A_k B_k A_s B_s\right)$ denotes the *jointly dephased state* given by

$$\overline{\rho} := \overline{\mathcal{Z}}\left(\rho\right) := \frac{1}{2^m} \sum_{\delta \in \{0,1\}^m} \left(\mathcal{Z}^{\delta}_{A_k} \otimes \mathcal{Z}^{\delta}_{B_k}\right)(\rho) \in \mathcal{D}\left(A_k B_k A_s B_s\right),$$

and finally, the *key-attacked state* $\hat{\rho}$ is given by

$$\hat{\rho} := \frac{1}{2^{2m}} \sum_{\alpha,\beta \in \{0,1\}^m} \rho^{\alpha\beta} \in \mathcal{D}\left(A_k B_k A_s B_s\right).$$

3.1.1 A Lower Bound on the Distillable Key of Key-Correlated States

In the following, we determine the rate of private key distillation from a particularly simple protocol consisting of local operations and classical communication. The statement and proof carry reminiscence of a previously shown lower bound on $E_D^{\rightarrow}(\rho)$ [1], and together these two results provide an indication of how to understand the gap between private key distillation and entanglement distillation in terms of quantum data hiding.

Before we proceed to the protocol, we first prove the statement below.

Lemma 44. Let $\rho \in \mathcal{D}(A_kB_kA_sB_s)$ and suppose $d_{A_k}, d_{B_k} = 2^m$ for $m \in \mathbb{N}$. If ρ is an m-bit key-correlated state, then

$$\mathrm{D}(\rho\|\hat{\rho}) = \mathrm{H}(A_kB_kA_sB_s)_{\hat{\rho}} - \mathrm{H}(A_kB_kA_sB_s)_{\rho} = \mathrm{I}(X : A_kB_kA_sB_s)_{\rho_{cq}}.$$

Proof. First, we note that $\mathcal{Z}^{\alpha}_{B_k^*}(\hat{\rho}) = \hat{\rho}$ for all $\alpha \in \{0,1\}^m$, so it follows from unitary invariance of the relative entropy that

$$\begin{aligned}\mathrm{D}(\rho\|\hat{\rho}) &= \frac{1}{2^m}\sum_{\alpha\in\{0,1\}^m} D\left(\mathcal{Z}^{\alpha}_{B_k}(\rho)\middle\|\mathcal{Z}^{\alpha}_{B_k}(\hat{\rho})\right)\\ &= \frac{1}{2^m}\sum_{\alpha\in\{0,1\}^m} D\left(\rho^{0^m\alpha}\middle\|\hat{\rho}\right)\\ &= \frac{1}{2^m}\sum_{\alpha\in\{0,1\}^m} \left(-\operatorname{Tr}\rho^{0^m\alpha}\log\hat{\rho} - \mathrm{H}(A_kB_kA_sB_s)_{\rho^{0^m\alpha}}\right)\\ &= \mathrm{H}(A_kB_kA_sB_s)_{\hat{\rho}} - \mathrm{H}(A_kB_kA_sB_s)_{\rho},\end{aligned}$$

where the last equality follows from unitary invariance of the von Neumann entropy and the observation that $\frac{1}{2^m}\sum_{\alpha\in\{0,1\}^m}\rho^{0^m\alpha} = \hat{\rho}$. Now note that $\hat{\rho} = \operatorname{Tr}_X \rho_{cq}$, and so

$$\begin{aligned}\mathrm{H}(A_kB_kA_sB_s)_{\hat{\rho}} - \mathrm{H}(A_kB_kA_sB_s)_{\rho} &= \mathrm{H}(X)_{\rho_{cq}} + \mathrm{H}(A_kB_kA_sB_s)_{\rho_{cq}} - \mathrm{H}(XA_kB_kA_sB_s)_{\rho_{cq}}\\ &= \mathrm{I}(X : A_kB_kA_sB_s)_{\rho_{cq}},\end{aligned}$$

which proves the desired statement. □

To derive a lower bound on the distillable secret key from a key-correlated state, we will now consider a particular protocol. Suppose Alice and Bob share a key-correlated state and Alice proceeds as follows: She measures her key system in the computational basis, simply keeps her shield system, and subsequently engages in a privacy amplification protocol with Bob in order to obtain a shared secret key. The statement below uses a result by Devetak and Winter [34] to compute the rate of key distillation of this protocol, which turns out to be optimal whenever the key-attacked state $\hat{\rho}$ is separable.

Proposition 45. Let $\rho \in \mathcal{D}(A_kB_kA_sB_s)$ and suppose $d_{A_k}, d_{B_k} = 2^m$ for $m \in \mathbb{N}$. If ρ be an m-bit key-correlated state, then

$$K_D^{\rightarrow}(\rho) \geq D(\rho\|\hat{\rho}),$$

and if $\hat{\rho}$ is separable, then

$$K_D(\rho) = D(\rho\|\hat{\rho}).$$

Proof. First note that $\mathcal{E}_{\text{Bell}}$ is a reversible protocol consisting of local operations and classical communication, which implies

$$K_D^{\rightarrow}(\rho) = K_D^{\rightarrow}(\rho_{qq}), \qquad D(\rho\|\hat{\rho}) = D(\rho_{qq}\|\hat{\rho}_{qq}).$$

Thus we may assume without loss of generality that ρ is an m-bit Bell key-correlated state, that is

$$\rho = \sum_{\alpha\in\{0,1\}^m} p_\alpha \varphi_{0^m\alpha} \otimes \rho_\alpha$$

where $(p_\alpha)_{\alpha\in\{0,1\}^m}$ is a probability distribution. Let E be a purifying system of A_sB_s, and let $|\psi_\alpha\rangle\langle\psi_\alpha| \in \mathcal{D}(A_sB_sE)$ be a purification of ρ_α for each $\alpha \in \{0,1\}^m$. Then

$$|\psi\rangle = \sum_{\alpha\in\{0,1\}^m} \sqrt{p_\alpha}\, |\varphi_{0^m\alpha}\rangle_{A_kB_k} \otimes |\psi_\alpha\rangle_{A_sB_sE} \otimes |\alpha\rangle_F,$$

is a purification of ρ, where $\{|\alpha\rangle\}_{\alpha\in\{0,1\}^m}$ is an orthonormal basis of F. Now suppose Alice engages in a key distillation protocol by measuring her system A_k in the computational basis and simply storing her system A_s. If we consider the key-attacked state $\hat{\psi} \in \mathcal{D}(A_kB_kA_sB_sEF)$, which is given by

$$\hat{\psi} = \frac{1}{2^m} \sum_{x\in\{0,1\}^m} |xx\rangle\langle xx|_{A_kB_k} \otimes \sum_{\alpha,\beta\in\{0,1\}^m} (-1)^{x\cdot(\alpha+\beta)} \sqrt{p_\alpha p_\beta}\, |\psi_\alpha\rangle\langle\psi_\beta|_{A_sB_sE} \otimes |\alpha\rangle\langle\beta|_F,$$

it follows from [34] that

$$K_D^{\rightarrow}(\rho) \geq \mathrm{I}(A_k : B_kB_s)_{\hat{\psi}} - \mathrm{I}(A_k : EF)_{\hat{\psi}} = \mathrm{H}(A_k|EF)_{\hat{\psi}}.$$

To simplify the calculation we consider a purification $|\Psi\rangle\langle\Psi| \in \mathcal{D}(A_kB_kA_sB_sEFG)$ of the key-attacked state, namely,

$$|\Psi\rangle = \frac{1}{\sqrt{2^m}} \sum_{x,\alpha\in\{0,1\}^m} (-1)^{\alpha\cdot x} \sqrt{p_\alpha}\, |xx\rangle^{A_kB_k} \otimes |\psi_\alpha\rangle^{A_sB_sE} \otimes |\alpha\rangle^F \otimes |x\rangle^G,$$

where $\{|x\rangle\}_{x\in\{0,1\}^m}$ is an orthonormal basis of G. A straightforward calculation now shows

$$\begin{aligned} K_D^{\rightarrow}(\rho) \geq \mathrm{H}(A_k|EF)_\Psi &= \mathrm{H}(B_kA_sB_sG)_\Psi - \mathrm{H}(A_kB_kA_sB_sG)_\Psi \\ &= \mathrm{H}(B_kA_sB_s)_\Psi - \mathrm{H}(A_kB_kA_sB_sG)_\Psi. \end{aligned} \tag{3.2}$$

First, we note

$$\Psi^{B_kA_sB_s} = \omega^{B_k} \otimes \sum_{\alpha\in\{0,1\}^m} p_\alpha \rho_\alpha^{A_sB_s} = \rho^{B_kA_sB_s},$$

which implies $\mathrm{H}(B_kA_sB_s)_\Psi = \mathrm{H}(B_kA_sB_s)_\rho = \mathrm{H}(A_kB_kA_sB_s)_{\hat{\rho}}$. Next, denote by $|\Psi'\rangle\langle\Psi'| \in \mathcal{D}(A_kB_kA_sB_sEF)$ the pure state

$$|\Psi'\rangle = \frac{1}{\sqrt{2^m}} \sum_{x,\alpha\in\{0,1\}^m} (-1)^{\alpha\cdot x} \sqrt{p_\alpha}\, |xx\rangle^{A_kB_k} \otimes |\psi_\alpha\rangle^{A_sB_sE} \otimes |\alpha\rangle^F,$$

and note $\mathrm{H}\left(A_kB_kA_sB_sG\right)_{\Psi} = \mathrm{H}\left(A_kB_kA_sB_s\right)_{\Psi'}$. Since we have

$$\begin{aligned}
\Psi'^{A_kB_kA_sB_s} &= \frac{1}{2^m} \sum_{x,y,\alpha\in\{0,1\}^m} p_\alpha \left(-1\right)^{\alpha\cdot(x+y)} |xx\rangle\langle yy|^{A_kB_k} \otimes \rho_\alpha^{A_sB_s} \\
&= \sum_{\alpha\in\{0,1\}^m} p_\alpha \varphi_{0^m\alpha}^{A_kB_k} \otimes \rho_\alpha^{A_sB_s}
\end{aligned}$$

it follows that $\mathrm{H}\left(A_kB_kA_sB_s\right)_{\Psi'} = \mathrm{H}\left(A_kB_kA_sB_s\right)_{\rho}$. Combining these observations with (3.2), we obtain

$$K_D^{\rightarrow}\left(\rho\right) \geq \mathrm{H}\left(A_k|EF\right)_{\Psi} \geq \mathrm{H}\left(A_kB_kA_sB_s\right)_{\hat{\rho}} - \mathrm{H}\left(A_kB_kA_sB_s\right)_{\rho},$$

and so the desired inequality follows from an application of Lemma 44.

To prove equality whenever $\hat{\rho} \in \mathcal{D}_{\mathrm{sep}}\left(A_kB_kA_sB_s\right)$, we note that

$$K_D^{\rightarrow}\left(\rho\right) \leq K_D\left(\rho\right) \leq E_R^{\infty}\left(\rho\right) \leq D\left(\rho\|\hat{\rho}\right),$$

which proves the desired statement. □

We note the striking resemblance between the statement of Proposition 45 and a previously shown lower bound on the distillable entanglement [1], namely,

$$E_D^{\rightarrow}\left(\rho\right) \geq D_{A_kA_s}^{\infty}\left(\rho\|\hat{\rho}\right) := \lim_{n\to\infty} \sup_{\Lambda\in\mathcal{M}_{\mathrm{all}}(A_kA_s)} \frac{1}{n} D\left(\Lambda\left(\rho\right)\|\Lambda\left(\hat{\rho}\right)\right). \tag{3.3}$$

for all key-correlated states $\rho \in \mathcal{D}\left(A_kB_kA_sB_s\right)$ with equality whenever $\hat{\rho}$ is separable. With this in mind, we have the first observation connecting the difference between distillable secret key and distillable entanglement of a state to the notion of hiding with respect to a locally restricted eavesdropper as discussed in Section 2.3.

Proposition 46. Let $\rho \in \mathcal{D}\left(A_kB_kA_sB_s\right)$ and suppose $d_{A_k}, d_{B_k} = 2^m$ for $m \in \mathbb{N}$. If ρ is an m-bit key-correlated state and $\hat{\rho} \in \mathcal{D}_{\mathrm{sep}}\left(A_kA_s : B_kB_s\right)$, then

$$K_D^{\rightarrow}\left(\rho\right) - E_D^{\rightarrow}\left(\rho\right) \geq H_{D,\mathcal{M}(A_kA_s)}\left(\rho_{cq}\right). \tag{3.4}$$

Remark. Contingent on the statement of Conjecture 10 being true, we have equality in the statement above.

Proof. By Proposition 45 and (3.3) we have

$$K_D^{\rightarrow}\left(A_kA_s : B_kB_s\right)_{\rho} - E_D^{\rightarrow}\left(A_kA_s : B_kB_s\right)_{\rho} = D\left(\rho\|\hat{\rho}\right) - D_{A_kA_s}^{\infty}\left(\rho\|\hat{\rho}\right),$$

and by Lemma 44 we thus have

$$\begin{aligned}
&K_D^{\rightarrow}\left(A_kA_s : B_kB_s\right)_{\rho} - E_D^{\rightarrow}\left(A_kA_s : B_kB_s\right)_{\rho} \\
&\qquad = \mathrm{I}\left(X : A_kB_kA_sB_s\right)_{\rho_{cq}} - \mathrm{I}_{A_kA_s}^{\infty}\left(X : A_kB_kA_sB_s\right)_{\rho_{cq}} \\
&\qquad = \mathrm{H}_{A_kA_s}^{\infty}\left(X|A_kB_kA_sB_s\right)_{\rho_{cq}} - \mathrm{H}\left(X|A_kB_kA_sB_s\right)_{\rho_{cq}}.
\end{aligned}$$

Finally, the upper bound from Theorem 21 yields the desired equality. □

We will now explore the gap between distillable secret key and distillable entanglement in terms of hiding classical data in the state of a quantum system. As the example below illustrates, the inequality (3.4) does not generalize to arbitrary states.

Example 47. Let $\mu \in \mathcal{D}\left(A_k B_k A_s B_s\right)$ and suppose $d_{A_k}, d_{B_k} = 2^m$ for $m \in \mathbb{N}$. Suppose $\mu \in \mathcal{D}\left(A_k B_k A_s B_s\right)$ is given by

$$\mu = \frac{1}{4}\left(\varphi_{00} + \varphi_{10}\right) \otimes \sigma_0 + \frac{1}{4}\left(\varphi_{01} + \varphi_{11}\right) \otimes \frac{1}{2}\left(\sigma_0 + \sigma_1\right),$$

where $\sigma_0, \sigma_1 \in \mathcal{D}\left(A_s B_s\right)$ denote the extremal Werner states of local dimension $d \in \mathbb{N}$. By calculations completely analogous to Example 37, we can show

$$H_{D,\mathcal{M}(A_k A_s)}\left(\mu\right) \geq 1 - \log\left(1 + \frac{1}{d}\right) + \frac{3}{4}\log\frac{3}{4} > 0, \qquad d \geq 5$$

while $K_D^{\rightarrow}\left(\mu\right) = E_D^{\rightarrow}\left(\mu\right) = 0$ since $\mu \in \mathcal{D}_{\text{sep}}\left(A_k A_s B_k B_s\right)$.

If we assume the statement of Conjecture 10, the result in Proposition 46 allows us to translate bounds concerning distinguishability problems to bounds on distillation protocols, and vice versa. For one, we could show that the lower bound on the rate of secure state distillation with respect to a PPT-restricted eavesdropper calculated in Corollary 30 is suboptimal.

Example 48. Let $\gamma \in \mathcal{D}\left(A_k B_k A_s B_s\right)$ be given by

$$\gamma = \frac{1}{2}\varphi_{00} \otimes \sigma_0 + \frac{1}{2}\varphi_{01} \otimes \sigma_1,$$

where σ_0, σ_1 denote the extremal Werner states of local dimension $d \in \mathbb{N}$. Also, let $\sigma \in \mathcal{D}\left(X A_s B_s\right)$ be given by

$$\sigma = \frac{1}{2}\left|0\right\rangle\!\left\langle 0\right| \otimes \sigma_0 + \frac{1}{2}\left|1\right\rangle\!\left\langle 1\right| \otimes \sigma_1.$$

As the maximally entangled states φ_{00}, φ_{01} are perfectly distinguishable by local measurements and one-way classical communication, it follows that

$$H_{D,\mathcal{M}(A_k A_s)}\left(\sigma\right) \geq H_{D,\mathcal{M}(A_k A_s)}\left(\gamma_{cq}\right) \overset{!}{=} 1 - E_D^{\rightarrow}\left(\gamma\right),$$

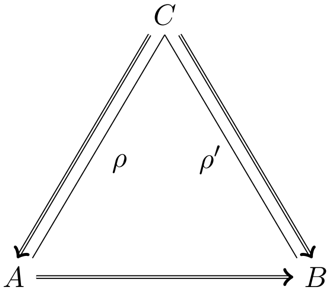
where the equality (!) follows from Proposition 46 contingent on the statement in Conjecture 10 being true. As the log-negativity $E_N\left(\gamma\right)$ is an upper bound on the distillable entanglement $E_D\left(\gamma\right)$ [35], it follows that

$$H_{D,\mathcal{M}(A_k A_s)}\left(\sigma\right) \overset{!}{\geq} 1 - E_N\left(\gamma\right) = 1 - \log\left(1 + \frac{2}{d}\right).$$

With simple techniques (although contingent on Conjecture 10 being true) we have achieved a lower bound on the hiding rate of σ, which is stronger than the statement in Corollary 30, which required substantial effort!

3.1.2 A Lower Bound on the Quantum Key Repeater Rate

In this section, we apply some of our findings in the context of long-distance quantum communication, where two parties, Alice and Bob, are unable to obtain a shared entangled state due to noise as discussed in [31]. Evidently, this prevents them from obtaining private states, and so an additional party, Charlie, is added to the setting to mediate the entanglement. We refer to this setup as a *quantum key repeater*, where Charlie is asked to assist Alice and Bob in obtaining a private state using only local operations and classical communication. More precisely, we consider the situation where Alice and Charlie share a state ρ^{AC}, and Bob and Charlie share a state ${\rho'}^{BC'}$. Charlie may apply an operation jointly to his systems CC' and communicate, say, the result of a measurement to Alice and Bob. Alice and Bob are then subsequently allowed to engage in a one-way LOCC protocol in order to obtain a private state. The setup is depicted below.



The optimal rate at which Charlie can assist Alice and Bob in achieving a private state from ρ, ρ' is called the *one-way quantum key repeater rate*, and we denote it by

$$R_D^{A\leftarrow CC'\rightarrow B}\left(A : CC' : B\right)_{\rho\otimes\rho'} .$$

We focus our attention on key-correlated states, so we may apply our findings in Proposition 45 in this context.

Proposition 49. Let $\rho \in \mathcal{D}\left(A_k C_k A_s C_s\right)$, $\rho' \in \mathcal{D}\left(B_k C'_k B_s C'_s\right)$ be m-bit key-correlated states. Then

$$R_D^{A\leftarrow CC'\rightarrow B}\left(A : CC' : B\right)_{\rho\otimes\rho'} \geq D_{CC'}^{\infty}\left(\overline{\rho\otimes\rho'}\middle\|\hat{\rho}\otimes\hat{\rho}'\right),$$

where

$$\overline{\rho\otimes\rho'} := \frac{1}{2^m}\sum_{\beta\in\{0,1\}^m} \rho^{\beta}\otimes{\rho'}^{\beta}, \qquad \rho^{\beta} = \mathcal{Z}_{C_k}^{\beta}\left(\rho\right), \quad \rho^{\beta} = \mathcal{Z}_{C'_k}^{\beta}\left(\rho'\right).$$

Proof. First apply the reversible LOCC one-way protocol $\mathcal{E}_{\mathrm{Bell}}$ from Corollary 2 [1] to obtain Bell key-correlated states, that is

$$\rho \mapsto \rho_{qq} = \frac{1}{2^m}\sum_{\alpha\in\{0,1\}^m} \varphi_{A_k^* C_k^*}^{0^m\alpha}\otimes\rho^{\alpha},$$

$$\rho' \mapsto \rho'_{qq} = \frac{1}{2^m}\sum_{\beta\in\{0,1\}^m} \varphi_{B_k^* {C'}_k^*}^{0^m\beta}\otimes{\rho'}^{\beta}.$$

If Bob and Charlie engage in a teleportation protocol using the entanglement in system $B_k^* C_k'^*$ to pass the state of system C_k^* to Bob's system B_k^*, the resulting state is given by

$$\xi = \frac{1}{2^m} \sum_{\delta \in \{0,1\}^m} \varphi^{0^m \delta}_{A_k^* B_k^*} \otimes \frac{1}{2^m} \sum_{\alpha+\beta=\delta} \rho^\alpha \otimes \rho'^\beta .$$

Upon measuring Charlie's systems CC' and communicating the measurement outcome to Alice and Bob, this becomes a Bell key-correlated state shared between Alice and Bob. Proposition 45 yields a lower bound on the distillable key of a key-correlated state, which in turn implies

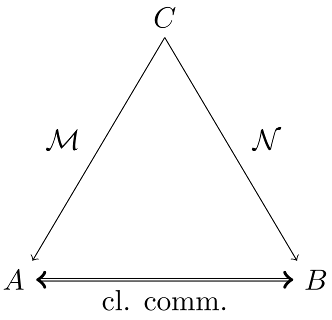
$$R_D^{A \leftarrow CC' \rightarrow B} \left(A : CC' : B \right)_{\rho \otimes \rho'} \geq \frac{1}{n} D_{C^n C'^n} \left(\xi^{\otimes n} \middle\| \hat{\xi}^{\otimes n} \right)$$

for all $n \in \mathbb{N}$. Finally, a straightforward calculation shows that

$$D_{C^n C'^n} \left(\xi^{\otimes n} \middle\| \hat{\xi}^{\otimes n} \right) = D_{C^n C'^n} \left(\overline{\rho \otimes \rho'}^{\otimes n} \middle\| \hat{\rho}^{\otimes n} \otimes \hat{\rho}'^{\otimes n} \right),$$

which proves the desired inequality. □

The statement of Proposition 49 indicates that the local distinguishability of states shared by Alice and Bob considered as one party, and Charlie considered as the other party, plays a role in determining the quantum key repeater rate. It has been shown [31] that even non-distillable entanglement shared between Bob and Charlie can assist Alice, Bob, and Charlie in achieving key repeater rates beyond the rates achievable by the standard quantum repeater protocol based on entanglement swapping. The lower bound in Proposition 49 can be used to show this same lower bound on the quantum key repeater rate in this example and in similar scenarios. The example in [31] was guided by the non-additivity of private capacity [36]. To accommodate this setting we will now consider a slightly different scenario, where Charlie no longer shares entanglement with Alice and Bob, but instead, he has a noisy channel to each of Alice and Bob. In this way, he can supply Alice and Bob with an entangled state, and from there Alice and Bob are allowed to engage in an LOCC protocol. The setup is depicted below.



We will now consider an example of a pair of channels from Charlie to Alice and Bob, where an optimal strategy for achieving private key shared between Alice and Charlie followed by Charlie passing his systems to Bob through a noisy channel is *not* optimal for establishing private key shared between Alice and Bob. This indicates that other resources than private states may be more useful in the quantum key repeater setup.

Example 50. Let A_k, A_s, C and E be quantum systems of dimensions $d_{A_k} = 2$, $d_{A_s} = 2$ $d_C = 3$ and $d_E = 2$, and denote by $U \in \mathrm{L}\left(C, A_k A_s E\right)$ the isometric operator given by

$$\begin{aligned} U \left|0\right\rangle_C &= \frac{1}{\sqrt{2}} \Big(\left|000\right\rangle_{A_k A_s E} + \left|011\right\rangle_{A_k A_s E} \Big) \,, \\ U \left|1\right\rangle_C &= \left|100\right\rangle_{A_k A_s E} \\ U \left|2\right\rangle_C &= \left|101\right\rangle_{A_k A_s E} \,. \end{aligned}$$

If we trace out the purifying system E, this is the platypus channel [37] from Charlie to Alice. Furthermore, let C_k and C_s be quantum systems of dimensions $d_{C_k} = 2$ and $d_{C_s} = 2$, and consider the state

$$\left|\psi\right\rangle = \frac{1}{\sqrt{2}} \left|000\right\rangle_{C C_k C_s} + p \left|110\right\rangle_{C C_k C_s} + q \left|211\right\rangle_{C C_k C_s} \,,$$

where $p^2 + q^2 = \frac{1}{2}$. If Charlie first applies the isometry V to his system C, we obtain the state

$$\left|\Psi\right\rangle = V \left|\psi\right\rangle = \frac{1}{\sqrt{2}} \left|0\right\rangle_{A_k} \left|\varphi\right\rangle_{A_s E} \left|00\right\rangle_{C_k C_s} + p \left|10010\right\rangle_{A_k A_s E C_k C_s} + q \left|10111\right\rangle_{A_k A_s E C_k C_s} \,,$$

where we recall $\left|\varphi\right\rangle = \frac{1}{\sqrt{2}} \left(\left|00\right\rangle + \left|11\right\rangle\right)$. The state $\Psi_{A_k A_s C_k C_s} = \mathrm{Tr}_E \left|\Psi\right\rangle\!\left\langle\Psi\right|$ is a key-correlated state shared by Alice and Charlie, and we may note that

$$\begin{aligned} \hat{\Psi}_{A_k A_s C_k C_s} = \frac{1}{4} \left|00\right\rangle\!\left\langle 00\right|_{A_k A_s} \otimes \left|00\right\rangle\!\left\langle 00\right|_{C_k C_s} + \frac{1}{4} \left|01\right\rangle\!\left\langle 01\right|_{A_k A_s} \otimes \left|00\right\rangle\!\left\langle 00\right|_{C_k C_s} \\ + p^2 \left|10\right\rangle\!\left\langle 10\right|_{A_k A_s} \otimes \left|10\right\rangle\!\left\langle 10\right|_{C_k C_s} + q^2 \left|10\right\rangle\!\left\langle 10\right|_{A_k A_s} \otimes \left|11\right\rangle\!\left\langle 11\right|_{C_k C_s} \,, \end{aligned}$$

which shows $\hat{\Psi} \in \mathcal{D}_{\mathrm{sep}}\left(A_k A_s : C_k C_s\right)$. Due to Proposition 45 we have

$$K_D \left(A_k A_s : C_k C_s\right)_\Psi = D(\Psi \| \hat{\Psi}) = 1 + \eta \left(\frac{1}{2} - q^2\right) + \eta \left(q^2\right) - \eta \left(\frac{3}{4} - q^2\right) - \eta \left(\frac{1}{4} + q^2\right) \,,$$

where we recall $\eta\left(x\right) = -x \log x$. Evidently, the optimal choice of parameter is $q^2 = \frac{1}{4}$, which yields a key rate at $K_D^{\rightarrow}\left(A_k A_s : C_k C_s\right)_\Psi = 1$.

Let B_k, B_s and F be quantum systems of dimension $d_{B_k} = 2$, $d_{B_s} = 2$ and $d_F = 2$. For $\lambda \in [0, 1]$ denote by $V_\lambda \in \mathrm{L}\left(C_s, B_s F\right)$ the isometric operator given by

$$V_\lambda \left|0\right\rangle = \left|00\right\rangle \,, \qquad V_\lambda \left|1\right\rangle = \sqrt{1 - \lambda} \left|10\right\rangle + \sqrt{\lambda} \left|01\right\rangle \,.$$

If we trace out the purifying system F, this is the amplitude damping channel from Charlie to Bob. If Charlie sends his key system through an identity channel and the shield system through an amplitude damping channel, it corresponds to applying the isometry $\mathbb{1} \otimes V_\lambda$ to his system $C_k C_s$, and we obtain the state

$$\begin{aligned} \left|\Psi_\lambda\right\rangle = \frac{1}{\sqrt{2}} \left|00\right\rangle_{A_k B_k} \left|\varphi_1\right\rangle_{A_s E} \left|00\right\rangle_{B_s F} + p \left|11\right\rangle_{A_k B_k} \left|0000\right\rangle_{A_s E B_s F} \\ + q \left|11\right\rangle_{A_k B_k} \left|01\right\rangle^{A_s E} \Big(\sqrt{1 - \lambda} \left|10\right\rangle + \sqrt{\lambda} \left|01\right\rangle \Big)_{B_s F} \,. \end{aligned}$$

Again, $\Psi_\lambda^{A_kA_sB_kB_s} = \mathrm{Tr}_{EF} |\Psi_\lambda\rangle\langle\Psi_\lambda|$ is a separable key-correlated state shared by Alice and Bob, and so it follows from Proposition 45 that

$$K_D^{\rightarrow}(A_kA_s : C_kC_s)_{\Psi_\lambda} = D(\Psi_\lambda \| \hat{\Psi}_\lambda).$$

Through a lengthy yet straightforward computation, we find an analytic expression for $D(\Psi_\lambda\|\hat{\Psi}_\lambda)$, which is shown in Appendix C along with a plot of the rate of distillable key as a function of $q \in \left[0, \frac{1}{\sqrt{2}}\right]$ for $\lambda = 1/2$. We saw that the optimal choice of q in order to achieve a private state between Alice and Charlie is $q = \frac{1}{2}$, but surprisingly the optimal choice of q in order to achieve key between Alice and Bob is ≈ 0.3 when $\lambda = \frac{1}{2}$. This indicates that in the setup with channels from Charlie to Alice and Bob individually, it is *not* optimal for Charlie to obtain a private state with Alice, and afterward pass his systems through a channel to Bob.

3.2 Encoding Data in Phase Orthogonal States

Consider a bipartite quantum system $A_kA_sB_kB_s$, and suppose the state of the quantum system is a private state $\gamma \in \mathcal{D}(A_kB_kA_sB_s)$. In the following, we will take inspiration from the observation that

$$\gamma \perp \left(\mathcal{Z}_{A_k}^{\alpha} \otimes \mathcal{Z}_{B_k}^{\beta}\right)(\gamma), \tag{3.5}$$

for all distinct $\alpha, \beta \in \{0,1\}^m$, which we elaborate on in Example 52 below. Alice and Bob can use this observation to encode classical data $\alpha \in \{0,1\}^m$ into the state of a quantum system by applying phase gates, and due to (3.5) the data is retrievable given global access to the system. We begin by introducing the appropriate generalization of encoding data into the state of a quantum system by applying certain phase gates. With this in place, we proceed to consider the connections between phase orthogonality and so-called local privacy.

Definition 51. Let $\mu \in \mathcal{D}(A_kB_kA_sB_s)$ and suppose $d_{A_k}, d_{B_k} = 2^m$ for $m \in \mathbb{N}$. Denote by

$$\mu^{\alpha\beta} := \left(\mathcal{Z}_{A_k}^{\alpha} \otimes \mathcal{Z}_{B_k}^{\beta}\right)(\mu), \qquad \alpha, \beta \in \{0,1\}^m.$$

We say that μ is an *m-bit joint A_kB_k-phase orthogonal state*, if $\mu^{\alpha\beta} \perp \mu$ for all distinct $\alpha, \beta \in \{0,1\}^m$.

An exceedingly simple example of a 1-bit joint A_kB_k-phase orthogonal state is given by

$$\mu = |{++}\rangle\langle{++}| \in \mathcal{D}(A_kB_k),$$

however, as we shall be particularly interested in the role of phase orthogonality in relation to privacy, we take particular note of the example below.

Example 52. Let $\gamma \in \mathcal{D}(A_kB_kA_sB_s)$ be an m-bit private state. We may write out any private state [6] as

$$\gamma = U(\varphi \otimes \sigma)U^\dagger$$

for some $U \in \mathrm{U}(A_kB_kA_sB_s)$ given by $U = \sum_{i\in\{0,1\}^m} |ii\rangle\langle ii| \otimes U_i$, where $U_i \in \mathrm{U}(A_sB_s)$, and some state $\sigma \in \mathcal{D}(A_sB_s)$. With this characterization in mind, we note that

$$\gamma^{\alpha\beta} = U(\varphi^{\alpha\beta} \otimes \sigma)U^\dagger = U\left(\varphi_{0^m(\alpha+\beta)} \otimes \sigma\right)U^\dagger = \gamma_{0^m(\alpha+\beta)},$$

It follows that $\gamma \perp \gamma^{\alpha\beta}$, if and only if $\varphi \perp \varphi^{\alpha\beta}$, and this holds exactly when $\alpha, \beta \in \{0,1\}^m$ are distinct.

With the appropriate definition of phase orthogonality in place, we proceed to show our first main result of this section concerning the trade-offs between phase orthogonality, correlation, and entanglement.

3.2.1 Phase Orthogonality, Correlation and Entanglement

In the following, we consider the interplay between phase orthogonality, correlation, and entanglement. We have already seen examples of a separable 1-bit joint A_kB_k-phase orthogonal states, namely,

$$\mu = |{+}{+}\rangle\langle{+}{+}| \in \mathcal{D}(A_kB_k)$$

and the state $\mu \in \mathcal{D}(A_kB_kA_sB_s)$ from Example 47. However, we may note that in both cases the state of the joint system A_kB_k of the key-attacked state is maximally mixed, that is, Alice and Bob have no correlation between their measurement outcomes. In Theorem 54 below we show how the combination of phase orthogonality and correlation gives rise to entanglement.

Lemma 53. Let $\mu \in \mathcal{D}(A_kB_kA_sB_s)$ and suppose $d_{A_k}, d_{B_k} = 2^m$ for $m \in \mathbb{N}$. Let μ be given by

$$\mu = \sum_{x,y\in\{0,1\}^m} \underbrace{\sum_{i,j\in\{0,1\}^m} |\varphi_{xi}\rangle\langle\varphi_{yj}| \otimes \mu_{xiyj}}_{\mu_{xy}} .$$

Then μ is joint A_kB_k-phase orthogonal, if and only if μ_{xx} is joint A_kB_k-phase orthogonal for all $x \in \{0,1\}^m$.

Proof. Suppose $\mu \perp \mu^{\alpha\beta}$ for all distinct $\alpha, \beta \in \{0,1\}^m$ and consider

$$\overline{\mu} = \overline{\mathcal{Z}}(\mu) = \sum_{x\in\{0,1\}^m} \underbrace{\sum_{i,j\in\{0,1\}^m} |\varphi_{xi}\rangle\langle\varphi_{xj}| \otimes \mu_{xixj}}_{\mu_{xx}} .$$

For a pair of distinct $\alpha, \beta \in \{0,1\}^m$ we have $\mu^{\delta(\alpha+\delta)} \perp \mu^{\delta'(\beta+\delta')}$ for all $\delta, \delta' \in \{0,1\}^m$, which in turn implies

$$\overline{\mu}^{0^m\alpha} = \frac{1}{2^m} \sum_{\delta\in\{0,1\}^m} \mu^{\delta(\alpha+\delta)} \perp \frac{1}{2^m} \sum_{\delta'\in\{0,1\}^m} \mu^{\delta'(\beta+\delta')} = \overline{\mu}^{0^m\beta},$$

and so each individual term $\mu_{xx}^{0^m\alpha} = \mu_{xx}^{\delta(\alpha+\delta)}$ of the jointly dephased state $\overline{\mu}^{0^m\alpha}$ must be orthogonal to each term $\mu_{x'x'}^{0^m\beta} = \mu_{x'x'}^{\delta'(\beta+\delta')}$ of $\overline{\mu}^{0^m\beta}$ for all distinct $\alpha, \beta \in \{0,1\}^m$.

Conversely, suppose $\mu_{xx}^{0^m\alpha} \perp \mu_{xx}^{0^m\beta}$ for all $x \in \{0,1\}^m$ and distinct $\alpha, \beta \in \{0,1\}^m$. It follows from the definition of μ_{xx} that $\mu_{xx}^{0^m\alpha} \perp \mu_{x'x'}^{0^m\beta}$ for all distinct $x, x' \in \{0,1\}^m$, and so we may infer

$$\overline{\mu}^{0^m\alpha} = \sum_{x\in\{0,1\}^m} \mu_{xx}^{0^m\alpha} \perp \sum_{x'\in\{0,1\}^m} \mu_{x'x'}^{0^m\beta} = \overline{\mu}^{0^m\beta}$$

for all distinct $\alpha, \beta \in \{0,1\}^m$. Now, suppose $\alpha', \beta' \in \{0,1\}^m$ are distinct, and note that the operator $\mu^{\alpha'\beta'}$ is among the terms in the summation

$$\overline{\mu}^{0^m(\alpha'+\beta')} = \frac{1}{2^m} \sum_{\delta \in \{0,1\}^m} \mu^{\delta(\alpha'+\beta'+\delta)},$$

namely, for $\delta = \alpha'$, and similarly $\mu = \mu^{0^m 0^m}$ is among the terms in the summation expression of $\overline{\mu}^{0^m 0^m}$. As $\overline{\mu}^{0^m 0^m} \perp \overline{\mu}^{0^m(\alpha'+\beta')}$ it follows that we also have $\mu \perp \mu^{\alpha'\beta'}$ as desired. □

Theorem 54. *Let $\mu' \in \mathcal{D}(A_k B_k A_s B_s)$ and suppose $d_{A_k}, d_{B_k} = 2^m$ for $m \in \mathbb{N}$. Write out μ as*

$$\mu' = \sum_{x,y \in \{0,1\}^m} \underbrace{\sum_{i,j \in \{0,1\}^m} |\varphi_{xi}\rangle\langle\varphi_{yj}| \otimes \mu'_{xiyj}}_{\mu_{xy}},$$

Let $\varepsilon \geq 0$ and suppose $\mu' \approx_\varepsilon \mu$ for some m-bit joint $A_k B_k$-phase orthogonal state μ. Then

$$E_R(\mu') \geq -\mathrm{H}\left(\{p_x\}_{x \in \{0,1\}^m}\right) + (1 - \sqrt{\varepsilon})\left(\left(1 - 4\sqrt{\varepsilon}\right) m - h\left(\sqrt{\varepsilon}\right)\right),$$

where $p_x = \operatorname{Tr} \mu'_{xx}$ for $x \in \{0,1\}^m$.

Proof. Let $\varepsilon \geq 0$ and suppose $\mu' \in \mathcal{D}(A_k B_k A_s B_s)$ is an ε-approximate m-bit joint $A_k B_k$-phase orthogonal state, that is, $\mu' \approx_\varepsilon \mu$ for some m-bit joint $A_k B_k$-phase orthogonal state $\mu \in \mathcal{D}(A_k B_k A_s B_s)$ given by

$$\mu = \sum_{x,i,y,j \in \{0,1\}^m} |\varphi_{xi}\rangle\langle\varphi_{yj}| \otimes \mu_{xiyj}, \qquad \overline{\mu} = \sum_{x \in \{0,1\}^m} \underbrace{\sum_{i,j \in \{0,1\}^m} |\varphi_{xi}\rangle\langle\varphi_{xj}| \otimes \mu_{xixj}}_{\mu_{xx}}.$$

The joint dephasing channel $\overline{\mathcal{Z}}(\cdot) = \frac{1}{2^m} \sum_{\alpha \in \{0,1\}^m} \mathcal{Z}_{A_k}^{\alpha} \otimes \mathcal{Z}_{B_k}^{\alpha}(\cdot)$ preserves the set of separable states, and so it follows that $E_R(\mu') \geq E_R(\overline{\mu}')$. Let $\sigma \in \mathcal{D}_{\text{sep}}(A_k A_s : B_k B_s)$ and note that $\mathcal{E}_{\text{Bell}} \in \mathcal{C}_{A \to B}(A_k : B_k)$ is reversible, which implies

$$D\left(\overline{\mu}' \middle\| \sigma\right) = D\left(\overline{\mu}'_{qq} \middle\| \mathcal{E}_{\text{Bell}}(\sigma)\right).$$

Now recall that $\overline{\mu}_{qq}$ and $\overline{\mu}'_{qq}$ are given by

$$\overline{\mu}_{qq} = \sum_{x \in \{0,1\}^m} p_x \gamma_{x,x0^m}, \qquad p_x \gamma_{x,x0^m} = \frac{1}{2^m} \sum_{\alpha \in \{0,1\}^m} \varphi_{x\alpha} \otimes \mu_{xx}^{0^m \alpha}$$

$$\overline{\mu}'_{qq} = \sum_{x \in \{0,1\}^m} p_x \gamma'_{x,x0^m}, \qquad p_x \gamma'_{x,x0^m} = \frac{1}{2^m} \sum_{\alpha \in \{0,1\}^m} \varphi_{x\alpha} \otimes \mu_{xx}^{\prime 0^m \alpha}$$

where $p_x = \operatorname{Tr} \mu'_{xx}$ for $x \in \{0,1\}^m$. Note that by Lemma 53 we have $\mu_{xx}^{0^m \alpha} \perp \mu_{xx}^{0^m \beta}$ for all distinct $\alpha, \beta \in \{0,1\}^m$, which in turn implies $\gamma_{x,x0^m}$ is a private state. Averaging over

$x \in \{0,1\}^m$, the states $\gamma'_{x,x0^m}$ approximates the private states $\gamma_{x,x0^m}$, and more precisely we have

$$\varepsilon \geq \frac{1}{2}\left\|\mu' - \mu\right\|_1 \geq \frac{1}{2}\left\|\overline{\mu}'_{qq} - \overline{\mu}_{qq}\right\|_1 = \sum_{x\in\{0,1\}^m} p_x \frac{1}{2}\left\|\gamma'_{x,x0^m} - \gamma_{x,x0^m}\right\|_1,$$

where the second inequality is due to the monotonicity of trace distance and the equality follows as $\varphi_{x\alpha} \perp \varphi_{x'\beta}$ for all distinct $x, x' \in \{0,1\}^m$ and all $\alpha, \beta \in \{0,1\}^m$. Let $\Sigma_\varepsilon \subseteq \{0,1\}^m$ denote the set of $x \in \{0,1\}^m$, where $\frac{1}{2}\left\|\gamma'_{x,x0^m} - \gamma_{x,x0^m}\right\|_1 \leq \sqrt{\varepsilon}$. Then

$$\begin{aligned}
D\left(\mu' \| \sigma\right) &\geq D\left(\sum_{x\in\{0,1\}^m} p_x \gamma'_{x,x0^m} \middle\| \mathcal{E}_{\mathrm{Bell}}\left(\sigma\right)\right) \\
&= -\operatorname{H}\left(\{p_x\}_{x\in\{0,1\}^m}\right) + \sum_{x\in\{0,1\}^m} p_x D\left(\gamma'_{x,x0^m} \middle\| \mathcal{E}_{\mathrm{Bell}}\left(\sigma\right)\right) \\
&\geq -\operatorname{H}\left(\{p_x\}_{x\in\{0,1\}^m}\right) + \sum_{x\in\Sigma_\varepsilon} p_x D\left(\gamma'_{x,x0^m} \middle\| \mathcal{E}_{\mathrm{Bell}}\left(\sigma\right)\right)
\end{aligned}$$

Now, recall $\mathcal{E}_{\mathrm{Bell}} \in \mathcal{C}_{A\to B}\left(A_k : B_k\right)$ and so preserves separable states. Furthermore, we note that Theorem 9 in [6] provides a lower bound on the relative entropy of entanglement of approximate private states, which allows us to infer that

$$\begin{aligned}
D\left(\mu' \| \sigma\right) &\geq -\operatorname{H}\left(\{p_x\}_{x\in\{0,1\}^m}\right) + \sum_{x\in\Sigma_\varepsilon} p_x\left(\left(1 - 4\sqrt{\varepsilon}\right)m - h\left(\sqrt{\varepsilon}\right)\right) \\
&\geq -\operatorname{H}\left(\{p_x\}_{x\in\{0,1\}^m}\right) + \left(1 - \sqrt{\varepsilon}\right)\left(\left(1 - 4\sqrt{\varepsilon}\right)m - h\left(\sqrt{\varepsilon}\right)\right),
\end{aligned}$$

where we in the second inequality have used that

$$\sum_{x\in\{0,1\}^m\setminus\Sigma_\varepsilon} p_x\sqrt{\varepsilon} \leq \sum_{x\in\{0,1\}^m\setminus\Sigma_\varepsilon} p_x \frac{1}{2}\left\|\gamma'_{x,x0^m} - \gamma_{x,x0^m}\right\|_1 \leq \varepsilon$$

to infer that $\sum_{x\in\Sigma_\varepsilon} p_x \geq 1 - \sqrt{\varepsilon}$. This proves the desired statement □

The statement of Theorem 54 has connections back to our lower bound on the rate of distillable private key in Proposition 45. Here, we saw that an m-bit key-correlated state $\rho \in \mathcal{D}\left(A_k B_k A_s B_s\right)$, which is additionally joint $A_k B_k$-phase orthogonal, is an m-bit private state. In fact, for an arbitrary state $\rho \in \mathcal{D}\left(A_k B_k A_s B_s\right)$ given by

$$\rho = \sum_{x,y\in\{0,1\}^m} \underbrace{\sum_{i,j\in\{0,1\}^m} |\varphi_{xi}\rangle\langle\varphi_{yj}| \otimes \mu_{xiyj}}_{\mu_{xy}}, \qquad \mu_{xiyj} \in \mathrm{L}\left(A_s B_s\right),$$

the statement of Proposition 45 generalizes through an identical argument to

$$K_D^{\rightarrow}\left(\rho\right) \geq D\left(\rho \| \hat{\rho}\right) - \operatorname{H}\left(\{p_x\}\right), \tag{3.6}$$

where $p_x = \operatorname{Tr}\mu_{xx}$ for $x \in \{0,1\}^m$, highlighting the fact that we can distill private key, whenever we have phase orthogonality *and* correlation. Furthermore, we may note that if we consider an m-bit key-correlated state $\mu \in \mathcal{D}\left(A_k B_k A_s B_s\right)$, then the state $\left(\mathcal{H}_{A_k}^{\otimes m} \otimes \mathcal{H}_{B_k}^{\otimes m}\right)(\mu)$ is joint $A_k B_k$-phase orthogonal. Finally, it is an easy consequence of Theorem 54 that when considering separable states, then phase orthogonality and correlation are mutually exclusive properties, as we shall see now.

Corollary 55. Let $\mu \in \mathcal{D}_{\text{sep}}\left(A_k A_s : B_k B_s\right)$ and suppose $d_{A_k}, d_{B_k} = 2^m$ for $m \in \mathbb{N}$. We have the following statements

- If μ is an m-bit joint $A_k B_k$-phase orthogonal state, then $\hat{\mu}_{qq}^{A_k^* B_k^*} = \omega^{A_k^* B_k^*}$.
- If μ is an m-bit key-correlated state, then $\overline{\mu}^{\alpha\beta} = \overline{\mu}$ for all $\alpha, \beta \in \{0,1\}^m$.

Proof. Suppose μ is an m-bit joint $A_k B_k$-phase orthogonal state given by

$$\mu = \sum_{x,y \in \{0,1\}^m} \underbrace{\sum_{i,j \in \{0,1\}^m} |\varphi_{xi}\rangle\langle\varphi_{yj}| \otimes \mu_{xiyj}}_{\mu_{xy}} .$$

Applying Theorem 54 we note that $\overline{\mu}_{qq}$ is given by

$$\overline{\mu}_{qq} = \sum_{x \in \{0,1\}^m} p_x \gamma_{x,x0^m}$$

for some probability distribution $(p_x)_{x \in \{0,1\}^m}$, where $\gamma_{x,x0^m}$ are x-bit flipped private states. Since μ is separable, it follows from Theorem 54 that

$$0 = E_R(\mu) \geq m - \mathrm{H}\left(\{p_x\}_{x \in \{0,1\}^m}\right) \geq 0,$$

which in turn implies $p_x = \frac{1}{2^m}$ for all $x \in \{0,1\}^m$. Finally, this implies that

$$\hat{\mu}_{qq}^{A_k^* B_k^*} = \sum_{x \in \{0,1\}^m} p_x \hat{\gamma}_{x,x0^m}^{A_k^* B_k^*} = \frac{1}{2^m} \sum_{x \in \{0,1\}^m} \hat{\varphi}_{x0^m}^{A_k^* B_k^*} = \omega^{A_k^* B_k^*},$$

which proves the first statement.

Suppose μ is an m-bit key-correlated state. It follows from Proposition 45 that

$$0 = E_R(\mu) \geq K_D^{\rightarrow}(\mu) \geq m - \mathrm{H}\left(X | A_k B_k A_s B_s\right)_{\mu_{cq}} \geq 0,$$

and so we must have $\overline{\mu}^{0^m\alpha} = \overline{\mu}^{0^m\beta}$ for all $\alpha, \beta \in \{0,1\}^m$. Applying the channel $\mathcal{Z}_{A_k}^{\alpha}$ yields

$$\overline{\mu}^{\alpha\beta} = \mathcal{Z}_{A_k}^{\alpha}\left(\overline{\mu}^{0^m\beta}\right) = \mathcal{Z}_{A_k}^{\alpha}\left(\overline{\mu}^{0^m\alpha}\right) = \overline{\mu}^{\alpha\alpha} = \overline{\mu},$$

which proves the desired statement. □

3.2.2 Locally Private States

We began our discussion of phase orthogonality as we saw that private states exhibit this very behaviour, however, as we saw in Example 47 there are joint phase orthogonal states with no distillable secret key, namely, separable states. In this section we take a step back and introduce a more general set of states, namely, locally private states, and as we shall see the promise of local privacy is equivalent to joint phase orthogonality.

Definition 56. Let $\mu \in \mathcal{D}(A_kB_kA_sB_s)$ and suppose $d_{A_k}, d_{B_k} = 2^m$ for $m \in \mathbb{N}$. Let $|\Psi\rangle\langle\Psi| \in \mathcal{D}(A_kA_sB_kB_sE)$ be a purification of μ. We say μ is a *locally private state*, if

$$\hat{\Psi}^{A_kB_kE} = \sum_{x\in\{0,1\}^m} p_x \hat{\varphi}_x^{A_kB_k} \otimes \Psi_x^E,$$

where $(p_x)_{x\in\{0,1\}^m}$ is a probability distribution.

Let $\varepsilon \geq 0$. More generally, we say that μ' is an *ε-approximate locally private state*, if $\mu' \approx_\varepsilon \mu$ for some locally private state $\mu \in \mathcal{D}(A_kB_kA_sB_s)$.

To motivate the choice of terminology, we show that a locally private state at most allows Eve to infer the correlation between Alice's and Bob's measurement outcomes, while she remains oblivious to their individual measurement outcomes.

Proposition 57. Let $\mu' \in \mathcal{D}(A_kB_kA_sB_s)$ and suppose $d_{A_k}, d_{B_k} = 2^m$ for $m \in \mathbb{N}$. Suppose μ' is an ε-approximate locally private state. For any purification $|\Psi'\rangle\langle\Psi'| \in \mathcal{D}(A_kB_kA_sB_sE)$ of μ', we have

$$\frac{1}{2}\left\|\hat{\Psi}'^{A_kB_kE} - \sum_{x\in\{0,1\}^m} p_x\hat{\varphi}_x \otimes \Psi_x^E\right\|_1 < \sqrt{\varepsilon}$$

for some probability distribution $(p_x)_{x\in\{0,1\}^m}$ and $\{\Psi_x\}_{x\in\{0,1\}^m} \subseteq \mathcal{D}(E)$. In particular, we have

$$\Delta(A_k|E)_{\hat{\Psi}'} < 2\sqrt{\varepsilon}, \qquad \Delta(B_k|E)_{\hat{\Psi}'} < 2\sqrt{\varepsilon}.$$

Proof. Let $\varepsilon > 0$ and suppose μ' is an ε-approximate locally private state, that is, $\mu' \approx_\varepsilon \mu$ for some locally private state $\mu \in \mathcal{D}(A_kB_kA_sB_s)$. Let $|\Psi'\rangle$ be a purification of μ', and note that due to Uhlmann's Theorem [38] we may choose a purification $|\Psi\rangle$ of μ, such that

$$F(\mu', \mu) = |\langle\Psi' \,|\, \Psi\rangle|^2.$$

It follows from two applications of the Fuchs-van de Graaf inequality [12] and the monotonicity of the fidelity function that

$$\frac{1}{2}\left\|\hat{\Psi}' - \hat{\Psi}\right\|_1 \leq \sqrt{1 - F\left(\hat{\Psi}', \hat{\Psi}\right)} \leq \sqrt{1 - F(\Psi', \Psi)} < \sqrt{\varepsilon},$$

which proves the desired inequality.

Finally, we have

$$\begin{aligned}\Delta(A_k|E)_{\hat{\Psi}} &= \frac{1}{2}\left\|\hat{\Psi}'^{A_kE} - \omega^{A_k} \otimes \Psi'^E\right\|_1 \\ &\leq \frac{1}{2}\left\|\hat{\Psi}'^{A_kE} - \omega^{A_k} \otimes \Psi^E\right\|_1 + \frac{1}{2}\left\|\omega^{A_k} \otimes \Psi^E - \omega^{A_k} \otimes \Psi'^E\right\|_1 \\ &\leq \frac{1}{2}\left\|\hat{\Psi}'^{A_kB_kE} - \sum_{x\in\{0,1\}^m} p_x\hat{\varphi}_x^{A_kB_k} \otimes \Psi_x^E\right\|_1 + \frac{1}{2}\left\|\omega^{A_k} \otimes \Psi^E - \omega^{A_k} \otimes \Psi'^E\right\|_1 \\ &\leq 2\sqrt{\varepsilon},\end{aligned}$$

and an analogous proof shows the last inequality. □

We now give a characterization of locally private states along the lines of (3.1). The characterization (3.1) was shown in [6], and the proof of the statement below runs along the very same lines.

Proposition 58. Let $\mu \in \mathcal{D}\left(A_kB_kA_sB_s\right)$ and suppose $d_{A_k}, d_{B_k} = 2^m$ for $m \in \mathbb{N}$. Then μ is a locally private state, if and only if

$$\mu = U\left(\sum_{z,z' \in \{0,1\}^m} \sqrt{p_z p_{z'}}\, |\varphi_{z0^m}\rangle\langle\varphi_{z'0^m}|^{A_kB_k} \otimes \sigma_{zz'}^{A_sB_s}\right)U^\dagger,$$

for some unitary $U = \sum_{x,z\in\{0,1\}^m} |x\,(x+z)\rangle\langle x\,(x+z)| \otimes U_{xz}$ with $U_{xz} \in \mathrm{U}\left(A_sB_s\right)$, and some operators $\sigma_{zz'} \in \mathrm{L}\left(A_sB_s\right)$.

Proof. Suppose μ is a locally private state, and let $|\Psi\rangle$ be a purification of μ with purifying system E. We may write out $|\Psi\rangle$ as

$$|\Psi\rangle_{A_kB_kA_sB_sE} = \sum_{x,y\in\{0,1\}^m} q_{xy}\, |xy\rangle_{A_kB_k} \otimes |\Psi_{xy}\rangle_{A_sB_sE}\,,$$

where $q_{xy} \geq 0$ for all $x, y \in \{0,1\}^m$, and so the key-attacked state is given by

$$\begin{aligned}\hat{\Psi}_{A_kB_kA_sB_sE} &= \sum_{x,y\in\{0,1\}^m} q_{xy}^2\, |xy\rangle\langle xy|_{A_kB_k} \otimes |\Psi_{xy}\rangle\langle\Psi_{xy}|_{A_sB_sE} \\ &= \sum_{x,z\in\{0,1\}^m} q_{x(x+z)}^2\, |x\,(x+z)\rangle\langle x\,(x+z)|_{A_kB_k} \otimes \left|\Psi_{x(x+z)}\right\rangle\left\langle\Psi_{x(x+z)}\right|_{A_sB_sE}\end{aligned}$$

By assumption of local privacy, it follows that the key-attacked state satisfies

$$\hat{\Psi}_{A_kB_kE} = \frac{1}{2^m}\sum_{x,z\in\{0,1\}^m} p_z\, |x\,(x+z)\rangle\langle x\,(x+z)| \otimes \Psi_z^E,$$

which in turn implies that for all $x, z \in \{0,1\}^m$ we have

$$q_{x(x+z)}^2 \Psi_{x(x+z)}^E = q_{x(x+z)}^2 \operatorname{Tr}_{A_sB_s} \left|\Psi_{x(x+z)}\right\rangle\left\langle\Psi_{x(x+z)}\right|^{A_sB_sE} = \frac{1}{2^m} p_z \Psi_z^E.$$

Evidently, we must have $q_{x(x+z)}^2 = \frac{1}{2^m}p_z$, and furthermore we may infer the existence of unitaries $U_{xz} \in \mathrm{U}\left(A_sB_s\right)$ such that

$$\left|\Psi_{x(x+z)}\right\rangle = \sum_i \sqrt{\lambda_{z,i}}\, U_{xz}\, |\psi_{z,i}\rangle^{A_sB_s} \otimes |\xi_{z,i}\rangle^E\,,$$

where $\{|\psi_{z,i}\rangle\}_i$, $\{|\xi_{z,i}\rangle\}_i$ are sets of orthonormal vectors. If we denote by $\sigma_{zz'} \in \mathrm{L}\left(A_sB_s\right)$ the operator

$$\sigma_{zz'} = \sum_{i,i'} \sqrt{\lambda_{z,i}\lambda_{z',i'}}\, \left|\left\langle\xi_{z,i}\,\middle|\,\xi_{z',i'}\right\rangle\right|^2 \left|\psi_{z,i}\right\rangle\left\langle\psi_{z',i'}\right|,$$

then we obtain

$$\mu = \mathrm{Tr}_E \left|\Psi\right\rangle\!\left\langle\Psi\right| = \sum_{x,z,x',z'\in\{0,1\}^m} \frac{1}{2^m}\sqrt{p_z p_{z'}}\,\left|x\,(x+z)\right\rangle\!\left\langle x'\,(x'+z')\right| \otimes U_{xz}\sigma_{zz'}U^\dagger_{x'z'}$$

$$= U\left(\sum_{z,z'\in\{0,1\}^m} \sqrt{p_z p_{z'}}\,\left|\varphi_{z0^m}\right\rangle\!\left\langle\varphi_{z'0^m}\right| \otimes \sigma_{zz'}\right)U^\dagger,$$

where $U = \sum_{x,z\in\{0,1\}^m} \left|x\,(x+z)\right\rangle\!\left\langle x\,(x+z)\right| \otimes U_{xz}$. This proves the desired statement. □

Having established a canonical form of locally private states, we proceed to one of our main results of this chapter, namely, a tight connection between local privacy and joint phase orthogonality. Before doing so we prove two results, which will serve as building blocks toward our final result. First, we generalize the observation that for a private state we have $\gamma = \overline{\gamma}$ to the more general setting of locally private states.

Lemma 59. Let $\mu \in \mathcal{D}\left(A_kB_kA_sB_s\right)$ and suppose $d_{A_k}, d_{B_k} = 2^m$ for $m \in \mathbb{N}$. Then μ is a locally private state, if and only if $\overline{\mu}$ is a locally private state.

Proof. Suppose μ is a locally private state, and note that by Proposition 58 we have

$$\mu = U\left(\sum_{z,z'\in\{0,1\}^m} \sqrt{p_z p_{z'}}\,\left|\varphi_{z0^m}\right\rangle\!\left\langle\varphi_{z'0^m}\right|^{A_kB_k} \otimes \sigma_{zz'}^{A_sB_s}\right)U^\dagger,$$

for some unitary $U = \sum_{x,z\in\{0,1\}^m} \left|x\,(x+z)\right\rangle\!\left\langle x\,(x+z)\right| \otimes U_{xz}$ with $U_{xz} \in \mathrm{U}\left(A_sB_s\right)$, and some operators $\sigma_{zz'} \in \mathrm{L}\left(A_sB_s\right)$. But then

$$\overline{\mu} = U\left(\sum_{z\in\{0,1\}^m} p_y\,\left|\varphi_{z0^m}\right\rangle\!\left\langle\varphi_{z0^m}\right|^{A_kB_k} \otimes \sigma_{zz}^{A_sB_s}\right)U^\dagger,$$

and again due to Proposition 58 we may infer that $\overline{\mu}$ is a locally private state.

For the converse implication, let $\mu \in \mathcal{D}\left(A_kB_kA_sB_s\right)$ be given by

$$\mu = \sum_{x,i,y,j\in\{0,1\}^m} \left|\varphi_{xi}\right\rangle\!\left\langle\varphi_{yj}\right| \otimes \mu_{xiyj}, \qquad \mu_{xiyj} \in \mathrm{L}\left(A_sB_s\right).$$

Let $\omega_F = \frac{1}{2^m}\sum_{\alpha\in\{0,1\}^m} \left|\alpha\right\rangle\!\left\langle\alpha\right|$, where $\left\{\left|\alpha\right\rangle\right\}_{\alpha\in\{0,1\}^m}$ is an orthonormal basis of system F. If we denote by $U \in \mathrm{U}\left(A_kB_kF\right)$ the unitary given by

$$U = \sum_{\alpha\in\{0,1\}^m} Z^\alpha_{A_k} \otimes Z^\alpha_{B_k} \otimes \left|\alpha\right\rangle\!\left\langle\alpha\right|_F,$$

then we have $\overline{\mu} = \mathrm{Tr}_F\,\mu_U$, where $\mu_U = U\mu\otimes\omega U^\dagger$. Now suppose $\overline{\mu}$ is an m-bit locally private state, and note that any purification $\left|\Psi_U\right\rangle\!\left\langle\Psi_U\right| \in \mathcal{D}\left(A_kB_kA_sB_sEF\right)$ of μ_U is also a purification of $\overline{\mu}$. This implies that μ_U is also a locally private state. For any purification $\left|\Psi\right\rangle\!\left\langle\Psi\right| \in \mathcal{D}\left(A_kB_kA_sB_sE\right)$ of μ, we may note that

$$\left|\Psi'\right\rangle = \left|\Psi\right\rangle \otimes \frac{1}{\sqrt{2^m}}\sum_{\alpha\in\{0,1\}^m} \left|\alpha\alpha\right\rangle^{FG}$$

is a purification of $\mu \otimes \omega_F$, and so $U\left|\Psi'\right\rangle$ is a purification of μ_U. Finally, it follows that

$$\hat{\Psi}^{A_kB_kE} = \operatorname{Tr}_{A_sB_s}\hat{\Psi} = \operatorname{Tr}_{A_sB_sFG}\hat{\Psi}' = \operatorname{Tr}_{A_sB_sFG}\left(U\hat{\Psi}'U^\dagger\right) = \sum_{x\in\{0,1\}^m} p_x\hat{\varphi}_x^{A_kB_k}\otimes\Psi_x^E,$$

where the last equality follows as μ_U is a locally private state. □

The next result is a natural generalization of the observation that a Bell key-correlated state is in fact a private state whenever the states of the shield system are orthogonal [1].

Lemma 60. Let $\mu \in \mathcal{D}\left(A_kB_kA_sB_s\right)$ and suppose $d_{A_k}, d_{B_k} = 2^m$ for $m \in \mathbb{N}$. Suppose μ is given by

$$\mu = \frac{1}{2^m}\sum_{x,\alpha\in\{0,1\}^m} p_x\left|\varphi_{x\alpha}\right\rangle\!\left\langle\varphi_{x\alpha}\right| \otimes \mu_{x\alpha x\alpha}, \qquad \mu_{x\alpha x\alpha} \in \mathcal{D}\left(A_sB_s\right)$$

where $\mu_{x\alpha x\alpha} \perp \mu_{y\beta y\beta}$ for all $x, y \in \{0,1\}^m$ and distinct $\alpha, \beta \in \{0,1\}^m$. Then μ is a locally private state.

Proof. Let $\mu_{x\alpha x\alpha} = \sum_i \lambda_{x\alpha,i}\left|\psi_{x\alpha,i}\right\rangle\!\left\langle\psi_{x\alpha,i}\right|$ be a spectral decomposition for $x, \alpha \in \{0,1\}^m$, and consider the purification $\left|\Psi_{x\alpha}\right\rangle\!\left\langle\Psi_{x\alpha}\right| \in \mathcal{D}\left(A_sB_sE\right)$ of $\mu_{x\alpha x\alpha}$ given by

$$\left|\Psi_{x\alpha}\right\rangle = \sum_i \sqrt{\lambda_{x\alpha,i}}\left|\psi_{x\alpha,i}\right\rangle \otimes \left|i_{x\alpha}\right\rangle,$$

where $\{\left|i_{x\alpha}\right\rangle\}_{x,\alpha,i}$ is an orthonormal basis of E. Now consider the purification $\left|\Psi\right\rangle\!\left\langle\Psi\right| \in \mathcal{D}\left(A_kB_kA_sB_sE\right)$ of μ given by

$$\left|\Psi\right\rangle = \frac{1}{\sqrt{2^m}}\sum_{x,\alpha\in\{0,1\}^m}\sqrt{p_x}\left|\varphi_{x\alpha}\right\rangle \otimes \left|\Psi_{x\alpha}\right\rangle.$$

It follows from the orthogonality assumption, namely, $\mu_{x\alpha x\alpha} \perp \mu_{x'\alpha'x'\alpha'}$ for all $x, x' \in \{0,1\}^m$ and distinct $\alpha, \alpha' \in \{0,1\}^m$, that the key-attacked state is given by

$$\hat{\Psi}^{A_kB_kE} = \frac{1}{2^m}\sum_{x,\alpha\in\{0,1\}^m} p_x\hat{\varphi}_x^{A_kB_k}\otimes\Psi_{x\alpha}^E = \sum_{x\in\{0,1\}^m} p_x\hat{\varphi}_x^{A_kB_k}\otimes\frac{1}{2^m}\sum_{\alpha\in\{0,1\}^m}\Psi_{x\alpha}^E,$$

which shows μ is a locally private state. □

We are now ready to prove the main result of this section, namely, a generalization of the following observation: If ρ is an m-bit key-correlated state, and $\rho \perp \mathcal{Z}_{B_k}^{\beta}\left(\rho\right)$ for all $\beta \in \{0,1\}^m$, then ρ_{qq} is an m-bit private state [1].

Theorem 61. Let $\mu \in \mathcal{D}\left(A_kB_kA_sB_s\right)$ and suppose $d_{A_k}, d_{B_k} = 2^m$ for $m \in \mathbb{N}$. Then μ is an m-bit joint A_kB_k-phase orthogonal state, if and only if μ_{qq} is an m-bit locally private state.

Proof. Consider first $\mu_{qq} = \mathcal{E}_{\text{Bell}}\left(\mu\right) \in \mathcal{D}\left(A_k^*B_k^*A_kB_kA_sB_s\right)$, and note that μ_{qq} is an m-bit locally private state, if and only if $\overline{\mu}_{qq}$ is a locally private state by Lemma 59. Now note $\overline{\mu}_{qq}$ is given by

$$\overline{\mu}_{qq} = \frac{1}{2^m}\sum_{x,\alpha\in\{0,1\}^m} p_x\left|\varphi_{x\alpha}\right\rangle\!\left\langle\varphi_{x\alpha}\right| \otimes \underbrace{\sum_{i,j\in\{0,1\}^m}\left|\varphi_{x(i+\alpha)}\right\rangle\!\left\langle\varphi_{x(j+\alpha)}\right| \otimes \mu_{xixj}}_{\mu_{xx}^{\alpha}},$$

and by Lemma 60 it is locally private if and only if $\mu_{xx}^{\alpha} \perp \mu_{xx}^{\beta}$ for all distinct $\alpha, \beta \in \{0,1\}^m$. Finally, this is equivalent to μ being an m-bit joint $A_k B_k$-phase orthogonal state by Lemma 53. This proves the desired statement. □

The content of Theorem 61 gives an understanding of the implication of joint phase orthogonality in terms of privacy. It is clear that with the added assumption of μ being key-correlated, then we may infer that μ_{qq} is in fact a private state.

3.3 The Hiding Rate in a Bipartite Setting

In this last section of the chapter, we will extend the notion of a rate of distillable hiding states in Chapter 2 to a generic state shared between two parties Alice and Bob sharing the quantum system AB. In Example 47 we saw that it is possible to hide classical data in a separable state of a quantum system. This means that whenever Alice and Bob are allowed to perform local operations and classical communication, they can produce states for hiding classical data with no prior resource. Thus we will restrict ourselves to local operations on Bob's system B and explore to what extent Bob can encode data into the state of their shared system that remains hidden as long as Alice and Bob are spatially separated.

Definition 62. Let $\rho \in \mathcal{D}(AB)$, and let $r \geq 0$. We say that r is an *achievable hiding rate*, if $r = 0$ or the following condition holds: Let $\varepsilon > 0$. For sufficiently large $n \in \mathbb{N}$ and $m = \lfloor rn \rfloor$, there exists an instrument $\Lambda \in \mathcal{I}_{\text{all}}(B^n \rangle X B')$ such that

$$\Lambda\left(\rho^{\otimes n}\right) = \mu \in \mathcal{D}\left(X A^n B'\right),$$

where X is a 2^m-dimensional quantum system, and μ is an m-bit ε-hiding state with respect to a $\mathcal{M}_{\text{all}}(A^n)$-restricted eavesdropper on $A^n B'$.

The *hiding rate* $H_D^{\rightarrow}(\rho)$ is the supremum over all achievable hiding rates.

Even with the restriction to local operations on Bob's system, it is not completely trivial, though not surprising, that the hiding rate of a generic state is indeed finite. We begin by proving this elementary fact.

Proposition 63. Let $\rho \in \mathcal{D}(AB)$. Then

$$H_D^{\rightarrow}(\rho) \leq \log d_A$$

Proof. Let $\varepsilon > 0$ and $n \in \mathbb{N}$. Let $\Lambda \in \mathcal{I}_{\text{all}}(B^n \rangle X B')$ and suppose $\Lambda(\rho^{\otimes n}) = \mu \in \mathcal{D}(X A^n B')$ is an m-bit ε-hiding state with respect to $\mathcal{M}(A^n)$. It follows from Fano's inequality (see Lemma 84 in Appendix A) that

$$\mathrm{H}\left(X \middle| A^n B'\right)_{\mu} \leq \varepsilon m + h(\varepsilon),$$

and similarly that

$$\mathrm{H}\left(X \middle| B'\right)_{\mu} \geq \mathrm{H}_{A^n}\left(X \middle| A^n B'\right)_{\mu} \geq m - 2\varepsilon m - g(\varepsilon).$$

Thus, it follows that

$$\begin{aligned} m &\leq \mathrm{H}\left(X\middle|B'\right)_{\mu} - \mathrm{H}\left(X\middle|A^{n}B'\right)_{\mu} + 3\varepsilon m + g\left(\varepsilon\right) + h\left(\varepsilon\right) \\ &= \mathrm{H}\left(XB'\right)_{\mu} - \mathrm{H}\left(B\right)_{\mu} - \mathrm{H}\left(XA^{n}B'\right)_{\mu} + \mathrm{H}\left(A^{n}B'\right)_{\mu} + 3\varepsilon m + g\left(\varepsilon\right) + h\left(\varepsilon\right) \\ &= \mathrm{I}\left(A^{n} : XB'\right)_{\mu} - \mathrm{I}\left(A^{n} : B'\right)_{\mu} + 3\varepsilon m + g\left(\varepsilon\right) + h\left(\varepsilon\right) \\ &\leq n\log d_{A} + 3\varepsilon m + g\left(\varepsilon\right) + h\left(\varepsilon\right). \end{aligned}$$

If we divide by n on both sides of the inequality, it follows that the hiding rate cannot exceed $\log d_A$. □

We now proceed to show a lower bound on the hiding rate of key-correlated states. Inspired by our previous work, a natural approach for Bob to encode data into the state of the shared system is to apply phase gates Z to his key system.

Proposition 64. Let $\rho \in \mathcal{D}\left(A_k B_k A_s B_s\right)$ and suppose $d_{A_k}, d_{B_k} = 2^m$ for some $m \in \mathbb{N}$. If ρ is an m-bit key-correlated state, then

$$\widetilde{H}^{\infty}_{\min,\mathcal{M}(A_k A_s)}\left(X|A_k B_k A_s B_s\right)_{\rho_{cq}} - \mathrm{H}\left(X|A_k B_k A_s B_s\right)_{\rho_{cq}} \leq H^{\rightarrow}_{D}\left(\rho\right).$$

Proof. Let $\Lambda \in \mathcal{I}_{\text{all}}\left(B_k \rangle X B_k\right)$ by given by

$$\Lambda\left(\cdot\right) = \frac{1}{2^m} \sum_{\beta \in \{0,1\}^m} |\beta\rangle\langle\beta|^{X} \otimes \mathcal{Z}^{\beta}_{B_k}\left(\cdot\right),$$

and note that $\Lambda\left(\rho\right) = \rho_{cq}$. As Bob may further perform classical processing of the state in system X, this implies that

$$H_{D,A}\left(\rho_{cq}\right) \leq H^{\rightarrow}_{D}\left(\rho\right),$$

and so Theorem 40 yields the desired lower bound. □

Recall the statement of Proposition 46, and note that with this extended definition of the hiding rate of hiding rate we have $H_{D,A}\left(\rho_{cq}\right) \leq H^{\rightarrow}_{D}\left(\rho\right)$, but the two quantities can, of course, be distinct. We do note, however, that contingent on the statement in Conjecture 10 being true, the hiding rate is an upper bound on the difference between distillable key and distillable entanglement of key-correlated states whenever the key-attacked state is separable.

The particular instrument applied to Bob's key system in Proposition 64 can never yield more hidden data than the amount of distillable secret key when ρ is a key-correlated state. This is an easy consequence of the observation that

$$H_{D,A}\left(\rho_{cq}\right) \leq \mathrm{I}\left(X : A_k A_s B_k B_s\right)_{\rho_{cq}} \leq K^{\rightarrow}_{D}\left(\rho\right),$$

where the last inequality was shown in Proposition 64. As we saw in Example 47 the presence of entanglement is not necessary to achieve a non-zero hiding rate, so we cannot hope to find any measure of entanglement to be an upper bound in general. We pose it as a very open question to place bounds or impose further restrictions on the set of free channels on the quantity $H^{\rightarrow}_{D}\left(\cdot\right)$ beyond the elementary observations above.

Chapter 4

Quantum Advantage in Randomness Extraction

Among the central aspects of cryptography, we find the information-theoretic task to obtain secret key from some only partially secret classical data. An adversarial eavesdropper may have encoded some side information about the partially secret data, and so one may engage in *randomness extraction* to ensure the side information gives no information about the final data [39]. Trivially, encoding information in an n-level quantum system allows an adversarial eavesdropper to have at least as much information as when storing a letter from an alphabet $\mathcal{E}$ of size $|\mathcal{E}| = n$. In this chapter, we will consider a particular protocol for randomness extraction, which is in fact exceedingly vulnerable to a quantum eavesdropper, while remaining secure against a classical eavesdropper! [2] This is closely related to the topic of Chapter 2, and in fact, this matter of concern was already briefly discussed in Example 32. If we think of the classical eavesdropper as a quantum eavesdropper with no quantum memory and being restricted to the set of measurements $\mathcal{M}_{\text{all}}(E\rangle Y)$, where $d_Y = n$, then the scenarios are essentially equivalent. We present our efforts to make this quantum advantage realizable on current hardware along with preliminary results on the performance of classical and quantum computers in this scenario.

4.1 The Cryptographic Scenario

Let $\mathcal{X}$ be an alphabet, and suppose Xavier samples a letter $x \in \mathcal{X}$ with respect to some probability distribution $(p_x)_{x\in\mathcal{X}}$. Suppose an adversarial eavesdropper Eve has an n-dimensional quantum system E, and she is allowed access to the string x. She encodes some side information into her system E by preparing some state $\rho_x \in \mathcal{D}(E)$ depending on the observed value of x. Next, Eve's access to the string is taken away, and she forgets everything about her observations except what is encoded in her quantum system. This situation is described by the cq-state

$$\rho_{cq}^{XE} = \sum_{x\in\mathcal{X}} p_x \, |x\rangle\langle x|^X \otimes \rho_x^E. \tag{4.1}$$

We are now in the scenario of Eve having partial knowledge about the data encoded in the state of system X, and so Xavier must engage in randomness extraction in order to obtain a secret string, such that access to Eve's system yields only negligible information about the secret string. In the following, we describe a particular protocol for privacy amplification, which we shall refer to as the *matching protocol.*

4.1.1 The Matching Protocol

The matching protocol is parameterized by some $\alpha \in \left(0, \frac{1}{2}\right]$, and in following it is understood that αn is an integer, that is, $\alpha = \frac{k}{n}$ for $k = 1, \ldots, \frac{\lfloor n/2 \rfloor}{n}$. We will restrict ourselves to considering alphabets given by $\mathcal{X}_n = \{0,1\}^n$, $\mathcal{X}'_{\alpha n} = \{0,1\}^{\alpha n}$ for some $n \in \mathbb{N}$.

We will now describe the family $\mathcal{F}_{\alpha,n}$ of functions $f \colon \mathcal{X}_n \to \mathcal{X}'_n$ used in the matching protocol. To do so, we introduce a distribution of the $2\alpha n$ indices among $\{1, \ldots, n\}$ into αn disjoint pairs, which corresponds to choosing an *α-matching of n indices* as is defined below.

Definition 65. Let $n \in \mathbb{N}$ and $\alpha \in (0, \frac{1}{2}]$. An *$\alpha$-matching of n indices* is a list of pairs of distinct indices $i_1, j_1, \ldots, i_{\alpha n}, j_{\alpha n} \in \{1, \ldots, n\}$, that is,

$$M = \left(\{i_1, j_1\}, \{i_2, j_2\}, \ldots, \{i_{\alpha n}, j_{\alpha n}\}\right), \tag{4.2}$$

and whenever $\alpha = \frac{1}{2}$ we simply refer to M as a *matching of n indices.*

We denote by $\mathcal{M}_{\alpha,n}$ the set of all α-matchings of n indices.

Remark. Whenever $n \in \mathbb{N}$ is clear from the context, we will simply refer to M as an *α-matching.*

With the notion of α-matchings in place, we can proceed to define the functions $f \colon \mathcal{X}_n \to \mathcal{X}'_{\alpha n}$. Given a string $x \in \mathcal{X}_n$ and a matching $M \in \mathcal{M}_{\alpha,n}$, we may compute a new string $x' \in \mathcal{X}'_{\alpha n}$ comprised of the letters

$$x'_k = x_{i_k} + x_{j_k}, \qquad k \in \{1, \ldots, \alpha n\},$$

where addition is carried out modulo 2. For each α-matching M this procedure corresponds to a function f_M used in the matching protocol, which we will now define.

Definition 66. Let $n \in \mathbb{N}$ and $\alpha \in (0, \frac{1}{2}]$. For each α-matching of n indices M given by (4.2) we define an *α-matching function of n letters* $f_M \colon \mathcal{X}_n \to \mathcal{X}_{\alpha n}$ by

$$f_M(x) \coloneqq (x_{i_1} \oplus x_{j_1}) \ldots (x_{i_{\alpha n}} \oplus x_{j_{\alpha n}}),$$

where it is understood that the parenthesized expressions above are concatenated into a string of zeros and ones of length αn. We denote by $\mathcal{F}_{\alpha,n}$ the set of all α-matching functions of n letters.

Remark. Whenever $n \in \mathbb{N}$ is clear from the context, we will simply refer to f_M as an *α-matching function.*

For $n \in \mathbb{N}$ and $\alpha \in (0, \frac{1}{2}]$ we refer to the proces of sampling a function $f \in \mathcal{F}_{\alpha,n}$ uniformly at random and applying it to $x \in \mathcal{X}$ as the *α-matching protocol.* For a cq-state $\rho \in \mathcal{D}(XE)$ given by (4.1) the state resulting from applying a particular α-matching function $f \in \mathcal{F}_{\alpha,n}$ is given by

$$f(\rho_{cq}) = \sum_{x \in \mathcal{X}} p_x \left|f(x)\right\rangle\!\left\langle f(x)\right| \otimes \rho_x = \sum_{x' \in \mathcal{X}'} \left|x'\right\rangle\!\left\langle x'\right| \otimes \sum_{x \in f^{-1}(\{x'\})} p_x \rho_x,$$

and so the state resulting from the α-matching protocol is

$$\mathcal{F}_{\alpha,n}\left(\rho_{cq}\right) = \frac{1}{\left|\mathcal{F}_{\alpha,n}\right|} \sum_{f \in \mathcal{F}_{\alpha,n}} f\left(\rho_{cq}\right) \otimes |f\rangle\langle f|^{F} .$$

Throughout this chapter, our primary concern will be the security of the α-matching protocol with respect to an eavesdropper, in particular how the security is dependent on whether the eavesdroppers information encoded into the state of a quantum system E or it is stored classically. More precisely, we will consider a setting concerning two parties Xavier and Eve. Suppose they engage in a protocol with parameter $\alpha \in (0, \frac{1}{2}]$ and $n \in \mathbb{N}$ described by the following steps:

- Xavier samples a string $x \in \mathcal{X}_n$ uniformly at random and encodes the outcome into the state of his system X. The resulting state is
$$\rho_c^X = \frac{1}{2^n} \sum_{x \in \mathcal{X}_n} |x\rangle\langle x|^X ,$$
where $\{|x\rangle\}_{x \in \mathcal{X}_n}$ is an orthonormal basis of system X.
- Eve is allowed temporary access to system X, and upon observing x she prepares the state ρ_x in her system E. Eve's access to system X is then taken away. The resulting state is
$$\rho_{cq}^{XE} = \frac{1}{2^n} \sum_{x \in \mathcal{X}_n} |x\rangle\langle x|^X \otimes \rho_x^E,$$
- Xavier samples an α-matching function $f \in \mathcal{F}_{\alpha,n}$ uniformly at random, and transmits the choice of f to Eve. Furthermore, Xavier computes $f\left(x\right)$ and stores the outcome in a system X'. The resulting state is
$$\mathcal{F}_{\alpha,n}\left(\rho_{cq}\right) = \frac{1}{2^n \left|\mathcal{F}_{\alpha,n}\right|} \sum_{f \in \mathcal{F}_{\alpha,n}, x \in \mathcal{X}_n} |f\left(x\right)\rangle\langle f\left(x\right)|^{X'} \otimes \rho_x^E \otimes |f\rangle\langle f|^F ,$$
where $\{|x'\rangle\}_{x' \in \mathcal{X}'_{\alpha n}}$, $\{|f\rangle\}_{f \in \mathcal{F}_{\alpha,n}}$ are orthonormal bases of X', F, respectively.
- Eve tries to guess one bit of $f\left(x\right)$ based on access to system EF. She outputs a bit $b \in \{0,1\}$ and a corresponding index $k \in \{1, \ldots, \alpha n\}$.
- We say that Eve has successfully guessed one bit of $f\left(x\right)$, if $f\left(x\right)_k = b$, that is, the kth bit of $f\left(x\right)$ is equal to b.

Before Eve tried to guess one bit of the string in the protocol above, the state describing the situation is given by $\mathcal{F}_{\alpha,n}\left(\rho_{cq}\right)$, and we denote the probability of success in the protocol above by

$$\mathrm{Pr}_{\text{guess-1-bit}}\left(X' \middle| EF\right)_{\mathcal{F}_{\alpha,n}\left(\rho_{cq}\right)} .$$

To make this interesting, we have to restrict the information encoded in the system E; when given access to the system X, Eve may simply choose to read of $x \in \mathcal{X}_n$ encode x into the state of her own system E. In this case, Eve is trivially able to infer the value of $f\left(x\right)$ once she is informed of Xavier's choice of $f \in \mathcal{F}_{\alpha,n}$. This shows that with no restrictions Eve has a perfect classical strategy. In order to identify a scenario, where a quantum strategy outperforms any classical strategy, we will have to refine our approach - for one we will have to exclude the strategy of storing the entire string. We will return to this point shortly.

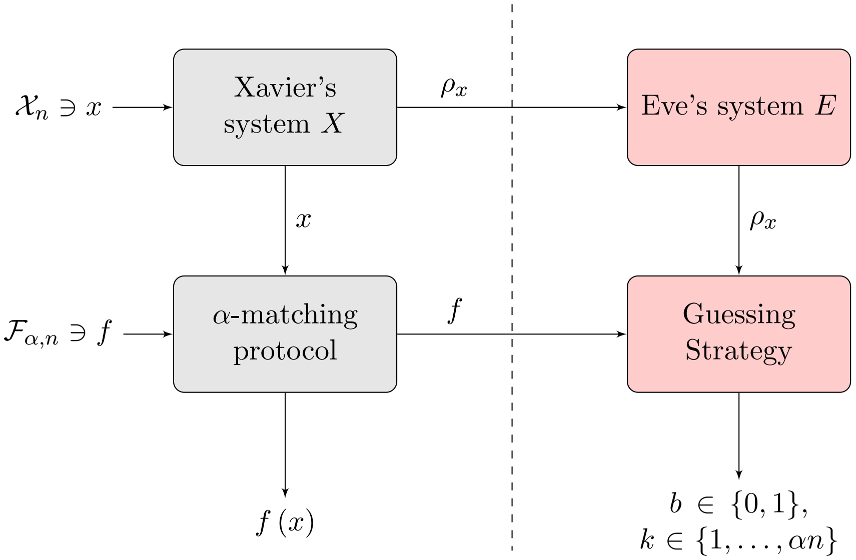


Figure 4.1: The diagram depicts the operational task, where first Xavier samples a string $x \in \mathcal{X}_n$ uniformly at random. Allowing Eve access to read of x, she encodes her information in the state ρ_x of her system E. Afterward, Eve is denied access to system X, and Xavier engages in the α-matching protocol. Finally, given the α-matching function $f \in \mathcal{F}_{\alpha,n}$ and the information encoded in the state ρ_x of system E, Eve outputs a guess on one of the bits of $f(x)$.

4.2 Lower Bound on Quantum Strategies

We approach the task of lower bounding the performance of a quantum strategy for guessing one bit of the string $x' = f(x)$ given access to a quantum system E in state ρ_x and a matching function f by describing a perfect quantum strategy. More precisely, we identify which quantum states to prepare in system E and what measurement strategy to employ in order to determine one bit of $x' = f(x)$. We shall see that when Xavier samples a string $x \in \{0,1\}^n$ uniformly at random, it suffices for Eve to have an n-level quantum system E at her disposal.

4.2.1 An Optimal Quantum Strategy

We now describe a quantum strategy for guessing one bit of the resulting string from the matching protocol, where the initial string x is sampled from $\mathcal{X}_n$ of size 2^n, and Eve's system is an n-level quantum system. Suppose Xavier samples a string $x \in \mathcal{X}_n = \{0,1\}^n$ uniformly at random. On observing x, Eve prepares the pure state $|\rho_x\rangle\langle\rho_x| \in \mathcal{D}(E)$ given by

$$|\rho_x\rangle = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} (-1)^{x_i} |i\rangle, \tag{4.3}$$

where $\{|i\rangle\}_{i=1}^{n}$ constitutes an orthonormal basis of system E. We will now see that the information encoded in this n-level quantum system is sufficient to extract one bit of $x' = f(x)$ with certainty for some matching function f.

Proposition 67. Let $\alpha \in (0, \frac{1}{2}]$ and suppose $n \in \mathbb{N}$ is even. Consider $\rho^{XE} \in \mathcal{D}(XE)$ given by

$$\rho^{XE} = \frac{1}{2^n} \sum_{x \in \mathcal{X}_n} |x\rangle\langle x|^X \otimes \rho_x^E, \qquad \text{where } \rho_x = \frac{1}{n} \sum_{i,j=1}^{n} (-1)^{x_i + x_j} |i\rangle\langle j|.$$

Then

$$\Pr_{\text{guess-1-bit}} \left(X' \middle| EF\right)_{\mathcal{F}_{\alpha,n}(\rho)} = \frac{1}{2} + \alpha.$$

Remark. This is essentially shown in [2], but we supply the proof here in order to support the proof of an implementation of an optimal guessing strategy.

Proof. Let $f_M \in \mathcal{F}_{\alpha,n}$ be an α-matching function. If we extend M to a matching M_{ext} by adding disjoint pairs of the remaining indices, we also have a corresponding matching function $f = f_{M_{\text{ext}}}$. Denote by Λ'_f the quantum instrument given by

$$\Lambda'_f(\cdot) = \sum_{k=1}^{n/2} |k\rangle\langle k| \otimes N_k \cdot N_k^{\dagger}, \qquad N_k = |i_k\rangle\langle i_k| + |j_k\rangle\langle j_k|,$$

where $\{i_k, j_k\}$ is the kth pair of the extended matching corresponding to the matching function $f = f_M$. For $x \in \mathcal{X}_n$ we have

$$\begin{aligned}\Lambda'_f(\rho_x) &= \sum_{k=1}^{n/2} |k\rangle\langle k| \otimes N_k \rho_x N_k^\dagger \\ &= \frac{1}{n} \sum_{k=1}^{n/2} |k\rangle\langle k| \otimes \left((-1)^{x_{i_k}} |i_k\rangle + (-1)^{x_{j_k}} |j_k\rangle\right)\left((-1)^{x_{i_k}} \langle i_k| + (-1)^{x_{j_k}} \langle j_k|\right).\end{aligned}$$

Next, denote by Λ''_f the measurement with POVM representation $P_{b,k}$ for $b \in \{0,1\}$, $k \in \left\{1, \dots \frac{n}{2}\right\}$ given by

$$P_{k,b} = |k\rangle\langle k| \otimes \Big(|i_k\rangle + (-1)^b |j_k\rangle\Big)\Big(\langle i_k| + (-1)^b \langle j_k|\Big).$$

Letting $\Lambda_f = \Lambda''_f \circ \Lambda'_f$, we obtain

$$\Lambda_f(\rho_x) = \frac{1}{n/2} \sum_{k=1}^{n/2} |k\rangle\langle k| \otimes |x_{i_k} \oplus x_{j_k}\rangle\langle x_{i_k} \oplus x_{j_k}| = \frac{1}{n/2} \sum_{k=1}^{n/2} |k\rangle\langle k| \otimes |f(x)_k\rangle\langle f(x)_k|$$

which shows that there exists a measurement $\Lambda_f \in \mathcal{M}(E)$, which allows us to infer 1 bit of $f(x)$ with probability 1. As this one bit is also a bit of $f_M(x)$ with probability 2α, this proves

$$\Pr_{\text{guess-1-bit}}\left(X' \middle| EF\right)_{\mathcal{F}_{\alpha,n}(\rho)} = \frac{1}{2} + \alpha$$

as desired. □

The assumption that $n \in \mathbb{N}$ is even makes sure that a matching includes all indices. When this is not the case, that is, $n \in \mathbb{N}$ is odd, then the probability of success of the quantum strategy above only drops slightly as described below.

Corollary 68. Let $\alpha \in (0, \frac{1}{2}]$ and suppose $n \in \mathbb{N}$ is odd. Consider $\rho^{XE} \in \mathcal{D}(XE)$ given by

$$\rho^{XE} = \frac{1}{2^n} \sum_{x \in \mathcal{X}_n} |x\rangle\langle x|^X \otimes \rho_x^E, \qquad \text{where } \rho_x = \frac{1}{n} \sum_{i,j=1}^{n} (-1)^{x_i + x_j} |i\rangle\langle j|.$$

Then

$$\Pr_{\text{guess-1-bit}}\left(X' \middle| EF\right)_{\mathcal{F}_{\frac{1}{2},n}(\rho)} \geq \frac{1}{2} + \alpha - \frac{1}{2n}.$$

Proof. This is shown analogously to Proposition 67. □

Observe that even with a significant restriction on the size of the quantum system E, a quantum eavesdropper is still able to retrieve one bit of the extracted string $x' = f(x)$ with high probability. This motivates the following proposal for a restriction, which ensures that the quantum protocol above is superior to any classical protocol: The eavesdropper's side information must be encoded in the state of an n-level quantum system or in an n-letter alphabet $\mathcal{E}$, that is, $|\mathcal{E}| = n$.

4.2.2 Implementation of Optimal Quantum Strategy

We proceed to describe an actual implementation of the two subprotocols in the quantum strategy above, namely, 1) the preparation of the state $\rho_x \in \mathcal{D}(E)$ for $x \in \mathcal{X}_n$ and 2) the measurement strategy Λ_f given a matching function $f \in \mathcal{F}_{\frac{1}{2},n}$. As we will present the performance results of running this protocol on current quantum hardware, the description of the implementation of the protocol is adapted to quantum systems consisting of multiple *qubits*, that is, 2-level quantum systems and unitary transformations are described in terms of a set of elementary gates.

In the following, we present two approaches to the desired state preparation and measurement subprotocols.

A protocol based on permutations of computational basis

Let $m \in \mathbb{N}$. In the following we will consider a 2^m-level quantum system E with computational basis given by $\{|x\rangle\}_{x\in\{0,1\}^m}$. As we will alternate between representing natural numbers $i \in \mathbb{N}$ in the decimal system and as strings of zeros and ones, we introduce the functions

$$\begin{aligned} \mathrm{dec}\colon \{0,1\}^m &\to \{0,\dots,2^m-1\}, & \mathrm{dec}(x) &\coloneqq \sum_{i=1}^{m} x_i \cdot 2^{m-i}, \\ \mathrm{bin}\colon \{0,\dots,2^m-1\} &\to \{0,1\}^m, & \mathrm{bin}(k) &\coloneqq \mathrm{dec}^{-1}(k). \end{aligned}$$

For a decimal representations of $n \in \mathbb{N}$, we will write $|n\rangle \coloneqq |\mathrm{bin}(n)\rangle$. We now present the first step towards preparing $|\rho_x\rangle$ defined by (4.3).

Lemma 69. Let $n \in \mathbb{N}$, and let $k \in \mathbb{N}_0$ be the integer satisfying $2^{k-1} < n \leq 2^k$. Then there is a unitary $U_{n,k}$ such that

$$U_{nk}\,|0\rangle^{\otimes(k-1)} = \frac{1}{\sqrt{n}}\sum_{i=0}^{n-1}|i\rangle,$$

where U_{nk} is constructed from one qubit gates and controlled Hadamard gates.

Proof. We proceed by the principles of mathematical induction. First, note the base case of our induction is trivial, so consider $k \in \mathbb{N}_0$ and $n \in \mathbb{N}$ satisfying $2^k < n \leq 2^{k+1}$. For $\theta \in [0, 2\pi)$ we denote by

$$R_y(\theta) = \begin{pmatrix} \cos\left(\frac{\theta}{2}\right) & -\sin\left(\frac{\theta}{2}\right) \\ \sin\left(\frac{\theta}{2}\right) & \cos\left(\frac{\theta}{2}\right) \end{pmatrix}, \qquad \mathrm{id}_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

and note that there exists a $\theta \in [0, 2\pi)$ such that

$$\left(R_y(\theta) \otimes \mathrm{id}_2^{\otimes k}\right)|0\rangle^{\otimes(k+1)} = \left(\sqrt{\frac{2^k}{n}}\,|0\rangle + \sqrt{\frac{n-2^k}{n}}\,|1\rangle\right) \otimes |0\rangle^{\otimes k}.$$

Denote by $N = n - 2^k \in \mathbb{N}$ and choose $K \in \mathbb{N}_0$ such that $2^{K-1} < N \leq 2^K$. Conditional on the first qubit being in state $|0\rangle$, we apply an Hadamard gate to each of the remaining

k qubits. Conditional on the first qubit being in state $|1\rangle$, we apply $U_{N,K}$ to the last K qubits. This produces the resulting state

$$\frac{1}{\sqrt{n}}\,|0\rangle \otimes \sum_{i=0}^{2^k-1} |i\rangle + \frac{1}{\sqrt{n}}\,|1\rangle \otimes \sum_{i=0}^{N-1} |i\rangle = \frac{1}{\sqrt{n}} \sum_{i=0}^{n-1} |i\rangle\,,$$

which proves the desired statement. □

As a side note, we may observe that for $m, n \in \mathbb{N}$ satisfying $n = 2^m$, a straight-forward calculation yields

$$H^{\otimes m}\,|0\rangle^{\otimes m} = \frac{1}{\sqrt{n}} \sum_{i=0}^{n-1} |i\rangle\,,$$

which in this special case gives an exceedingly simple way to do this first step towards preparing $|\rho_x\rangle$ of (4.3). With Lemma 69 in place, we simply need to encode the bits of $x \in \{0,1\}^n$ as a relative phase. In order to do so, we take a step back and consider an arbitrary permutation $\pi\colon \{1,\dots,n\} \to \{1,\dots,n\}$. Next, we denote by $U_\pi \in \mathrm{U}(E)$ the unitary transformation given by

$$U_\pi\,|x\rangle \coloneqq |\pi(x)\rangle\,.$$

As we will see now, implementing the unitary U_π for a particular permutation π will allow us to prepare the state $|\rho_x\rangle$ in (4.3).

Lemma 70. Let $n \in \mathbb{N}$ and $x \in \mathcal{X}_n = \{0,1\}^n$. Let $\pi_x\colon \{1,\dots,n\} \times \{0,1\} \to \{1,\dots,n\} \times \{0,1\}$ be given by

$$\pi_x(i,b) = (i, x_i \oplus b)\,.$$

Then π_x is a permutation for all $x \in \mathcal{X}$, and

$$U_{\pi_x}\left(\frac{1}{\sqrt{n}} \sum_{i=1}^{n} |i\rangle\,|-\rangle\right) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} (-1)^{x_i}\,|i\rangle\,|-\rangle\,. \tag{4.4}$$

Proof. To see π_x is a permutation, we may note that

$$\pi_x(i,b) = \pi_x\left(i', b'\right)$$

if and only if $i = i'$ and $x_i \oplus b = x_{i'} \oplus b'$, which in turn also implies $b = b'$. This shows π_x is injective, and hence bijective.

Applying U_{π_x} to each term in the sum in (4.4) yields

$$U_{\pi_x}\,|i\rangle \otimes |-\rangle = |\pi_x(i,0)\rangle - |\pi_x(i,1)\rangle = |i\rangle\,|x_i\rangle - |i\rangle\,|\overline{x}_i\rangle = (-1)^{x_i}\,|i\rangle\,|-\rangle\,,$$

which proves the desired statement. □

It has been shown that all permutations of bit strings $\pi\colon \{0,1\}^m \to \{0,1\}^m$ can be implemented by *NOT*, *CNOT* and *CCNOT* [10], and so any quantum hardware supporting these three quantum gates will allow us to prepare the state ρ_x for any $x \in \{0,1\}^n$. Furthermore, it has been shown that all permutations of $\{0,1\}^3$ can be implemented with at most 8 gates, and the circuits can be found in reasonable time [40].

In the following, we shall refer to the measurement Λ_f for $f \in \mathcal{F}_{\frac{1}{2},n}$ performed in the proof of Proposition 67 as the *matching measurement*. We will show that implementing the unitary U_{π_f} for a particular permutation π_f of $\{0,1\}^m$ is sufficient in order to implement a protocol, which outputs measurement statistics equivalent to those output by the matching measurement.

Lemma 71. Let $n \in \mathbb{N}$, and suppose $n = 2^m$ for some $m \in \mathbb{N}$. For a matching measurement Λ_{f_M} corresponding to some matching M, let

$$M = \Big(\{i_1, j_1\}, \ldots, \{i_{\frac{1}{2}n}, j_{\frac{1}{2}n}\} \Big),$$

and denote by $\pi_M \colon \{0,1\}^m \to \{0,1\}^m$ the permutation given by

$$\pi_M(\mathrm{bin}(i_k)) = \mathrm{bin}(k)\,0, \qquad \pi_M(\mathrm{bin}(j_k)) = \mathrm{bin}(k)\,1.$$

There exists a measurement implemented by U_{π_M}, the Hadamard gate, and a measurement in the computational basis, which provide statistics equivalent to those of the matching measurement Λ_{f_M}.

Proof. Consider $x \in \mathcal{X}_n$ and note

$$|\rho_x\rangle = \frac{1}{\sqrt{n}} \sum_{i=0}^{n-1} (-1)^{x_i} |i\rangle = \frac{1}{\sqrt{n/2}} \sum_{k=0}^{n/2-1} \frac{1}{\sqrt{2}} \Big((-1)^{x_{i_k}} |i_k\rangle + (-1)^{x_{j_k}} |j_k\rangle \Big).$$

Applying the unitary U_{π_M} yields the state

$$U_{\pi_M} |\rho_x\rangle = \frac{1}{\sqrt{n/2}} \sum_{k=0}^{n/2-1} |k\rangle \frac{1}{\sqrt{2}} \Big((-1)^{x_{i_k}} |0\rangle + (-1)^{x_{j_k}} |1\rangle \Big).$$

If we subsequently apply $\mathrm{id}^{\otimes(m-1)} \otimes \mathcal{H}$, we obtain

$$\frac{1}{\sqrt{n/2}} \sum_{k=0}^{n/2-1} (-1)^{x_{i_k}} |k\rangle |x_{i_k} \oplus x_{j_k}\rangle .$$

Finally, measuring in the computational basis gives the desired result, namely,

$$\frac{1}{n/2} \sum_{k=0}^{n/2-1} |k\rangle\langle k| \otimes |x_{i_k} \oplus x_{j_k}\rangle\langle x_{i_k} \oplus x_{j_k}| = \frac{1}{n/2} \sum_{k=0}^{n/2-1} |k\rangle\langle k| \otimes |f_M(x)_k\rangle\langle f_M(x)_k| ,$$

which proves the desired statement. □

As mentioned in the preliminaries, it is possible to implement U_π for any permutation π of the computational basis vectors simply by using NOT-gates, $CNOT$-gates, and $CCNOT$-gates [10]. Just as we remarked upon previously, it is for small values of $m \in \mathbb{N}$ feasible to implement the unitary U_π for all permutations π of $\{0,1\}^m$ [41]. Naturally, this implies that it is feasible to implement the matching measurement, as the only additional gate employed is a single Hadamard gate.

A special protocol for state preparation

In the following, we shall consider an alternative implementation of the state preparation protocol, where the demanding step is again encoding the bits of $x \in \{0,1\}^n$ as a relative phase in the n-level quantum system E. As before we restrict ourselves to $n = 2^m$ for some $m \in \mathbb{N}$ as this corresponds to running the protocol on m qubits.

Proposition 72. Let $n \in \mathbb{N}$ and let $x \in \mathcal{X}_n$. The transformation

$$\frac{1}{\sqrt{n}} \sum_{i=1}^{n} |i\rangle \mapsto |\rho_x\rangle = \frac{1}{\sqrt{n}} \sum_{i=0}^{n-1} (-1)^{x_i} |i\rangle$$

can be done with at most $2n-1$ gates, where $n-1$ gates are $CNOT$ gates and n gates are single qubit rotations.

Proof. The diagonal unitary $U \in \mathrm{U}(E)$ given by

$$U = \begin{pmatrix} (-1)^{x_1} & 0 & \dots & 0 \\ 0 & (-1)^{x_2} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & (-1)^{x_n} \end{pmatrix}$$

yields the desired state, and it has been shown that this can be done with at most $2n-1$ gates [42], more precisely using $n-1$ $CNOT$ gates and n single qubit rotations. □

Although the implementation above is indeed relatively simple, the number of gates required to prepare the state $|\rho_x\rangle$ is exponential in the number of qubits $m \in \mathbb{N}$. There are slight optimizations of the circuit in Proposition 72 allowing us to reduce the number of gates [41], however, even with this approach will still require a number of gates exponential in the number of qubits in most situations. We refer to Appendix B to review the code for constructing the relevant circuit given a string $x \in \mathcal{X}_n$.

4.3 Upper Bounds on Classical Strategies

As we saw in the previous section, sampling $x \in \mathcal{X}_n$ for $n \in \mathbb{N}$ and restricting the eavesdropper to an n-level quantum system E does not impact her ability to infer at least one bit of the string $x' = f(x)$ when given $f \in \mathcal{F}_{\frac{1}{2},n}$. In the following we consider the classical scenario, where Eve is allowed to store a letter from an n-letter alphabet $\mathcal{E}$, that is, $|\mathcal{E}| = n$, after observing the string $x \in \mathcal{X}_n$. The resulting state is given by a cc-state (classical-classical state) $\rho_{cc}^{XE} \in \mathcal{D}(XE)$, that is,

$$\rho_{cc}^{XE} = \sum_{x \in \mathcal{X}_n} p_x |x\rangle\langle x|^X \otimes \sum_{e \in \mathcal{E}} p_{e|x} |e\rangle\langle e|^E ,$$

where $\{|e\rangle\}_{e \in \mathcal{E}}$ is an orthonormal basis of E, and $p_{e|x}$ is the probability of Eve storing $e \in \mathcal{E}$ conditional on observing $x \in \mathcal{X}_n$.

As previously, we shall focus our attention on the situation of the alphabet $\mathcal{X}_n$ given by $\mathcal{X}_n = \{0,1\}^n$, and the probability distribution $(p_x)_{x\in\mathcal{X}_n}$ being uniform, that is, $p_x = \frac{1}{2^n}$. The state resulting from applying a particular α-matching function $f \in \mathcal{F}_{\alpha,n}$ is given by

$$f\left(\rho_{cc}\right) = \sum_{x\in\mathcal{X}_n} |f\left(x\right)\rangle\langle f\left(x\right)| \otimes \sum_{e\in\mathcal{E}} p_{e|x}\, |e\rangle\langle e|^E$$

and so the state resulting from the α-matching protocol is given by

$$\mathcal{F}_{\alpha,n}\left(\rho_{cc}\right) = \frac{1}{|\mathcal{F}_{\alpha,n}|} \sum_{f\in\mathcal{F}_{\alpha,n}} f\left(\rho_{cc}\right) \otimes |f\rangle\langle f|^F .$$

We will now consider upper bounds on the performance of any classical strategy for inferring one bit of the resulting string $x' = f\left(x\right)$ when given $f \in \mathcal{F}_{\frac{1}{2},n}$.

4.3.1 An Asymptotic Upper Bound

We will present the matter of concern in the terminology of randomness extraction. The purpose of randomness extraction is to obtain an almost uniformly random string given an imperfect source of randomness using only a small amount of randomness, which we shall refer to as a *random seed*. A map achieving this task is referred to as a *randomness extractor*. Furthermore, we are interested in the additional property that the joint distribution of the resulting string and the random seed is approximately uniformly random, which is formalized in the definition below.

Definition 73. Let $\mathcal{X}, \mathcal{X}'$ be alphabets, let $k \in \mathbb{N}$ and $\varepsilon > 0$. A (k,ε)-strong extractor is a family $\mathcal{F}$ of functions $f\colon \mathcal{X} \to \mathcal{X}'$ with a corresponding probability distribution $\left(p_f\right)_{f\in\mathcal{F}}$ such that

$$\Delta\left(X'|EF\right)_{\mathcal{F}(\rho_{cc})} \leq \varepsilon$$

for all cc-states ρ_{cc} with $\mathrm{H}_{\min}\left(X|E\right)_{\rho_{cc}} \geq k$.

If we interpret the definition above in a cryptographic setting, it corresponds to assuming the eavesdropper Eve has some knowledge of the distribution of X, so the distribution of X is not uniform from Eve's perspective. It has been shown [2] that the family of α-matching functions of $n \in \mathbb{N}$ letters $\mathcal{F}_{\alpha,n}$ equipped with the uniform probability distribution is a strong randomness extractor for $\alpha \in \left(0, \frac{1}{4}\right]$.

Theorem 74. Let $n \in \mathbb{N}$. There exists a constant $\gamma > 0$ such that for all $\varepsilon > 0$, $\alpha \in \left(0, \frac{1}{4}\right]$ and any cc-state $\rho_{cc}^{XE} \in \mathcal{D}\left(XE\right)$ with $\mathrm{H}_{\min}\left(X|E\right)_{\rho_{cc}} \geq n - \gamma\varepsilon\sqrt{n/\alpha}$, we have

$$\Delta\left(X'|EF\right)_{\mathcal{F}_{\alpha,n}(\rho_{cc})} \leq \varepsilon.$$

The statement in Theorem 74 above implies that the α-matching protocol gives rise to a $\left(n - \gamma\varepsilon\sqrt{n/\alpha}, \varepsilon\right)$-strong extractor for $\alpha \in \left(0, \frac{1}{4}\right]$. This is particularly interesting as we saw in the previous section that $\mathcal{F}_{\alpha,n}$ is *not* a strong extractor against with respect to a quantum eavesdropper with information encoded in the state of an n-level quantum system.

As mentioned above we will be particularly interested in the situation of $\mathcal{E}$ being an n-letter alphabet, that is, $|\mathcal{E}| = n$. To simplify the analysis, we note that the requirement imposed on the conditional min-entropy above can be achieved by a simple bound on the size of the alphabet $\mathcal{E}$ that Eve uses to store her side information.

Corollary 75. There exists a constant $\gamma > 0$ such that for all $\varepsilon > 0$, $\alpha \in \left(0, \frac{1}{4}\right]$ and any cc-state $\rho_{cc}^{XE} \in \mathcal{D}(XE)$ with $\rho_{cc}^{X} = \omega^{X}$ and $|\mathcal{E}| \leq 2^{\gamma\varepsilon\sqrt{n/\alpha}}$, we have

$$\Delta\left(X' \middle| EF\right)_{\mathcal{F}_{\alpha,n}(\rho_{cc})} \leq \varepsilon.$$

Proof. By the chain rule of the min-entropy we have

$$\mathrm{H}_{\min}(X|E)_{\rho_{cc}} \geq \mathrm{H}(X) - \log|\mathcal{E}| \geq n - \gamma\varepsilon\sqrt{n/\alpha},$$

and so the result follows from Theorem 74. □

In order to adapt the upper bound established in Corollary 75 to our scenario, namely, guessing one bit of $x' = f(x)$ for $f \in \mathcal{F}_{\alpha,n}$, we prove the following result.

Corollary 76. There exists a constant $\gamma > 0$ such that for all $\varepsilon > 0$, $\alpha \in \left(0, \frac{1}{4}\right]$ and any cc-state $\rho_{cc}^{XE} \in \mathcal{D}(XE)$ with $\rho_{c}^{X} = \omega^{X}$ and $|\mathcal{E}| \leq 2^{\gamma\varepsilon\sqrt{n/\alpha}}$, we have

$$\mathrm{Pr}_{\text{guess-1-bit}}\left(X' \middle| EF\right)_{\mathcal{F}_{\alpha,n}(\rho_{cc})} - \frac{1}{2} \leq \varepsilon.$$

Proof. Suppose the optimal strategy for correctly guessing one bit of $x' = f(x)$ given access to system EF succeeds with probability $\frac{1}{2} + p$ for some $p \geq 0$. We will employ this strategy in order to devise a strategy for distinguishing $\mathcal{F}_{\alpha,n}(\rho_{cc})$ from $\omega^{X'} \otimes \rho_{cc}^{EF}$.

Now, suppose we are presented with either $\mathcal{F}_{\alpha,n}(\rho_{cc})$ or $\omega^{X'} \otimes \rho_{cc}^{EF}$ with equal probability. Employing the optimal strategy for correctly guessing one bit of the state encoded in system X' given access to system E, we get a bit $b \in \{0,1\}$ and an index $k \in \{1,\ldots,n\}$ corresponding to guessing the kth bit of x is b. In system X' we may simply read the kth bit of x', namely, x'_k. Our strategy for distinguishing $\mathcal{F}_{\alpha,n}(\rho_{cc})$ and $\omega^{X'} \otimes \rho_{cc}^{EF}$ is now the following

- If $b = x'_k$, then our guess is $\mathcal{F}_{\alpha,n}(\rho_{cc})$.
- If $b \neq x'_k$, then our guess is $\omega^{X'} \otimes \rho_{cc}^{EF}$.

Finally, to evaluate our probability of success in distinguishing $\mathcal{F}_{\alpha,n}(\rho_{cc})$ or $\omega^{X'} \otimes \rho_{cc}^{EF}$, we note that given $\mathcal{F}_{\alpha,n}(\rho_{cc})$ our probability of success is $\frac{1}{2} + p$, and given $\omega^{X'} \otimes \rho_{cc}^{EF}$ our probability of success is $\frac{1}{2}$. As ρ_{cc}^{XE} and $\omega^{X} \otimes \rho_{cc}^{E}$ are equally likely, this strategy succeeds with probability $\frac{1}{2} + \frac{1}{2}p$.

As the total variational distance between two probability distributions equals the bias when trying to distinguish them using an optimal strategy, it follows that

$$\mathrm{Pr}_{\text{guess-1-bit}}\left(X' \middle| EF\right)_{\mathcal{F}_{\alpha,n}(\rho_{cc})} - \frac{1}{2} \leq \Delta\left(X' \middle| EF\right)_{\mathcal{F}_{\alpha,n}(\rho_{cc})} \leq \varepsilon,$$

which proves the desired statement. □

In Corollary 76 we have established an upper bound on the probability of correctly guessing one bit in the string $x \in \mathcal{X}_n$. From a practical point of view the statement in Corollary 76 is, however, not very useful. The universal constant $\gamma > 0$ in Corollary 75 can optimally be chosen $\gamma \approx 0.0752$ [43], and this implies that for small $\varepsilon > 0$ we need large values of

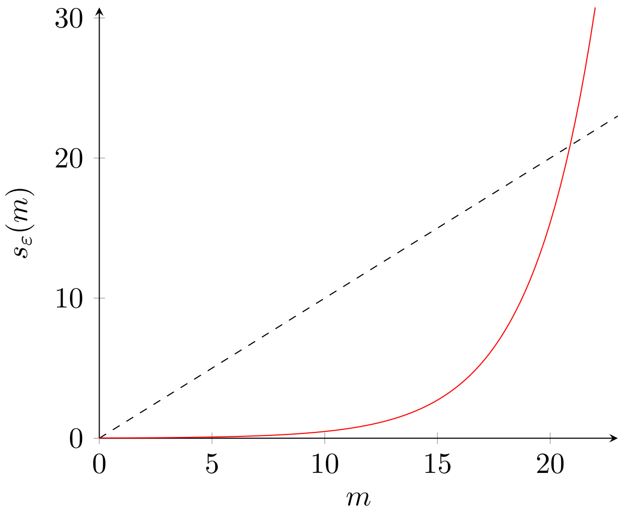


Figure 4.2: The red graph corresponds to $s_\varepsilon(m) = 2\gamma\varepsilon\sqrt{2^m}$ for $\varepsilon = 0.1$, and the dashed line is the graph of the function $m \mapsto m$.

$n \in \mathbb{N}$ to fulfill the assumptions of the statement, namely, $|\mathcal{E}| = n \leq 2^{\gamma\varepsilon\sqrt{n/\alpha}}$. Consider $m, n \in \mathbb{N}$ with $n = 2^m$, that is, m is the number of bits we can use to encode classical side information. In Figure 4.2 we illustrate for $\varepsilon = 0.1$ the length of strings n needed in order to infer the upper bound

$$\Pr_{\text{guess-1-bit}}\left(X' \middle| EF\right)_{\mathcal{F}_{\alpha,n}(\rho_{cc})} \leq 0.60$$

for $\alpha \in (0, \alpha)$. The result in Corollary 76 only applies for $n \approx 2^{22}$, that is, Eve is using 22 bits of storage. However, we may note that for, say $\varepsilon = 0.025$, which corresponds to

$$\Pr_{\text{guess-1-bit}}\left(X' \middle| EF\right)_{\mathcal{F}_{\alpha,n}(\rho_{cc})} \leq 0.525,$$

we "only" need to consider strings of length $n \approx 2^{26}$, that is, 26 bits of storage. In other words, once we reach approximately 25 (qu)bits of storage, then we can ensure the probability of success for a classical eavesdropper decreases rapidly.

Unfortunately, it is an unrealistic expectation of current quantum hardware to run complicated circuits on n-level quantum systems for such large values of $n \in \mathbb{N}$. Furthermore, it is not a trivial task to obtain efficient circuits to implement the quantum protocol [41]. To improve upon the upper bounds on the performance of classical eavesdropper Eve for small values of $n \in \mathbb{N}$, we will in the following upper bound the performance of a classical eavesdropper for small values of $n \in \mathbb{N}$ by simply evaluating Eve's average performance. For $\alpha = \frac{1}{2}$ an eavesdropper can always store the parity of the sum of all x_i for $i \in \{1, \ldots, n\}$ and this equals the parity of the sum of all z_j for $j \in \{1, \ldots, n/2\}$, so Theorem 74 does not generalize to $\alpha = \frac{1}{2}$. The choice of $\alpha \leq \frac{1}{4}$ in our exposition is, however, arbitrary, and α can be chosen arbitrarily close to $\frac{1}{2}$. Thus, the upper bounds for $\alpha \in \left(0, \frac{1}{2}\right)$ suggest that the probability of success of a classical eavesdropper for our particular task is still relatively low.

4.3.2 Upper Bounds on the Performance of a Classical Eavesdropper

Consider $n \in \mathbb{N}$ and choose $\alpha = \frac{1}{2}$ in the matching protocol. We approach the task of upper bounding the success probability of any classical strategy naively, that is, our first effort will be to evaluate the success probability of all possible classical strategies. Recall the task at hand is to guess one bit of the resulting string $x' = f(x) \in \mathcal{X}'_{\frac{1}{2}n}$ given the matching function f and side information E.

Any classical strategy consists of two components, namely, storing information depending on the observed string $x \in \mathcal{X}_n$, and choosing a strategy for guessing one bit of the string $x' = f(x) \in \mathcal{X}_{\alpha n}$ given the stored information $e \in \mathcal{E}$ and the matching function f. Due to convexity, it follows that there exists an optimal strategy, which is deterministic. Upon observing the string $x \in \mathcal{X}_n$, Eve may choose to store a letter of the alphabet $\mathcal{E}$; the strategy for storage of information is thus equivalent to choosing a function $s\colon \mathcal{X}_n \to \mathcal{E}$. Given a storage function s and a matching function f, we can actually identify an optimal strategy for guessing one bit of the string $x' = f(x)$.

Proposition 77. Let $\mathcal{X}_n = \{0,1\}^n$ and $\mathcal{E}$ be an alphabet, and consider a cc-state $\rho_{cc}^{XE} \in \mathcal{D}(XE)$ given by

$$\rho_{cc}^{XE} = \frac{1}{2^n} \sum_{x \in \mathcal{X}_n} |x\rangle\langle x|^X \otimes \sum_{e \in \mathcal{E}} p_{e|x} \, |e\rangle\langle e|^E ,$$

where $p_{e|x}$ denotes the probability of Eve storing $e \in \mathcal{E}$ upon observing $x \in \mathcal{X}_n$. After applying the matching protocol the state is

$$\mathcal{F}_{\frac{1}{2},n}(\rho_{cc}) = \frac{1}{|\mathcal{F}_{\frac{1}{2},n}|} \sum_{f \in \mathcal{F}_{\frac{1}{2},n}} f(\rho_{cc}) \otimes |f\rangle\langle f|^F .$$

If we denote by $p_{bk|fe}$ the probability of observing the bit b at the kth position of $x' \in \mathcal{X}'_{\frac{1}{2}n}$ given the matching $f \in \mathcal{F}_{\frac{1}{2}n}$ and the side information $e \in \mathcal{E}$, then the probability of correctly guessing one bit of z given the matching and side information is

$$\frac{1}{|\mathcal{F}_{\frac{1}{2},n}|} \sum_{f \in \mathcal{F}_{\frac{1}{2},n},\, e \in \mathcal{E}} p_e \max_{\substack{b \in \{0,1\}, \\ k \in \{1,\ldots,n/2\}}} p_{bk|fe}.$$

Proof. In general, a classical strategy for guessing one bit of x' based on the information stored in systems FE is given by

$$\lambda = \left(\lambda_{bk\,|\,fe}\right)_{b,k,f,e} \qquad \text{where } 0 \leq \lambda_{bk\,|\,fe}, \sum_{\substack{b \in \{0,1\}, \\ k \in \{1,\ldots,n/2\}}} \lambda_{bk\,|\,fe} = 1,$$

where the operational interpretation of $\lambda_{bk\,|\,fe}$ is the probability of Eve making the guess b for the kth bit of x', when she observes $f \in \mathcal{F}_{\frac{1}{2}n}$, $e \in \mathcal{E}$. The maximal probability of guessing a bit of x' correctly is thus given by

$$\max_{\lambda} \lambda^T p = \max_{\lambda} \sum_{\substack{b \in \{0,1\},\, k \in \{1,\ldots,n/2\}, \\ f \in \mathcal{F}_{\frac{1}{2},n},\, e \in \mathcal{E}}} \lambda_{bk|fe} p_{bk|fe} p_{fe}.$$

The optimal strategy for guessing one bit of x' given a matching function f and the information encoded in the state of system E is to simply select the most likely observation, which yields

$$\max_{\lambda} \lambda^T p = \frac{1}{\left|\mathcal{F}_{\frac{1}{2},n}\right|} \sum_{f \in \mathcal{F}_{\frac{1}{2},n},\, e \in \mathcal{E}} p_e \max_{\substack{b \in \{0,1\},\\ k \in \{1,\ldots,n/2\}}} p_{bk|fe},$$

where we have additionally noted that the matching is sampled uniformly at random independent of Eve's information encoded in system E. □

Identifying an optimal storage strategy for small values of $n \in \mathbb{N}$

The observation that the optimal strategy for guessing one bit of x' given the matching function f and the side information e is simply guessing for the most likely combination of bit b and index k of x' naturally reduces the overall number of strategies to the number of storage strategies. From now on we will focus exclusively on the storage functions $s\colon \{0,1\}^n \to \{1,\ldots,n\}$, and we denote by S_n the set of all such storage functions. It is clear that the size of S_n increases rapidly as n increases, more precisely $|S_n| = n^{2^n}$, and for even small values of $n \in \mathbb{N}$ we have

n	2	3	4
$\lvert S_n\rvert$	16	6,561	4,294,967,296

Checking all possible $|S_n| = n^{2^n}$ choices of a storage function is infeasible for $n \geq 4$, but as we shall see in the following we may reduce this number notably.

As we have identified Eve's optimal strategy for guessing one bit of x' when the matching function f and side information e is given, we note that this strategy can be heuristically explained as follows: Upon observing side information $e \in \mathcal{E}$, Eve computes the preimage of e under her storage function s, that is, $s^{-1}(\{e\})$. Given the matching function f, she computes the strings $x' = f(x)$ for each $x \in s^{-1}(\{e\})$. Observing all the possible strings x', she finally chooses the index k, where most strings x' agree, and outputs as b the typical value of x'_k. As this strategy utilizes only the partition of $\mathcal{X}_n$ that a storage function naturally gives rise to, namely,

$$\mathcal{X}_n = \bigcup_{e \in \mathcal{E}} f^{-1}(\{e\}),$$

we only have to consider all the various partitions of $\mathcal{X}_n$ in at most $|\mathcal{E}|$ parts. Without loss of generality, we may assume Eve's strategy uses all letters of $\mathcal{E}$, so we will consider the number of partitions of $\mathcal{X}_n$ into exactly $|\mathcal{E}|$ parts. Again, we focus on the situation of $|\mathcal{E}| = n$, we may denote by T_n the set of all partitions of $\mathcal{X}_n$ in at n parts, and now we have

n	2	3	4	5
$\lvert T_n\rvert$	7	966	171,798,901	193,257,076,459,811,283,150

Evidently, this is a significant improvement, however, the number of partitions in n parts of $\mathcal{X}_n$ still increases rapidly as n increases. As further improvement is needed, we direct the attention of the reader to the observation that $f(x) = f(\overline{x})$ for all $f \in \mathcal{F}_{\frac{1}{2},n}$, where $\overline{x}$ denotes the string resulting from flipping all bits of x.

Lemma 78. Let $\mathcal{X}_n$ and $\mathcal{E}$ be alphabets. There exists an optimal storage function $s\colon \mathcal{X}_n \to \mathcal{E}$ such that

$$s(x) = s(\overline{x})$$

for all $x \in \mathcal{X}_n$.

Proof. Suppose $s\colon \mathcal{X}_n \to \mathcal{E}$ is an optimal choice of storage function. Let $y \in \mathcal{X}_n$ and define $t\colon \mathcal{X}_n \to \mathcal{E} \cup \{\perp\}$ by

$$t(x) = s(x), \qquad t(y) = t(\overline{y}) = \perp$$

for $x \in \mathcal{X}_n \setminus \{y, \overline{y}\}$. We may think of t as a storage function with the added possible value $\perp$, and we may combine this with the optimal guessing strategy from Proposition 77 except that on observing $\perp$ the guessing protocol aborts resulting in a failed attempt at guessing. Suppose we use storage function t, and let p^t_e denote the probability of observing $e \in \mathcal{E}$, and let $p^t_{bk|fe}$ denote the probability of b being the kth bit of x' given a matching function f and stored information $e \in \mathcal{E}$. For each $e \in \mathcal{E}$ we may define

$$t_e(x) = t(x), \qquad t_e(x_0) = e, \qquad t_e(\overline{x}_0) = \perp,$$

and choose the storage function t_{e_0} with the highest probability of guessing one bit of x'. Relative to t, the added probability of success in using the storage function t_{e_0} is

$$\frac{1}{2^n |\mathcal{F}_{\frac{1}{2},n}|} \sum_{f \in \mathcal{F}_{\frac{1}{2},n}} \max_{\substack{b \in \{0,1\},\\ k \in \{1,\ldots,n/2\}}} \Bigg(\left|\left\{x \in t_{e_0}^{-1}(\{e_0\}) \,\middle|\, f(x)_k = b\right\}\right| \\ - \left|\left\{x \in t^{-1}(\{e_0\}) \,\middle|\, f(x)_k = b\right\}\right| \Bigg),$$

where we have used that f is deterministic. Again, define for each $e \in \mathcal{E}$ the storage function $t_{e_0 e}\colon \mathcal{X}_n \to \mathcal{E}$ by

$$t_{e_0 e}(x) = t(x), \qquad t_{e_0 e}(x_0) = e_0, \qquad t_{e_0 e}(\overline{x}_0) = e,$$

and note that using the storage function $t_{e_0 e_0}$ increases the probability of success by

$$\frac{1}{2^n |\mathcal{F}_{\alpha,n}|} \sum_{f \in \mathcal{F}_{\alpha,n}} \max_{\substack{b \in \{0,1\},\\ k \in \{1,\ldots,n/2\}}} \Bigg(\left|\left\{x \in t_{e_0 e_0}^{-1}(\{e_0\}) \,\middle|\, f(x)_k = b\right\}\right| \\ - \left|\left\{x \in t_{e_0}^{-1}(\{e_0\}) \,\middle|\, f(x)_k = b\right\}\right| \Bigg).$$

Since $f(y) = f(\overline{y})$ for all $f \in \mathcal{F}_{\alpha,n}$, this increase is lower bounded by the previous increase, which was chosen optimally. This implies $t_{e_0 e_0}$ is an optimal choice of storage function among the candidates $t_{ee'}$, hence at least as good as s. Repeating this process for all $x \in \mathcal{X}_n$ yields the desired result. □

If we denote by T^*_n the set of partitions of $\mathcal{X}_n$ into exactly n parts, where $x, \overline{x}$ are elements of the same part for all $x \in \mathcal{X}_n$, then Lemma 78 ensures an optimal partition is among the elements of T^*_n. Additionally, this limits the search space for an optimal partition significantly as illustrated in the table below.

n	2	3	4	5
$\lvert T_n^* \rvert$	1	6	1.701	1,096,190,550

We may interject that the rate of growth of $|T_n^*|$ for increasing $n \in \mathbb{N}$ is, however, not improved by much; the observation that we may assume $x, \overline{x}$ are always in the same part of an optimal partition of $\mathcal{X}_n$ merely reduces the counting problem of all partitions of a set of size 2^n to a set of size 2^{n-1}. Nevertheless, this reduction does allow us to identify optimal storage functions $s_n \colon \mathcal{X}_n \to \{0, 1, \ldots, n-1\}$ for small values of $n \in \mathbb{N}$.

For $n = 2$ the storage function $s_2 \colon \mathcal{X}_2 \to \{0, 1\}$ given by

x	00	01	10	11
$s_2(x)$	0	1	1	0

allows Eve to infer one bit of $f(x)$ with certainty.

For $n = 3$ the storage function $s_3 \colon \mathcal{X}_3 \to \{0, 1, 2\}$ given by

x	000	001	010	011	100	101	110	111
$s_3(x)$	0	0	1	2	2	1	0	0

is an optimal storage strategy with success probability at $83.\overline{3}\%$.

For $n = 4$ the storage function $s_4 \colon \mathcal{X}_4 \to \{0, 1, 2, 3\}$ given by

x	0000	0001	0010	0011	0100	0101	0110	0111	1000	1001	1010	1011	1100	1101	1110	1111
$s_4(x)$	0	0	1	1	2	2	3	3	3	3	2	2	1	1	0	0

allows Eve to infer one bit of $f(x)$ with certainty.

This naive approach to analyzing the classical strategies is at the current stage insufficient to show an upper bound on the performance of classical strategies for $n \geq 5$, which could potentially allow us to identify an advantage using current quantum hardware. Although our efforts have improved the computation time, we are still far from able to calculate an upper bound on the performance of an optimal classical strategy by checking all strategies for $n = 8$ corresponding to three bits. However, random sampling from all strategies has yielded classical strategies with success probability 100% for $n = 8, 16$.

We note, however, that if we generalize the structure of the three examples of optimal strategies above, then we may formulate a classical storage strategy given by the pattern described below to challenge the performance of our quantum strategy.

Definition 79. Let $n \in \mathbb{N}$. We define the *symmetric storage strategy*

$$s_{\text{sym}} \colon \mathcal{X}_n \to \{0, 1, \ldots, n-1\}$$

by listing the elements of $\mathcal{X}_n$ in lexicographic order and specifying s_{sym} on subsets as follows. Choose $q, r \in \mathbb{N}_0$ such that $2^{n-1} = q \cdot n + r$, where $0 \leq r < n$. First, we divide the numbers $0, 1, \ldots, rq + r$ into r groups of size $q + 1$ as follows

$$\operatorname{dec}(x) = \underbrace{0, \ldots, q}_{s_{\text{sym}}(x)=0}, \underbrace{q+1, \ldots, 2q+1}_{s_{\text{sym}}(x)=1}, \ldots, \underbrace{(r-1)\,q + (r-1), \ldots, rq + (r-1)}_{s_{\text{sym}}(x)=r-1}.$$

Secondly, we divide the numbers $rq+r+1,\ldots,2^{n-1}$ into $n-r$ groups of size q as follows

$$\operatorname{dec}(x)=\underbrace{rq+r,\ldots,(r+1)\,q+(r-1)}_{s_{\mathrm{sym}}(x)=r},\ldots,\underbrace{(n-1)\,q+r,\ldots,nq+(r-1)}_{s_{\mathrm{sym}}(x)=n-1}.$$

This defines $s_{\mathrm{sym}}(x)$ for all $x\in\mathcal{X}_n$ with $x_1=0$, so we may extend the definition to the entire domain by letting $s_{\mathrm{sym}}(x)=s_{\mathrm{sym}}(\overline{x})$.

To benchmark the symmetric storage strategy, we may compare it to the naïve strategy of sampling $\log n$ distinct indices $i\in\{1,\ldots,n\}$ and encoding the value of x_i in our n-level memory system. If both indices $\{i_k,j_k\}$ of at least one pair of the matching M corresponding to the matching function f_M are among the $\log n$ bits of x in the memory, then Eve can infer one bit of $x'=f_M(x)$ with certainty; otherwise she will resort to guessing at random. We refer to this strategy as the *birthday strategy* due to its reminiscence of the Birthday Paradox [44]. We assume our classical strategies are run on error-free hardware, and below we have calculated the theoretical (and essentially practical) performance of both classical strategies.

Lemma 80 (Performance of classical strategies)**.** Let $m\in\mathbb{N}$ and denote by $n=2^m$. Then

$$\Pr_{\mathrm{cl,birthday}}(\mathrm{Success})=1-2^{m-1}\cdot\binom{2^{m-1}}{m}\binom{2^m}{m}^{-1}$$

and for $m\geq 3$ we have

$$\Pr_{\mathrm{cl,sym}}(\mathrm{Success})=1-2^{m}\cdot\binom{2^{m-1}}{m+1}\binom{2^m}{m+1}^{-1}$$

Proof. Let $x\in\{0,1\}^n$. To evaluate the probability of an eavesdropper Eve guessing a bit of $f_M(x)$ given her side information $s_{\mathrm{sym}}(x)$ and the matching M, assume without loss of generality that $s_{\mathrm{sym}}(x)=0$. Then Eve knows x is among the strings

$$\begin{array}{cc}
0\ldots00\ldots00 & 1\ldots11\ldots11\\
0\ldots00\ldots01 & 1\ldots11\ldots10\\
\vdots & \vdots\\
0\ldots01\ldots11 & 1\ldots10\ldots00,
\end{array}$$

where the first $m+1$ bits are either all 0s or all 1s. If the matching M includes a pair of indices among the first $m+1$ indices, then Eve can infer a bit of $f_M(x)$ with certainty, and otherwise, her best strategy is to guess at random. The number of matchings with no pair among the first $m+1$ bits is

$$N_{\mathrm{rand}}=\frac{(n-m-1)!}{(n-2m-2)!}\cdot\frac{(n-2m-2)!}{2^{n/2-m-1}\cdot(n/2-m-1)!}=\frac{(2^m-m-1)!}{2^{2^{m-1}-m-1}\cdot(2^{m-1}-m-1)!},$$

where we have used $m\geq 3$ to ensure the positivity of the expressions in factorials. Furthermore, the total number of matchings is

$$N_{\mathrm{total}}=\frac{n!}{2^{n/2}\cdot(n/2)!}=\frac{2^m!}{2^{2^{m-1}}\cdot 2^{m-1}!}.$$

With this in mind, it follows by straightforward computation that

$$\begin{aligned}\Pr_{\text{cl,sym}}(\text{Success}) &= 1 - \frac{1}{2}\frac{N_{\text{rand}}}{N_{\text{total}}} \\ &= 1 - 2^m \cdot \frac{2^{m-1}!}{(2^{m-1}-m-1)!\,(m+1)!}\frac{(2^m-m-1)!\,(m+1)!}{2^m!} \\ &= 1 - 2^m \cdot \binom{2^{m-1}}{m+1}\binom{2^m}{m+1}^{-1}.\end{aligned}$$

The performance of the birthday strategy is determined analogously. □

4.4 Comparison of Quantum and Classical Strategies

In the previous sections, we have seen that restricting Eve to an n-level storage system barely limits a quantum eavesdropper's ability to infer one bit of the string $x' \in \mathcal{X}'_{\frac{1}{2}n}$ when given the matching function $f \in \mathcal{F}_{\frac{1}{2},n}$, while an eavesdropper with classical memory must essentially resort to random guessing for large values of $n \in \mathbb{N}$. For small values of $n \in \mathbb{N}$, we saw that there exists an optimal storage strategy, namely, the symmetric storage strategy, and in the following, we will compare the performance of the symmetric storage strategy with an actual implementation of an optimal quantum strategy. To fit the quantum hardware architecture made available by IonQ, we will restrict ourselves to $n = 2^m$ for some $m \in \mathbb{N}$, that is, we compare the performance of the quantum and classical strategies with memory restricted to m (qu)bits.

4.4.1 Implementation of Quantum Protocol

In the following, we will describe the circuits to be run on IonQ's device in order to analyze the performance of an eavesdropper Eve with quantum memory and quantum processing. We will use the result in Proposition 72 to device a circuit that prepares the desired state, and for $n = 2^m$ where $m = 1, 2$ we will use the result in Lemma 71 to extract the relevant statistics from a measurement in the computational basis for given matching functions. For $n = 2^m$ with $m = 3, 4, 5, 6$, we will restrict ourselves to analyzing the situation of a single, simple matching function. The code to generate a circuit for preparing the relevant state $\rho_x \in \mathcal{D}(E)$ given $x \in \{0,1\}^n$ is given in Appendix B.

Implementation of quantum protocol on one qubit, $n = 2^m$ with $m = 1$

For $x \in \{0,1\}^2$ there are only two states to prepare, namely,

$$|\rho_{00}\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle), \qquad |\rho_{01}\rangle = \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle),$$

which can be done by running the circuits

```
T  : |0|                  T  : |0|1|

q0 : -H-                  q0 : -H-Z-

T  : |0|                  T  : |0|1|
```

There is only one matching function $f\colon \{0,1\}^2 \to \{0,1\}$, and implementing this such that a final measurement in the computational basis corresponds to Eve's guess of $x' = f(x)$ corresponds to adding an Hadamard gate.

To estimate the performance on 1 qubit, we have implemented the two combinations of the preparation of $|\rho_x\rangle$ for $x = 00, 01$ and measurement Λ_{f_M} for $M = (\{1,2\})$. Each implementation is run 500 times.

Implementation of quantum protocol on two qubits, $n = 2^m$ with $m = 2$

For $x \in \{0,1\}^2$ there are eight states to prepare. For example, the circuit for preparing $|\rho_x\rangle$ with $x = 0010$ is given by

```
T  : |0|    |    1    |    2    |    3    |    4|

q0 : -H-Rz(-1.57)-C-----------------C-
                  |                 |
q1 : -H-Rz(1.57)--X-Rz(-1.57)-X-

T  : |0    |    1    |    2    |    3    |    4|
```

and the circuits for preparing the remaining $|\rho_x\rangle$ are of similar structure [42, 41]. The code for generating the relevant circuits can be found in Appendix B.

There are three choices of a matching function f_M with matchings M given by

$$M_0 = (\{0,1\},\{2,3\}), \quad M_1 = (\{0,2\},\{1,3\}), \quad M_2 = (\{0,3\},\{1,2\}).$$

Inspired by Lemma 71, we implement the corresponding measurements by permuting the computational basis elements, such that after a measurement in the computational basis, then the first bit of the measurement outcome tells Eve which bit of $x' = f(x)$ to make a guess at, and the second bit tells her what to guess for. This is implemented by the three circuits below

```
T : |0|          T : |0   |  1|        T : |0|1|

q0 : ---         q0:-SWAP----          q0 : -X---
                      |                       |
q1 : -H-         q1:-SWAP--H-          q1 : -C-H-

T : |0|          T : |0   |  1|        T : |0|1|
```

To estimate the performance on 2 qubits, we have implemented the four combinations of the preparation of $|\rho_x\rangle$ for $x = 0001, 0111$ and measurement Λ_{f_M} for

$$M = (\{1,2\}, \{3,4\}), (\{1,4\}, \{2,3\}).$$

Each implementation is run 500 times.

Implementation of quantum protocol on multiple qubits, $n = 2^m$ for $m = 3, 4, 5, 6$

For $x \in \{0,1\}^m$ there are 2^{m-1} states to prepare. For example, the circuit for preparing $|\rho_x\rangle$ with $x = 00000001$ is given by

```
T  : |0 |    1    | 2 |   3    | 4|5|   6    | 7 |    8     | 9 |   10   | 11|

q0 : -H-Rz(-0.79)-C----------C-C----------------------C-------------
                  |          | |                      |
q1 : -H-Rz(-0.79)-X-Rz(0.79)-X-|----------C-----------|----------C--
                               |          |           |          |
q2 : -H-Rz(-0.79)-------------X-Rz(0.79)-X-Rz(-0.79)-X-Rz(0.79)-X--

T  : |0 |    1    | 2 |   3    | 4|5 |  6    | 7 |    8     | 9 |   10   | 11|
```

and the circuits for preparing the remaining $|\rho_x\rangle$ are of similar structure [42, 41]. This can also be done on four, five, and six qubits. The code for generating the relevant circuits can be found in Appendix B.

For $m = 3$, we can within reasonable computing time identify an optimal circuit for implementing the measurement as in Lemma 71. Most of such implementations, however, make use of the Tofolli gate, which is not supported by IonQ, and so we will only evaluate the performance of the quantum protocol for the matching function f_M with M given by

$$M = (\{1,2\}, \{3,4\}, \{5,6\}, \{7,8\}).$$

We make a similarly simplified analysis of the performance of the quantum protocol for $m = 4, 5, 6$. Each implementation is run 500 times.

4.4.2 Comparison of Performance

Collecting the results of the implementations of the previous section, we are now in a position to compare the performance of a classical and quantum eavesdropper Eve. The table in Figure 4.3 shows that we were not able to show a real-life advantage for the eavesdropper using current quantum hardware in this particular setup.

n	2	4	8	16	32	64
$\Pr_{\text{cl,sym}}$ (Success)	100%	100%	88.57%	79.49%	71.72%	65.32%
$\Pr_{\text{cl,birthday}}$ (Success)	50.00%	66.67%	71.43%	69.23%	65.29%	61.32%
$\Pr_{\text{cl,optimal}}$ (Success)	100%	100%	100%	100%	?	?
$\Pr_{\text{q,real}}$ (Success)	99.8%	97.35%	96.93%†	94.07%†	79.45%†	64.3%†

Figure 4.3: For $n = 2, 4, 8, 16, 32$ and $n = 64$ the table shows the performance of the symmetric storage strategy, the birthday strategy, and the optimal classical strategy and the real performance of the quantum strategy from Section 4.2, respectively. For $n = 8, 16, 32$ and $n = 64$, we have tested the implementation of the quantum strategy against the most simple choice of a matching function and for only a few randomly sampled strings $x \in \{0, 1\}^n$.

As we note that there exist perfect classical strategies for $m \leq 4$, and we see that the performance of our quantum strategy, when implemented on IonQ's hardware rapidly, drops as the number of qubits m exceeds 4, it seems that we are still far from being able to apply the asymptotic bounds on success probability when using a classical memory [2]. Actually, we note that the symmetric storage strategy outperforms the quantum strategy for $m = 6$. One further point is that we have only analyzed the performance of the quantum algorithm when the matching function is chosen to minimize the demands of the quantum implementation. In addition, it is computationally heavy [41] to generate circuits for applying the correct measurements corresponding to arbitrary matchings using the approach we presented here, and this will add significantly to the circuit depth of this implementation.

Conclusion

Let us reflect upon what we have learned and what we would still like to learn from our acquaintance with our friends Xavier, Alice, and Bob, and the eavesdropper Eve.

We began in Chapter 2 by considering a two-party setting involving Xavier and Eve sharing a quantum system with classical data encoded into Xavier's system, while Eve held some quantum side information. Here, we saw that with no restrictions on the capabilities of Eve, the amount of secure information to be extracted per copy is exactly given by the conditional von Neumann entropy of the state. In the more general setting of Eve having some imperfection in her quantum memory, we saw that it is indeed possible for Xavier to extract more secure information per copy even with a deterministic strategy in certain cases. This allowed us to motivate the introduction of hidden information, and we saw that it is possible to hide information at non-zero rates. However, two questions concerning the distillation of secure information and hidden information remain open. Firstly, the exact rate at which Xavier can extract secure information with respect to a restricted eavesdropper remains undecided, since the proof of the quantum asymptotic equipartition property (QAEP) does not directly generalize to our setting; more precisely, the original proof of the QAEP [16] relies on the smoothed min-entropy being lower bounded (with some error term) by a Rényi-like conditional entropy, which is additive. When Eve is allowed to act jointly on multiple copies of the state, the additivity property is lost. Recent generalizations of the QAEP [45, 46] are similarly very restrictive when it comes to the correlations between individual rounds, and so we expect that different tools are necessary in order to prove Conjecture 10. The second question that remains open is whether the rate at which secure states can be distilled by a deterministic strategy is equal to the optimal rate using probabilistic strategies. The question arises from the adaptation of the leftover hash lemma (LHL) [17] to our setting, where the optimal strategy of Eve is suddenly no longer evident; when Eve has perfect quantum memory, it is of course optimal for her to do nothing. The proof of the LHL relies on the interplay between the trace norm, the Hilbert-Schmidt norm, the pretty good measurement, and the conditional min-entropy. Unfortunately, there is no natural analog of the pretty good measurement in a generic set of channels, so the adaptation of the deterministic LHL to our setting remains open. Both of these questions would be interesting to pursue further.

We changed gears in Chapter 3 in order to study the connections between privacy, entanglement, and data hiding. Here, we proved Proposition 45, which provides a simple lower bound on the amount of distillable key analogous to a previously shown lower bound on the distillable entanglement [1]. Combining these two results gave an indication that the amount of classical data that can be hidden in a key-correlated state may be related to the difference between the rates of distillable key and distillable entanglement. Furthermore, the lower bound in Proposition 45 gives a simple tool to lower bound the rate of a quantum key repeater in certain special cases. Taking a step back, we introduced a notion of joint phase orthogonality and saw in Theorem 54 that the properties of joint phase orthogonality and key-correlation cannot occur simultaneously without entanglement. Inspired by this, we saw an easy generalization of Proposition 45 to arbitrary states, which indicates how the amount of distillable key is lower bounded by the amount of correlation and joint phase orthogonality. Finally, we introduced a rate of data hiding of a bipartite state with direct connections back to the notion of the rate of hiding state distillation introduced in Chapter 2. We do, however, remain curious and unknowledgeable about this quantity beyond elementary observations such as it being finite. Our work on phase orthogonality was motivated by the conjecture that the difference between distillable key and distillable entanglement is described by a notion of hiding [7], but we do not expect our quantity to encapture this difference exactly. At the current level of our understanding, the potential to hide data may be necessary to see a gap between distillable private key and distillable entanglement, but it may not be sufficient.

Finally, in Chapter 4 we explored to what extent the difference in the rates of secure state distillation with respect to an unrestricted and a restricted eavesdropper is of practical relevance using current quantum hardware. On IonQ's quantum device, we have analyzed the relative performance of an eavesdropper with quantum memory and an eavesdropper with classical memory in the very particular task of guessing a bit of a certain partially secure string. As this was discussed from a theoretical perspective in Chapter 2, we saw in Example 32 that a gap between the respective performances exists theoretically, however, current quantum hardware is noisy. The theoretical advantage of a quantum eavesdropper is large for long strings, however, this is also when the quantum device becomes exceedingly noisy. Our analysis of short strings shows that even though a quantum eavesdropper may have a theoretical advantage, we cannot verify that this translates to a practical advantage using IonQ's quantum device. For further research into this particular setup as an example of a relevant task with an advantageous quantum strategy, we propose two directions. Clearly, it is necessary to improve the upper bounds on the performance of a classical eavesdropper for longer strings than was possible here. Furthermore, an efficient implementation of the unitaries that permute the computational basis vectors in terms of native gates may give sufficiently short circuits in order to achieve quantum advantage.

Looking back at the contents of Chapters 2-4, we note that understanding the advantage in securing classical data with respect to a restricted eavesdropper plays a role in understanding the difference between key distillation and entanglement distillation. More precisely, we have seen lower bounds on key distillation (resp. entanglement distillation) in terms of correlation and state discrimination (resp. local state discrimination). Finally, we saw Chapter 4 that the theoretical advantage of a quantum eavesdropper does not carry over to a practical advantage - yet!

Appendix A

Miscellaneous

Theorem 81. Let $\varepsilon > 0$ and $\alpha \in (1, 2]$. For $\rho \in \mathcal{D}(AB)$ we have

$$\mathrm{H}_{\min}^{\varepsilon}(A|B)_{\rho} \geq \mathrm{H}_{\alpha}(A|B)_{\rho} - \frac{1}{\alpha - 1} \log \frac{2}{\varepsilon^2}.$$

Proof. This was shown in [16]. □

Theorem 82 (Quantum Asymptotic Equipartition Property)**.** Let $\rho^{AB} \in \mathrm{D}(AB)$. Then

$$\lim_{\varepsilon \to 0} \lim_{n \to \infty} \frac{1}{n} \mathrm{H}_{\min}(A^n|B^n)_{\rho^{\otimes n}} = \mathrm{H}(A|B)_{\rho}.$$

Proof. This was shown in [16]. □

Lemma 83 (Leftover Hash Lemma)**.** Consider a cq-state $\rho^{XE} \in \mathcal{D}(XE)$. Let $\mathcal{X}'$ be an alphabet and let $\mathcal{F}$ be a family of functions $f \colon \mathcal{X} \to \mathcal{X}'$ with associated probability distribution $(p_f)_{f \in \mathcal{F}}$. Denote by

$$\mathcal{F}(\rho) = \sum_{x' \in \mathcal{X}'} |x'\rangle\langle x'|^{X'} \otimes \sum_{f \in \mathcal{F}} p_f \sum_{x \in f^{-1}(\{x'\})} p_x \rho_x^E \otimes |f\rangle\langle f|^F.$$

For all $\varepsilon \geq 0$ we have

$$\Delta\left(X' \middle| EF\right)_{\mathcal{F}(\rho)} \leq 2\varepsilon + \sqrt{2^{\log|\mathcal{X}'| - \mathrm{H}_{\min}^{\varepsilon}(X|E)_{\rho}}}.$$

Proof. This was shown in [17]. □

Lemma 84 (Fano's Inequality)**.** Let $\mathcal{X}$ be an alphabet and consider an ensemble $\{p_x \rho_x\}_{x \in \mathcal{X}} \subseteq \mathcal{D}(A)$. For $\varepsilon > 0$ suppose we have $\mathrm{Pr}_{\text{guess}}(X|A) \geq 1 - \varepsilon$. Then

$$\mathrm{H}(X|A)_{\rho} \leq \varepsilon \log|\mathcal{X}| + h(\varepsilon).$$

Proof. This was shown in [47]. □

Theorem 85. Let $\rho, \rho' \in \mathcal{D}(A)$, and suppose $\varepsilon \geq 0$ is given by

$$\varepsilon = \frac{1}{2}\left\|\rho - \rho'\right\|.$$

Then

$$\left|\mathrm{H}(A)_{\rho} - \mathrm{H}(A)_{\rho'}\right| \leq \varepsilon \log d_A + h(\varepsilon).$$

Proof. This was shown in [20]. □

Theorem 86. Let $\varepsilon \geq 0$ and $\rho, \rho' \in \mathcal{D}(AB)$. If

$$\frac{1}{2}\left\|\rho - \rho'\right\| \leq \varepsilon,$$

then

$$\left|\mathrm{H}\left(A|B\right)_{\rho} - \mathrm{H}\left(A\right)B_{\rho'}\right| \leq 2\varepsilon \log d_A + g\left(\varepsilon\right).$$

Proof. This was shown in [20]. □

Appendix B

Implementation of State Preparation Protocol

For a given $x \in \{0,1\}^{2^m}$ for $m \in \mathbb{N}$ the code generate a circuit implementing a unitary U satisfying

$$U \left|0\right\rangle^{\otimes m} = \frac{1}{\sqrt{2^m}} \sum_{i=0}^{2^m-1} (-1)^{x_i} \left|\operatorname{bin} i\right\rangle ,$$

where $\operatorname{bin}(i) \in \{0,1\}^m$ denotes the binary representation of $i \in \mathbb{N}$. Without any further ado, we present the generating code.

```
import math
import copy

def innerProd(xs, ys):
    n = len(xs)
    innerProd = 0
    for i in range(0,n):
        innerProd = innerProd + xs[i]*ys[i]
    return innerProd

### STATE PREPARATION: Gray Circuits ###
def findGrayCode(m):
    grayCode = [[0],[1]]
    for i in range(1,m):
        grayCodeTemp = grayCode.copy()
        grayCode0 = [bs + [0] for bs in grayCodeTemp]
        grayCodeTemp.reverse()
        grayCode1 = [bs + [1] for bs in grayCodeTemp]
        grayCode = grayCode0 + grayCode1
    return grayCode
```

```
def bitChanges(m):
    if m == 0:
        return []
    prior = bitChanges(m-1)
    return prior + [m-1] + prior

def findPhases(m, xs):
    phases = []
    for bs in someList(m):
        Arow = []
        for cs in allBinary(m):
            Arow.append((-1)**innerProd(bs, cs))
        phase = 2*math.pi*innerProd(Arow, xs)/(2**m)
        phases.append(phase)
    return phases

def someList(m):
    if m == 1:
        return [[0],[1]]
    lastHalf = findGrayCode(m-1)
    for code in lastHalf:
        code.append(1)
    firstHalf = someList(m-1)
    for code in firstHalf:
        code.append(0)
    return firstHalf + lastHalf

def grayCircuit(m, phases):
    if m == 1:
        return [['RZ',0,phases[1]]]
    halfCircuit0 = grayCircuit(m-1,phases)
    halfCircuit1 = []
    grayCode = [[1] + index for index in findGrayCode(m-1)]
    cs = bitChanges(m-1) + [m-2]
    for i in range(0,2**(m-1)):
        c = cs[i]
        halfCircuit1.append(['RZ',m-1,phases[i + 2**(m-1)]])
        halfCircuit1.append(['CX',[c],m-1])
    return halfCircuit0 + halfCircuit1
```

```
def stateCircuit(m, xs):
    phases = findPhases(m, xs)
    gates = grayCircuit(m, phases)
    circuit = Circuit()
    for i in range(0,m):
        circuit.h(i)
    for gate in gates:
        if gate[0] == 'RZ':
            t = gate[1]
            theta = gate[2]
            circuit.rz(t,theta)
        if gate[0] == 'CX':
            c = gate[1][0]
            t = gate[2]
            circuit.cnot(c,t)
    return circuit
```

Appendix C

Rate of Distillable Key from a Platypus State

The resulting state from the protocol in Example 50 is given by

$$
\begin{aligned}
\left|\Psi_\lambda\right\rangle = \frac{1}{\sqrt{2}}\left|00\right\rangle_{A_kB_k}\left|\varphi_1\right\rangle_{A_sE}\left|00\right\rangle_{B_sF} + p\left|11\right\rangle_{A_kB_k}\left|0000\right\rangle_{A_sEB_sF} \\
+ q\left|11\right\rangle_{A_kB_k}\left|01\right\rangle^{A_sE}\left(\sqrt{1-\lambda}\left|10\right\rangle + \sqrt{\lambda}\left|01\right\rangle\right)_{B_sF}.
\end{aligned}
$$

Tracing out the purifying systems EF yields the state shared by Alice and Bob, that is

$$
\begin{aligned}
\Psi_\lambda = \hat{\Psi}_\lambda &+ \frac{1}{2}\left|00\right\rangle\!\left\langle 11\right| \otimes \left(p\left|00\right\rangle\!\left\langle 00\right| + q\sqrt{1-\lambda}\left|10\right\rangle\!\left\langle 01\right|\right) \\
&+ \frac{1}{2}\left|11\right\rangle\!\left\langle 00\right| \otimes \left(p\left|00\right\rangle\!\left\langle 00\right| + q\sqrt{1-\lambda}\left|01\right\rangle\!\left\langle 10\right|\right),
\end{aligned}
$$

where $\hat{\Psi}_\lambda \in \mathcal{D}\left(A_kB_kA_sB_s\right)$ is the key-attacked state given by

$$
\begin{aligned}
\hat{\Psi}_\lambda = \frac{1}{2}\left|00\right\rangle\!\left\langle 00\right| \otimes \frac{1}{2}\left(\left|0\right\rangle\!\left\langle 0\right| + \left|1\right\rangle\!\left\langle 1\right|\right) \otimes \left|0\right\rangle\!\left\langle 0\right| + p^2\left|11\right\rangle\!\left\langle 11\right| \otimes \left|00\right\rangle\!\left\langle 00\right| \\
+ q^2\left|11\right\rangle\!\left\langle 11\right| \otimes \left|0\right\rangle\!\left\langle 0\right| \otimes \left(\lambda\left|0\right\rangle\!\left\langle 0\right| + (1-\lambda)\left|1\right\rangle\!\left\langle 1\right|\right).
\end{aligned}
$$

It follows from Lemma 44 that

$$
K_D^{\rightarrow}\left(\Psi_\lambda\right) = D\left(\Psi_\lambda \middle\| \hat{\Psi}_\lambda\right) = \mathrm{H}\left(A_kA_sB_kB_s\right)_{\hat{\Psi}_\lambda} - \mathrm{H}\left(A_kA_sB_kB_s\right)_{\Psi_\lambda},
$$

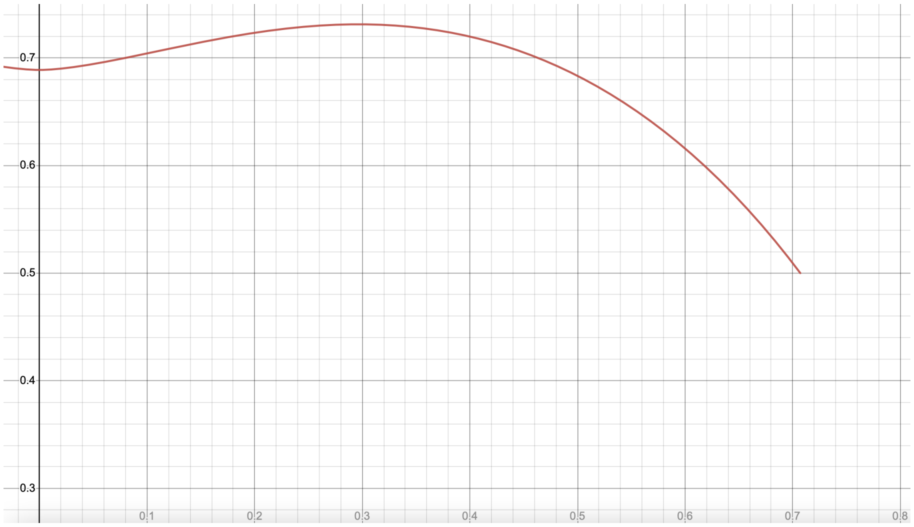


Figure C.1: We plot $K_{D}^{\rightarrow}\left(\Psi_{\lambda}\right)$ as a functions of the parameter $q \in \left[0, \frac{1}{\sqrt{2}}\right]$.

which reduces the problem to finding the eigenvalues of Ψ_λ and $\hat{\Psi}_\lambda$. This yields

$$\begin{aligned}
K_{D}^{\rightarrow}\left(\Psi_{\lambda}\right) = 1 &- \left(\frac{1}{2} - (1-\lambda)\, q^{2}\right) \log\left(\frac{1}{2} - (1-\lambda)\, q^{2}\right) \\
&- \left((1-\lambda)\, q^{2}\right) \log\left((1-\lambda)\, q^{2}\right) \\
&+ \frac{\frac{3}{4} - (1-\lambda)\, q^{2} - \sqrt{\left(\frac{3}{4} - (1-\lambda)\, q^{2}\right)^{2} - (\lambda)\, q^{2}}}{2} \\
&\qquad \log\left(\frac{\frac{3}{4} - (1-\lambda)\, q^{2} - \sqrt{\left(\frac{3}{4} - (1-\lambda)\, q^{2}\right)^{2} - (\lambda)\, q^{2}}}{2}\right) \\
&+ \frac{\frac{3}{4} - (1-\lambda)\, q^{2} + \sqrt{\left(\frac{3}{4} - (1-\lambda)\, q^{2}\right)^{2} - (\lambda)\, q^{2}}}{2} \\
&\qquad \log\left(\frac{\frac{3}{4} - (1-\lambda)\, q^{2} + \sqrt{\left(\frac{3}{4} - (1-\lambda)\, q^{2}\right)^{2} - (\lambda)\, q^{2}}}{2}\right) \\
&+ \left(\frac{1}{4} + (1-\lambda)\, q^{2}\right) \log\left(\frac{1}{4} + (1-\lambda)\, q^{2}\right).
\end{aligned}$$

For $\lambda = \frac{1}{2}$ we have shown in Figure C.1 how the rate of distillable key changes with the choice of parameter $q \in \left[0, \frac{1}{\sqrt{2}}\right]$.

Bibliography

[1] Matthias Christandl and Roberto Ferrara. Private states, quantum data hiding, and the swapping of perfect secrecy. *Physical Review Letters*, 119(22), 2017.

[2] Dmytro Gavinsky, Julia Kempe, Iordanis Kerenidis, Ran Raz, and Ronald de Wolf. Exponential separations for one-way quantum communication complexity, with applications to cryptography. 2006.

[3] Peter W. Shor. Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer. *SIAM Journal on Computing*, 26(5):1484–1509, oct 1997.

[4] Charles H. Bennett and Gilles Brassard. An update on quantum cryptography. In *Annual International Cryptology Conference*, 1985.

[5] Artur K. Ekert. Quantum cryptography based on bell's theorem. *Phys. Rev. Lett.*, 67:661–663, Aug 1991.

[6] Karol Horodecki, Michał Horodecki, Paweł Horodecki, and Jonathan Oppenheim. Secure key from bound entanglement. *Physical Review Letters*, 94(16), 2005.

[7] Roberto Ferrara. *An Information-Theoretic Framework for Quantum Repeaters*. PhD thesis, 2018.

[8] Michał Horodecki, Paweł Horodecki, and Ryszard Horodecki. Separability of n-particle mixed states: necessary and sufficient conditions in terms of linear maps. *Physics Letters A*, 283(1-2):1–7, may 2001.

[9] E. M. Rains. Bound on distillable entanglement. *Physical Review A*, 60(1):179–184, jul 1999.

[10] Scott Aaronson, Daniel Grier, and Luke Schaeffer. The classification of reversible bit operations, 2015.

[11] William Matthews, Stephanie Wehner, and Andreas Winter. Distinguishability of quantum states under restricted families of measurements with an application to quantum data hiding. *Communications in Mathematical Physics*, 291(3):813–843, Aug 2009.

[12] Christopher A. Fuchs and Jeroen van de Graaf. Cryptographic distinguishability measures for quantum mechanical states, 1997.

[13] Marco Tomamichel, Roger Colbeck, and Renato Renner. Duality between smooth min- and max-entropies. *IEEE Transactions on Information Theory*, 56(9):4674–4681, sep 2010.

[14] Ashley Montanaro. On the distinguishability of random quantum states. *Communications in Mathematical Physics*, 273(3):619–636, mar 2007.

[15] Carl W. Helstrom. Quantum detection and estimation theory. *Journal of Statistical Physics*, 1(2):231–252, June 1969.

[16] Marco Tomamichel, Roger Colbeck, and Renato Renner. A fully quantum asymptotic equipartition property. *IEEE Transactions on Information Theory*, 55(12):5840–5847, Dec 2009.

[17] Marco Tomamichel, Christian Schaffner, Adam Smith, and Renato Renner. Leftover hashing against quantum side information. *IEEE Transactions on Information Theory*, 57(8):5524–5535, Aug 2011.

[18] Pattrawut Chansangiam. A survey on operator monotonicity, operator convexity, and operator means. *International Journal of Analysis*, 2015:1–8, 11 2015.

[19] Koenraad M R Audenaert. A sharp continuity estimate for the von neumann entropy. *Journal of Physics A: Mathematical and Theoretical*, 40(28):8127–8136, jun 2007.

[20] R. Alicki and M. Fannes. Continuity of quantum conditional information. 2003.

[21] Andreas Winter. Tight uniform continuity bounds for quantum entropies: Conditional entropy, relative entropy distance and energy constraints. *Communications in Mathematical Physics*, 347(1):291–313, mar 2016.

[22] Robert Konig, Renato Renner, and Christian Schaffner. The operational meaning of min- and max-entropy. *IEEE Transactions on Information Theory*, 55(9):4337–4347, Sep 2009.

[23] Robert König, Renato Renner, Andor Bariska, and Ueli Maurer. Small accessible quantum information does not imply security. *Physical Review Letters*, 98(14), apr 2007.

[24] M. Fekete. Über die verteilung der wurzeln bei gewissen algebraischen gleichungen mit ganzzahligen koeffizienten. *Mathematische Zeitschrift*, 17:228–249, 1918.

[25] Robert T. Konig and Barbara M. Terhal. The bounded-storage model in the presence of a quantum adversary. *IEEE Transactions on Information Theory*, 54(2):749–762, feb 2008.

[26] A.S Holevo. Statistical decision theory for quantum systems. *Journal of Multivariate Analysis*, 3(4):337–394, 1973.

[27] Reinhard F. Werner. Quantum states with einstein-podolsky-rosen correlations admitting a hidden-variable model. *Phys. Rev. A*, 40:4277–4281, Oct 1989.

[28] William Matthews and Andreas Winter. On the chernoff distance for asymptotic LOCC discrimination of bipartite quantum states. *Communications in Mathematical Physics*, 285(1):161–174, jul 2008.

[29] D.P. DiVincenzo, D.W. Leung, and B.M. Terhal. Quantum data hiding. *IEEE Transactions on Information Theory*, 48(3):580–598, mar 2002.

[30] K. Audenaert, J. Eisert, M. B. Plenio, and R. F. Werner. Entanglement properties of the harmonic chain. *Physical Review A*, 66(4), oct 2002.

[31] Stefan Bäuml, Matthias Christandl, Karol Horodecki, and Andreas Winter. Limitations on quantum key repeaters. *Nature Communications*, 6(1), 2015.

[32] A.S. Holevo. The capacity of the quantum channel with general signal states. *IEEE Transactions on Information Theory*, 44(1):269–273, 1998.

[33] Karol Horodecki, Michal Horodecki, Pawel Horodecki, and Jonathan Oppenheim. General paradigm for distilling classical key from quantum states. *IEEE Transactions on Information Theory*, 55(4):1898–1929, apr 2009.

[34] Igor Devetak and Andreas Winter. Distillation of secret key and entanglement from quantum states. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 461(2053):207–235, Jan 2005.

[35] G. Vidal and R. F. Werner. Computable measure of entanglement. *Physical Review A*, 65(3), feb 2002.

[36] Graeme Smith and Jon Yard. Quantum communication with zero-capacity channels. *Science*, 321(5897):1812–1815, sep 2008.

[37] Felix Leditzky, Debbie Leung, Vikesh Siddhu, Graeme Smith, and John A. Smolin. The platypus of the quantum channel zoo. *To be published*, 41(4):915–940, Jan 2012.

[38] A. Uhlmann. The "transition probability" in the state space of a *-algebra. *Reports on Mathematical Physics*, 9(2):273–279, 1976.

[39] Mario Berta, Omar Fawzi, and Stephanie Wehner. Quantum to classical randomness extractors. *IEEE Transactions on Information Theory*, 60(2):1168–1192, feb 2014.

[40] Vivek V. Shende, Aditya K. Prasad, Igor L. Markov, and John P. Hayes. Reversible logic circuit synthesis. 2002.

[41] Matthew Amy, Parsiad Azimzadeh, and Michele Mosca. On the controlled-NOT complexity of controlled-NOT–phase circuits. *Quantum Science and Technology*, 4(1):015002, sep 2018.

[42] Norbert Schuch and Jens Siewert. Programmable networks for quantum algorithms. *Physical Review Letters*, 91(2), jul 2003.

[43] Marco Ugo Gambetta. Few-qubit behaviour and noise-robustness of quantum-classical separations in communication complexity, 2019.

[44] W.W. Rouse Ball and H.S.M. Coxeter. *Mathmatical Recreations & Essays: 12th Edition.* University of Toronto Press, 1974.

[45] Frédéric Dupuis, Omar Fawzi, and Renato Renner. Entropy accumulation. *Communications in Mathematical Physics*, 379(3):867–913, sep 2020.

[46] Tony Metger, Omar Fawzi, David Sutter, and Renato Renner. Generalised entropy accumulation, 2022.

[47] R. Fano. *Transmission of Information: A Statistical Theory of Communications.* The MIT Press, Cambridge, MA, 1961.