Uncertainty relations for multiple measurements with applications

Omar Fawzi

School of Computer Science
McGill University, Montréal
August, 2012

A thesis submitted to McGill University in partial fulfillment of the requirements of the degree of PhD.

Copyright ©Omar Fawzi 2012.
Contents

Abstract ................................................................. iii
Résumé ................................................................. iv
Contents of the thesis .............................................. v
Acknowledgements .................................................. vi
Notation ................................................................. vii

1 Introduction ......................................................... 1
  1.1 Quantum information science .............................. 1
  1.2 Conjugate coding ............................................. 2
  1.3 Uncertainty relations for quantum measurements .... 3
    1.3.1 Context ............................................... 3
    1.3.2 Summary of the contributions ......................... 5
  1.4 Uncertainty relations in the presence of quantum side information .... 5
    1.4.1 Context ............................................... 5
    1.4.2 Summary of contributions ............................ 6

2 Preliminaries ..................................................... 7
  2.1 Classical information theory ............................... 7
  2.2 Representation of physical systems ...................... 8
    2.2.1 Quantum states ................................... 8
    2.2.2 Evolution of quantum systems ................... 9
    2.2.3 Distance measures ................................ 10
    2.2.4 Information measures ............................. 12
  2.3 Quantum computation ..................................... 15

3 Uncertainty relations: Definitions and constructions .... 17
  3.1 Background ................................................ 17
  3.2 Metric uncertainty relations ............................. 20
Abstract

Uncertainty relations express the fundamental incompatibility of certain observables in quantum mechanics. Far from just being puzzling constraints on our ability to know the state of a quantum system, uncertainty relations are at the heart of why some classically impossible cryptographic primitives become possible when quantum communication is allowed. This thesis is concerned with strong notions of uncertainty relations and their applications in quantum information theory.

One operational manifestation of such uncertainty relations is a purely quantum effect referred to as information locking. A locking scheme can be viewed as a cryptographic protocol in which a uniformly random $n$-bit message is encoded in a quantum system using a classical key of size much smaller than $n$. Without the key, no measurement of this quantum state can extract more than a negligible amount of information about the message, in which case the message is said to be “locked”. Furthermore, knowing the key, it is possible to recover, that is “unlock”, the message. We give new efficient constructions of bases satisfying strong uncertainty relations leading to the first explicit construction of an information locking scheme. We also give several other applications of our uncertainty relations both to cryptographic and communication tasks.

In addition, we define objects called QC-extractors, that can be seen as strong uncertainty relations that hold against quantum adversaries. We provide several constructions of QC-extractors, and use them to prove the security of cryptographic protocols for two-party computations based on the sole assumption that the parties’ storage device is limited in transmitting quantum information. In doing so, we resolve a central question in the so-called noisy-storage model by relating security to the quantum capacity of storage devices.
Résumé

Les relations d’incertitude expriment l’incompatibilité de certaines observables en mécanique quantique. Les relations d’incertitude sont utiles pour comprendre pourquoi certaines primitives cryptographiques impossibles dans le monde classique deviennent possibles avec de la communication quantique. Cette thèse étudie des notions fortes de relations d’incertitude et leurs applications à la théorie de l’information quantique.

Une manifestation opérationnelle de telles relations d’incertitude est un effet purement quantique appelé verrouillage d’information. Un système de verrouillage peut être considéré comme un protocole cryptographique dans lequel un message aléatoire composé de \( n \) bits est encodé dans un système quantique en utilisant une clé classique de taille beaucoup plus petite que \( n \). Sans la clé, aucune mesure sur cet état quantique ne peut extraire plus qu’une quantité négligeable d’information sur le message, auquel cas le message est “verrouillé”. Par ailleurs, connaissant la clé, il est possible de récupérer ou “déverrouiller” le message.

Nous proposons de nouvelles constructions efficaces de bases vérifiant de fortes relations d’incertitude conduisant à la première construction explicite d’un système de verrouillage. Nous exposons également plusieurs autres applications de nos relations d’incertitude à des tâches cryptographiques et des tâches de communication.

Nous définissons également des objets appelés QC-extracteurs, qui peuvent être considérés comme de fortes relations d’incertitude qui tiennent contre des adversaires quantiques. Nous fournissons plusieurs constructions de QC-extracteurs, que nous utilisons pour prouver la sécurité de protocoles cryptographiques pour le calcul sécurisé à deux joueurs en supposant uniquement que la mémoire des joueurs soit limitée en ce qui concerne la transmission d’information quantique. Ce faisant, nous résolvons une question centrale dans le modèle de mémoire brouillée en mettant en relation la sécurité et la capacité quantique de la mémoire.
Contents of the thesis

This thesis is mainly based on two papers. The first one is joint work with Patrick Hayden and Pranab Sen [Fawzi et al. 2011] and is presented in Chapters 3 and 4. The second paper is presented in Chapter 5 and is joint work with Mario Berta and Stephanie Wehner [Berta et al. 2012].
Acknowledgements

I was very fortunate to be supervised by Luc Devroye and Patrick Hayden during the last four years. I wish to thank them for their generosity with their time and for their guidance. I learned almost all of what I know about probability from Luc and almost all what I know about classical and quantum information theory from Patrick. I also thank them for their financial support. I also consider Pranab Sen as my supervisor. I thank him for all his extremely clear explanations. I should also thank Claude Crépeau, Aram Harrow and Louis Salvail for accepting to be on my committee.

I had the chance of collaborating with many other great researchers. I would like to thank my other co-authors for all what they taught me: Anil Ada, Mario Berta, Nicolas Broutin, Arkadev Chattopadhyay, Nicolas Fraiman, Phuong Nguyen, Ivan Savov, Stephanie Wehner, Mark Wilde. Special thanks to Anil and Ivan for countless hours of entertaining discussions and Mark for carefully reading this thesis. Part of the work presented in this thesis was done while I was visiting the Center for Quantum Technologies in Singapore. I wish to thank Miklós Sántha for inviting me. I also benefited from various useful discussions with many colleagues including Louigi Addario-Berry, Abdulrahman Al-lahham, Kamil Bradler, Frédéric Dupuis, Nicolas Dutil, Laszlo Egri, Hamza Fawzi, Hussein Fawzi, Jan Florjanczyk, Hamed Hatami, Kassem Kalach, Ross Kang, Marc Kaplan, Jamie King, Sean Kennedy, Pascal Koiran, Christian Konrad, Nima Lashkari, Debbie Leung, Zhentao Li, Frédéric Magniez, Joe Renes, Renato Renner, Miklós Sántha, Thomas Vidick, Andreas Winter, Jürg Wullschleger.

I am also grateful to all the staff of the School of Computer Science for providing an excellent environment. I certainly have to thank many friends who made life in Montréal so enjoyable. At the risk of forgetting many, I’ll try to mention some of them: Ahmad, Alaa, Kosai, Mazen, Mohammed’s, Mustafa, Nour, Omar, Qusai, Rayhan, Saeed, Tamer, Tarek. And it is a pleasure to thank my parents and brothers for everything.
## Notation

**Common**

| Symbol | Description |
|--------|-------------|
| log    | Binary logarithm. |
| ln     | Natural logarithm. |
| \( \mathbb{R} \) | Real numbers. |
| \( \mathbb{C} \) | Complex numbers. |
| \( M^\dagger \) | Conjugate transpose of the matrix \( M \). |
| \([n]\) | Set \( \{1, \ldots, n\} \). |
| \( d_H \) | Hamming distance \( d_H(x, y) = \{i : x_i \neq y_i\} \). |
| \( w \) | Hamming weight \( w(x) = \{i : x_i \neq 0\} \). |
| \( p_X \) | The distribution of a random variable \( X \). |
| \( \Pr \{E\} \) | Probability of the event \( E \). |
| \( \mathbf{E} \{X\} \) | Expectation of a random variable \( X \). |
| \( \mathbf{E}_y \{f(x, y)\} \) | Expectation over \( y \) and fixed \( x \). |
| \( f \circ g \) | Composition of the functions \( f \) and \( g \). |

**Spaces**

| Symbol | Description |
|--------|-------------|
| \( A, B, C, \ldots \) | Hilbert spaces associated with the systems \( A, B, C, \ldots \) |
| \( A \simeq A' \) | \( A' \) is a copy of \( A \). |
| \( d_A \) | Dimension of the space \( A \). |
| \( AB \) | Tensor product \( A \otimes B \) or composite system \( AB \). |
| \( \mathcal{L}(A, B) \) | Space of linear operators from \( A \) to \( B \). |
| \( \mathcal{L}(A) \) | \( \mathcal{L}(A, A) \). |

**Vectors**

| Symbol | Description |
|--------|-------------|
| \( |\psi\rangle^A, |\phi\rangle^A, \ldots \) | Vectors belonging to \( A \). |
| \( \langle \psi |^A \) | Dual vectors in \( \mathcal{L}(A, \mathbb{C}) \). |
| \( \langle \psi | \phi \rangle \) | Inner product of the vectors \( |\psi\rangle \) and \( |\phi\rangle \). |
### Operators

| Operator | Description |
|----------|-------------|
| $S(A)$ | Set of density operators on $A$. |
| $S_\leq(A)$ | Set of sub-normalized density operators on $A$. |
| $\rho^A = \rho_A, \psi^A, \ldots$ | Density operators on $A$. |
| $i|_A = i|_A^A$ | Identity map on $A$ or $\mathcal{L}(A)$. |
| $\|X\|_1$ | Trace norm of the operator $X$. |
| $\|X\|_2$ | Hilbert-Schmidt norm of the operator $X$. |

### Distance measures for operators

| Measure | Description |
|---------|-------------|
| $\Delta(\rho, \sigma)$ | Trace distance between $\rho$ and $\sigma$. |
| $F(\rho, \sigma)$ | Fidelity between $\rho$ and $\sigma$. |
| $\tilde{F}(\rho, \sigma)$ | Generalized fidelity between $\rho$ and $\sigma$. |
| $P(\rho, \sigma)$ | Purified distance between $\rho$ and $\sigma$. |

### Measures of information

| Measure | Description |
|---------|-------------|
| $H(A)_\rho$ | von Neumann entropy of the density operator $\rho^A$. |
| $H(A|B)_\rho$ | Conditional von Neumann entropy of $\rho^{AB}$. |
| $I(A; B)_\rho$ | Mutual information of the density operator $\rho^{AB}$. |
| $H_{\min}(A|B)_{\rho|\sigma}$ | Min-entropy of $\rho^{AB}$ relative to $\sigma^B$. |
| $H_{\min}(A|B)_\rho$ | Conditional min-entropy of $\rho^{AB}$ given $B$. |
| $H_{\max}(A|B)_\rho$ | Conditional max-entropy of $\rho^{AB}$ given $B$. |
| $H^\epsilon_{\min}(A|B)_\rho$ | Smooth min-entropy of $\rho^{AB}$ given $B$. |
| $H^\epsilon_{\max}(A|B)_\rho$ | Smooth max-entropy of $\rho^{AB}$ given $B$. |
| $H_2(A|B)_{\rho|\sigma}$ | Collision entropy of $\rho^{AB}$ relative to $\sigma^B$. |
| $h_2(\epsilon)$ | Binary entropy $h_2(\epsilon) = -\epsilon \log \epsilon - (1 - \epsilon) \log(1 - \epsilon)$. |
Chapter 1

Introduction

1.1 Quantum information science

Even though Turing machines are abstract mathematical constructions, they are widely believed to capture a universal notion of computation in our physical world. This is reflected by the Church-Turing thesis, which states that any computation performed on a physical device can also be performed by a Turing machine. The main reason the Church-Turing thesis is believed is that all known models for (reasonable) physical computation mechanisms were shown to be simulatable by a Turing machine. In fact, the strong Church-Turing thesis states that any computation performed efficiently by some physical device can be computed efficiently by a Turing machine.

Consider now the problem of information transmission through a physical channel. How to model such a channel? A natural answer is to associate for each possible input a probability distribution on the possible outputs of the channel. The randomness is used to model our ignorance or lack of control of some phenomena happening in the transmission. There is a feeling that a better understanding of the physical process can always be incorporated in the model by adjusting the probabilities assigned to each outcome. As for Turing machines, there is a belief that the most general way to model a physical information channel is using probability distributions.

When taking into account quantum theory, these assumptions should be re-examined. According to quantum physics, the state of a physical system, e.g., a potential computing device, need not be represented by some string of characters written on a tape, but can potentially be a superposition of many strings. Like for waves, different parts of the system could interfere with each other. Could such a model define a different notion of computation? There is by now significant evidence that this might be the case. 

Shor
 showed that one could make use of wave-like properties in a quantum system to factor integers efficiently, a problem that is believed to be hard for classical computers. In addition, many other computational tasks seem to be much more natural and efficiently implementable for a quantum computer, in particular concerning the simulation of quantum mechanical systems [Feynman, 1982]. For information transmission, imagine a channel that carries information using photon polarization, a property which is known to be best described by a quantum state. In this case, modelling the channel as a distribution over the outputs for each possible input is incomplete. In fact, it turns out that one can use quantum mechanical properties not only to increase the rate at which information is transmitted but also to perform tasks that are simply impossible using only “classical” communication.

1.2 Conjugate coding

An example of a task that becomes possible when quantum properties are used is key distribution. Suppose Alice and Bob are far apart and they want to exchange a secret over email. To achieve unconditional security\(^1\), it is well-known that they have to share a large private key about which the adversary does not have any information. How can they obtain such a key by communicating over a public channel? This task is impossible to achieve with unconditional security using only classical communication.

In groundbreaking work, Bennett and Brassard [1984] based on an idea of Wiesner [1983] devised a simple protocol for key distribution using quantum communication\(^2\). One of the key ideas of the protocol is to use “conjugate coding” [Wiesner, 1983]. Even though it cannot store (reliably) more than one bit of information, there are several ways of encoding one bit in the polarization of a photon. We can encode in the “rectilinear” basis, e.g., \(0 \mapsto H\) (horizontal) and \(1 \mapsto V\) (vertical), or in the “diagonal” basis, e.g., \(0 \mapsto M\) (main diagonal) and \(1 \mapsto A\) (anti-diagonal). This is a valid encoding because in both cases, the states corresponding to \(0\) and \(1\) are perfectly distinguishable. However, an observer that does not know which one of the two encodings was used cannot recover the encoded bit perfectly. In fact, if he performs a measurement in the rectilinear basis and the actual state was \(M\) (which belongs to the diagonal basis and encodes \(0\)), then the result will be \(H\) (corresponding to \(0\)) with probability \(1/2\) and \(V\) (corresponding to \(1\)) with probability \(1/2\).

We stress that this type of encoding in the polarization of a single photon does not have a

\(^1\) Unconditional security means that security doesn’t rest on unproven computational assumptions.

\(^2\) Note that this protocol can be and is implemented with today’s technology. In fact, encryptors based on quantum key distribution can actually be bought from a handful of companies.
classical analogue. Assume we have two perfectly distinguishable classical states $A$ and $B$. One can define two possible encodings for bits: $0 \mapsto A$ and $1 \mapsto B$, or $0 \mapsto B$ and $1 \mapsto A$. An adversary who ignores which encoding was used cannot obtain any information about the encoded bit by seeing $A$ or $B$. But given that we see $A$, we know that the encoded bit is $0$ if the first encoding was chosen and it is $1$ if the second encoding was chosen. For the quantum encoding described above, for all possible states $H, V, M$ or $A$, it is not possible to have a definite encoded value for both the rectilinear and diagonal bases. This is a form of the uncertainty principle: either the “rectilinear value” or the “diagonal value” of a state has to be undetermined. This idea of encoding in conjugate bases is at the heart of the whole field of quantum cryptography that takes advantage of the uncertainty principle and related ideas to guarantee privacy properties; see [Gisin et al., 2002, Scarani et al., 2009] for surveys. The results in this thesis can be seen as stronger versions of conjugate coding that use multiple (more than two) encodings.

The following more technical sections describe the context and the main results in this thesis.

1.3 Uncertainty relations for quantum measurements

1.3.1 Context

The uncertainty principle was first formulated by Heisenberg [1927] and it states that the position and momentum of a particle cannot both have definite values. It was then generalized by Robertson [1929] to arbitrary observables that do not commute. Here, we consider modern formulations of the uncertainty principle for which the measure of uncertainty is an entropic quantity. Entropic uncertainty relations were introduced in Bialynicki-Birula and Mycielski [1975], Deutsch [1983], Hirschman [1957] and have found many applications in quantum information theory. For example, such relations are the main ingredients in the proofs of security of protocols for two-party computations in the bounded and noisy quantum storage models [Damgård et al., 2005, 2007, König et al., 2012]. A simple example of an entropic uncertainty relation was given by Maassen and Uffink [1988]. Let $B_+$ denote a “rectilinear” or computational basis of $\mathbb{C}^2$ and $B_x$ be a “diagonal” or Hadamard basis and let $B_{+n}$ and $B_{x^n}$ be the corresponding bases obtained on the tensor product space $(\mathbb{C}^2)^\otimes n$. All vectors in the rectilinear basis $B_{+n}$ have an inner product with all vectors in the diagonal basis $B_{x^n}$ upper bounded by $2^{-n/2}$ in absolute value. The uncertainty relation of Maassen and Uffink [1988] states that for any quantum state on
1.3. Uncertainty relations for quantum measurements

$n$ qubits described by a unit vector $|\psi\rangle \in (\mathbb{C}^2)^\otimes n$, the average measurement entropy satisfies

$$\frac{1}{2} \left( H(p_{B_{+}^n,|\psi\rangle}) + H(p_{B_{x}^n,|\psi\rangle}) \right) \geq \frac{n}{2},$$

(1.1)

where $p_{B,|\psi\rangle}$ denotes the outcome probability distribution when $|\psi\rangle$ is measured in basis $B$ and $H$ denotes the Shannon entropy. Equation (1.1) expresses the fact that the outcome of at least one of two measurements cannot be well predicted, even after knowing which measurement was performed.

A surprising application of entropic uncertainty relations is the effect known as information locking [DiVincenzo et al., 2004] (see also Leung [2009]). Suppose Alice holds a uniformly distributed random $n$-bit string $X$. She chooses a random basis $K \in \{+^n, x^n\}$ and encodes $X$ in the basis $B_K$. This random quantum state $\mathcal{E}(X,K)$ is then given to Bob. How much information about $X$ can Bob, who does not know $K$, extract from this quantum system via a measurement? To better appreciate the quantum case, observe that if $X$ were encoded in a classical state $\mathcal{E}_c(X,K)$, then $\mathcal{E}_c(X,K)$ would “hide” at most one bit about $X$; more precisely, the mutual information between $X$ and $\mathcal{E}_c(X,K)$ is at least $n - 1$. For the quantum encoding $\mathcal{E}$, one can show that for any measurement that Bob applies on $\mathcal{E}(X,K)$ whose outcome is denoted $I$, the mutual information between $X$ and $I$ is at most $n/2$ [DiVincenzo et al., 2004]. The $n/2$ missing bits of information about $X$ are said to be locked in the quantum state $\mathcal{E}(X,K)$. If Bob had access to $K$, then $X$ can be easily obtained from $\mathcal{E}(X,K)$: The one-bit key $K$ can be used to unlock $n/2$ bits about $X$.

A natural question is whether it is possible to lock more than $n/2$ bits in this way. In order to achieve this, the key $K$ has to be chosen from a larger set. In terms of uncertainty relations, this means that we need to consider $t > 2$ bases to achieve an average measurement entropy larger than $n/2$ (equation (1.1)). In this case, the natural candidate is a set of $t$ mutually unbiased bases, the defining property of which is a small inner product between any pair of vectors in different bases. Surprisingly, it was shown by Ballester and Wehner [2007] and Ambainis [2010] that there are up to $t = 2^{n/2}$ mutually unbiased bases $\{B_1, B_2, \ldots, B_t\}$ that only satisfy an average measurement entropy of $n/2$, which is only as good as what can be achieved with two measurements (1.1). In other words, looking at the pairwise inner product between vectors in different bases is not enough to obtain uncertainty relations stronger than (1.1).

To achieve an average measurement entropy of $(1 - \epsilon)n$ for small $\epsilon$ while keeping the number of bases subexponential in $n$, the only known constructions are probabilistic and computationally inefficient [Hayden et al., 2004].
1.3.2 Summary of the contributions

Chapter 3 We introduce the notion of a metric uncertainty relation and connect it to low-distortion embeddings of $\ell_2$ into $\ell_1$. A metric uncertainty relation also implies an entropic uncertainty relation. We prove that random bases satisfy uncertainty relations with a stronger definition and better parameters than previously known. Our proof is also considerably simpler than earlier proofs. We give efficient constructions of bases satisfying metric uncertainty relations. The bases are computable by quantum circuits of almost linear size. These constructions are obtained by adapting an explicit norm embedding due to Indyk (2007) and an extractor construction of Guruswami et al. (2009).

Chapter 4 We prove that any metric uncertainty relation leads to a locking scheme. Applying the results of Chapter 3, we show the existence of locking schemes with key size independent of the message length. Moreover, using the efficient constructions, we give the first explicit strong information locking scheme. Moreover, we present a locking scheme that can in principle be implemented with current technology. We use our locking schemes to construct hiding fingerprints as defined by Gavinsky and Ito (2010).

We also apply our metric uncertainty relations to exhibit communication protocols that perform equality testing of $n$-qubit states. We prove that this task can be performed by a single message protocol using $O(\log(1/\epsilon))$ qubits and $n$ bits of communication, where $\epsilon$ is an error parameter. We also give a single message protocol that uses $O(\log^2 n)$ qubits, where the computation of the sender is efficient.

1.4 Uncertainty relations in the presence of quantum side information

1.4.1 Context

Suppose that we are now looking for a stronger notion of uncertainty. We want the outcome to be unpredictable even if the adversary, who is trying to predict the outcome of the measurement, holds a system that is entangled with the system being measured. Let Alice hold a system $A$ and Eve hold $E$, and the two systems are maximally entangled. How well can Eve predict the outcome of measurements in bases $\{B_1, \ldots, B_t\}$? It turns out that because Alice and Eve are maximally entangled, Eve can perfectly predict the outcome that Alice obtains. In this case, there is no uncertainty at all from the point of view of Eve. The
interesting question is then: Can we obtain some uncertainty if Eve holds some quantum side information about the system $A$ but is not maximally entangled with it? The amount of uncertainty in the measurement outcomes should then be a function of some quantum correlation measure between Alice and Eve. We should note here that unlike classical side information which can usually be handled easily, quantum side information can behave in unexpected ways; see for example the work on randomness extractors against quantum adversaries [Gavinsky et al., 2007, König et al., 2005, Renner and König, 2005]. In beautiful recent work, Renes and Boileau [2009] and Berta et al. [2010] showed that in fact one can extend the uncertainty relation in equation (1.1) to allow for quantum side information. Related uncertainty relations that hold in the presence of quantum memory have proven to be a very useful tool in security proofs for quantum key distribution [Furrer et al., 2011, Tomamichel and Renner, 2011, Tomamichel et al., 2012].

But as in the previous section, just two measurements are in many cases not sufficient to obtain the desired amount of uncertainty. Before this work, uncertainty relations that hold when the adversary has a quantum memory were known only for two measurements.

1.4.2 Summary of contributions

Chapter 5 We introduce QC-extractors by analogy to classical randomness extractors, which are objects that found many applications in theoretical computer science, and relate them to uncertainty relations with quantum side information. Using techniques similar to the ones used for proving decoupling results, we give several constructions of QC-extractors based on unitary two-designs, complete sets of mutually unbiased bases and single-qudit unitaries. These naturally lead to uncertainty relations in terms of the min-entropy and in terms of the von Neumann entropy. This gives the first uncertainty relations in the presence of quantum side information for more than two measurements. Moreover, we use the uncertainty relation for single-qubit measurements to finally link the security of two-party secure function evaluation to the ability of the parties’ storage device to store quantum information [Wehner et al., 2008]. Previously, the security could only be shown when the classical capacity [König et al., 2012] or entanglement cost [Berta et al., 2011a] of the storage device was limited.
Chapter 2

Preliminaries

The objective of this chapter is to introduce some notations and results that will be used throughout this thesis. We start with a very brief section about classical information theory before moving to the description of quantum systems.

2.1 Classical information theory

Random variables are usually denoted by capital letters $X, K, \ldots$, while $p_X$ denotes the distribution of $X$, i.e., $\Pr \{ X = x \} = p_X(x)$. The notation $X \sim p$ means that $X$ has distribution $p$. $\text{unif}(S)$ is the uniform distribution on the set $S$. To measure the distance between probability distributions on a finite set $\mathcal{X}$, we use the total variation distance or trace distance $\Delta(p, q) = \frac{1}{2} \sum_{x \in \mathcal{X}} |p(x) - q(x)|$. We also have $\Delta(p, q) = \max_{A \subseteq \mathcal{X}} (\sum_{x \in A} p(x) - \sum_{x \in A} q(x))$.

We will also write $\Delta(X, Y)$ for $\Delta(p_X, p_Y)$. When $\Delta(X, Y) \leq \epsilon$, we say that $X$ is $\epsilon$-close to $Y$. A useful characterization of the trace distance is $\Delta(p, q) = \max_{X \sim p, Y \sim q} \Pr \{ X = Y \}$ (this equality is sometimes attributed to Doeblin [1938]). Another useful measure of closeness between distributions is the fidelity $F(p, q) = \sum_{x \in \mathcal{X}} \sqrt{p(x)q(x)}$ also known as the Bhattacharyya distance and related to the Hellinger distance. We have the following relation between the fidelity and the trace distance:

$$1 - F(p, q) \leq \Delta(p, q) \leq \sqrt{1 - F(p, q)^2}. \quad (2.1)$$

The Shannon entropy of a distribution $p$ on $\mathcal{X}$ is defined as $H(p) = -\sum_{x \in \mathcal{X}} p(x) \log p(x)$ where the log is taken here and throughout the thesis to be base two. We will also write $H(X)$ for $H(p_X)$. The conditional entropy is defined by $H(X|Y) = H(XY) - H(Y)$. It also has the property that $H(X|Y) = E_y \{ H(X|Y = y) \}$. The mutual information between
two random variables $X$ and $Y$ is defined as $I(X; Y) = H(X) + H(Y) - H(X, Y)$. The min-entropy of a distribution $p$ is defined as $H_{\min}(p) = -\log \max_x p(x)$. We say that a random variable $X$ is a $k$-source if $H_{\min}(X) \geq k$.

## 2.2 Representation of physical systems

We briefly describe the notation and the basic facts about quantum theory that will be used in this thesis. We refer the reader to [Nielsen and Chuang (2000), Wilde (2011)] for more details.

### 2.2.1 Quantum states

The state of a (pure) quantum system is represented by a unit vector in a Hilbert space. For the purpose of this thesis, a Hilbert space is a finite-dimensional complex inner product space. Quantum systems are denoted $A, B, C \ldots$ and are identified with their corresponding Hilbert spaces. The dimension of $A$ is denoted $d_A$. It is important to note that all unit vectors represent valid physical states and for any two different vectors, one can perform an experiment for which the two states have a different observable behaviour. Vectors in $A$ are denoted by “kets” $|\psi\rangle_A \in A$ and dual vectors (i.e., linear functions from $A$ to $\mathbb{C}$) are denoted by “bras” $\langle \phi |$, so that $\langle \phi | (|\psi\rangle) = \langle \phi | \psi \rangle$ is simply the inner product between the vectors $|\phi\rangle$ and $|\psi\rangle$. Performing the product in the other direction $|\psi\rangle \langle \phi |$, we obtain a linear transformation mapping $A$ to itself. In particular, $|\psi\rangle\langle \psi |$ is the orthogonal projector onto the span of $|\psi\rangle$. If we fix a basis of the Hilbert space, then we can represent $|\psi\rangle$ as a column vector $\vec{v}$ and the dual vector $\langle \psi |$ can be represented by $\vec{v}^\dagger$, where $M^\dagger$ represents the conjugate transpose of the matrix $M$. In this thesis, every Hilbert space $A$ comes with a preferred orthonormal basis $\{ |a\rangle \}_{a \in [d_A]}$ that we call the computational basis. The elements of this basis are labeled by integers in $[d_A] \equiv \{1, \ldots, d_A\}$. Often, the Hilbert spaces we consider are composed of $n$ qubits, i.e., have the form $(\mathbb{C}^2)^\otimes n$. In this case, the computational basis will also be labeled by strings in $\{0, 1\}^n$.

In order to model our ignorance of the description of a quantum system, we can consider distributions $\{p_1, \ldots, p_r\}$ over quantum states $\{|\psi_1\rangle, \ldots, |\psi_r\rangle\}$. It is well known that such a distribution over states is best described by a density operator $\rho = \sum_{i=1}^r p_i |\psi_i\rangle\langle \psi_i |$ acting on $A$. We denote by $\mathcal{L}(A, B)$ the set of linear transformations from $A$ to $B$ and we write

\[ \text{Technically, quantum states are actually rays rather than unit vectors in the Hilbert space, so two vectors that only differ by a global phase represent the same state.} \]
\( \mathcal{L}(A) \) for \( \mathcal{L}(A, A) \). Observe that a density operator is a Hermitian positive semidefinite operator with unit trace. Conversely, any unit trace Hermitian operator \( \rho \) with non-negative eigenvalues is a valid density operator. If we write \( \rho^A = \sum_i p_i |\psi_i\rangle \langle \psi_i| \) where \( \{|\psi_i\rangle\}_i \) form an orthonormal eigenbasis for \( \rho \), we can interpret the state of \( A \) as being \( |\psi_i\rangle \) with probability \( p_i \). In particular, the density operator associated with a pure state \( |\psi\rangle \) is \( |\psi\rangle \langle \psi| \) and it will be abbreviated by omitting the ket and bra: \( \psi \defeq |\psi\rangle \langle \psi| \). We use \( S(A) \) to denote the set of density operators acting on \( A \). The Hilbert space on which a density operator \( \rho \in S(A) \) acts is sometimes denoted by a superscript or subscript, as in \( \rho^A \) or \( \rho_A \). This notation is also used for pure states \( |\psi\rangle^A \in A \).

In order to describe the joint state of a system \( AB \), the associated state space is the tensor product Hilbert space \( A \otimes B \), which is sometimes simply denoted \( AB \). If \( \rho^{AB} \) describes the joint state on \( AB \), the state on the system \( A \) is described by the partial trace \( \rho^A \defeq \text{tr}_B \rho^{AB} \). The partial trace \( \text{tr}_B : \mathcal{L}(A \otimes B) \to \mathcal{L}(A) \) is defined as \( \text{tr}_B[\rho^{AB}] = \sum_b (\text{id}_A \otimes \langle b|) \rho^{AB} (\text{id}_A \otimes |b\rangle) \), where \( \{|b\rangle\} \) is an orthonormal basis of \( B \).

A classical system can easily be described using this formalism. A distribution \( \{p_i\} \) over \([d]\) is represented by \( \rho = \sum_{i \in [d]} p_i |i\rangle \langle i| \). A state on \( XB \) is said to be classical on \( X \) if there exists a basis \( \{|x\rangle\} \) of \( X \) and a set of (non-normalized) operators \( \rho_x \) on \( B \) such that

\[
\rho^{XB} = \sum_x |x\rangle \langle x|_X \otimes \rho^B_x. \tag{2.2}
\]

### 2.2.2 Evolution of quantum systems

The operations that change the state of a closed quantum system \( A \) are unitary transformations on \( A \). Recall that \( U \in \mathcal{L}(A) \) is unitary if \( UU^\dagger = U^\dagger U = \text{id} \). After applying such a transformation, the state of system \( A \) evolves from \( \rho \) to \( U \rho U^\dagger \). We can also consider a system \( AB \) and act by a unitary on \( A \) to obtain the state \( U_A \rho_A U_A^\dagger = (U \otimes \text{id}_B) \rho_{AB} (U \otimes \text{id}_B)^\dagger \).

Another important class of quantum operations are measurements. The most general way to obtain classical information from a quantum state is by performing a measurement. A measurement is described by a positive operator-valued measure (POVM), which is a set \( \{P_1, \ldots, P_s\} \) of positive semidefinite operators that sum to the identity. If the state of the quantum system is represented by the density operator \( \rho \), the probability of observing the outcome labeled \( i \) is \( \text{tr}[P_i \rho] \) for all \( i \in \{1, \ldots, s\} \). Whenever \( \{P_i\} \) are orthogonal projectors, we say that \( \{P_i\} \) is a projective measurement. A simple class of measurements that will be extensively used in this thesis are measurements in a basis \( B \). The measurement in the basis \( B = \{ |e_i\rangle \}_{i \in [d_A]} \) is defined by the POVM described by the operators \( \{|e_i\rangle \langle e_i| \}_{i \in [d_A]} \).
so that we obtain outcome $i$ with probability $\text{tr}[e_i\langle e_i|\rho] = \langle e_i|\rho|e_i\rangle$ whenever the state of the system is $\rho$. In particular, if the measurement is in the computational basis, we use the special notation $p_\rho(a) = \text{tr}[|a\rangle\langle a|\rho]$, and $p_\psi(a) = |\langle a|\psi\rangle|^2$ whenever the state $\rho = |\psi\rangle\langle \psi|$ is pure.

More generally, we can represent the evolution of any quantum system by a completely positive trace preserving (CPTP) map $\mathcal{E}_{A \rightarrow C}$. A map is called positive if for any positive operator $\rho$, $\mathcal{E}(\rho)$ is also positive. It is called completely positive if for any quantum system $B$, the map $\mathcal{E} \otimes \text{id}_B : \mathcal{L}(A \otimes B) \rightarrow \mathcal{L}(C \otimes B)$ is positive. Because this is the most general kind of quantum operation, a CPTP map is also called a quantum channel.

We can view a measurement as a quantum channel that maps a quantum system to a classical one. In particular, the map that performs a measurement in the computational basis can be written as:

$$\mathcal{M}(\cdot)_{A \rightarrow X} = \sum_a \langle a|\cdot|a\rangle |a\rangle\langle a|,$$

where $\{|a\rangle\}$ is the computational basis of $A$. Note that we renamed the system $X$ to emphasize that it is a classical system. We will also use extensively in Chapter 5 the map

$$\mathcal{T}(\cdot)_{A \rightarrow A_1} = \sum_{a_1a_2} \langle a_1a_2|\cdot|a_1a_2\rangle |a_1\rangle\langle a_1|,$$

where $\{|a_1\rangle\}, \{|a_2\rangle\}$ are the computational bases of $A_1, A_2$ respectively. A small calculation readily reveals that this map can be understood as tracing out $A_2$, and then measuring the remaining system $A_1$ in the basis $\{|a_1\rangle\}$. Note that the outcome of the measurement map is classical in the basis $\{|a_1\rangle\}$ on $A_1$.

### 2.2.3 Distance measures

We will employ two well known distance measures between quantum states. The first is the distance induced by the $\ell_1$-norm defined by $\|M\|_1 = \text{tr}[\sqrt{M^\dagger M}]$. For $\rho, \sigma \in \mathcal{S}(A)$, $\|\rho - \sigma\|_1$ is the sum of the absolute values of the eigenvalues of $\rho - \sigma$. As in the classical case, one half of the $\ell_1$-norm of a difference of two density operators, also known as the trace distance $\Delta(\rho, \sigma) = \frac{1}{2} \cdot \|\rho - \sigma\|_1$, is related to the success probability of distinguishing two states $\rho$ and $\sigma$ given with a priori equal probability [Helstrom, 1967]:

$$\Delta(\rho, \sigma) = \max_{0 \leq \Lambda \leq \text{id}} \text{tr}[\Lambda(\rho - \sigma)].$$

(2.5)

The second distance measure we use is the purified distance. To define it, we first define the fidelity between two states $\rho, \sigma \in \mathcal{S}(A)$ by $F(\rho, \sigma) = \|\sqrt{\rho}\sqrt{\sigma}\|_1$. Note that if $\rho = |\psi\rangle\langle \psi|$
is pure, then $F(\rho, \sigma) = \sqrt{\langle \psi | \sigma | \psi \rangle}$. Another useful characterization of the fidelity is with Uhlmann’s theorem. Before stating the theorem, we need to define the important notion of a purification. A purification of a density operator $\rho \in \mathcal{S}(A)$ is a pure state $|\rho\rangle \in AR$ such that $\text{tr}_R[\rho^{AR}] = \rho^A$. Such a purification always exists, for example one can choose $R$ to be a copy of $A$ and $|\rho\rangle = \sum_i \sqrt{p_i} |\psi_i\rangle^A |\psi_i\rangle^R$, where $\{|\psi_i\rangle\}_i$ is an eigenbasis for $\rho^A$.

**Theorem 2.2.1** (Uhlmann’s theorem [Uhlmann, 1976]). Let $\rho, \sigma \in \mathcal{S}(A)$ and let $|\rho\rangle^{AR}$ and $|\sigma\rangle^{AR}$ be purifications of $\rho$ and $\sigma$. Then we have

$$F(\rho, \sigma) = \max_U |\langle \rho | U^R \otimes \text{id}^A | \sigma \rangle|.$$

See e.g., [Wilde, 2011, Theorem 9.2.1] for a proof. We will also need the concept of generalized fidelity between two possibly sub-normalized positive operators $\rho, \sigma$, which can be defined as [Tomamichel et al., 2010],

$$\tilde{F}(\rho, \sigma) = F(\rho, \sigma) + \sqrt{(1 - \text{tr}[\rho])(1 - \text{tr}[\sigma])}.$$

Note that if at least one of the states is normalized, then the generalized fidelity is the same as the fidelity, i.e., $\tilde{F}(\rho, \sigma) = F(\rho, \sigma)$. The purified distance between two possibly subnormalized states $\rho, \sigma$ is then defined as:

$$P(\rho, \sigma) = \sqrt{1 - \tilde{F}(\rho, \sigma)^2}, \quad (2.6)$$

and is a metric on the set of sub-normalized states [Tomamichel, 2012, Tomamichel et al., 2010].

Observe that for pure states $\sqrt{1 - F(|\rho\rangle\langle \rho|, |\sigma\rangle\langle \sigma|)^2} = \frac{1}{2} |||\rho\rangle\langle \rho| - |\sigma\rangle\langle \sigma||_1$. Hence, by Uhlmann’s theorem, we can think of the purified distance between two normalized states as the minimal trace distance between any two purifications of the states $\rho$ and $\sigma$. The purified distance is indeed closely related to the trace distance, as for any two states $\rho, \sigma$ we have [Fuchs and van de Graaf, 1999, Tomamichel et al., 2010]:

$$\frac{1}{2} \| \rho - \sigma \|_1 \leq P(\rho, \sigma) \leq \sqrt{2} \| \rho - \sigma \|_1. \quad (2.7)$$

It is furthermore easy to see that for normalized states the factor 2 on the right hand side can be improved to 1.

For any distance measure, we can define an $\epsilon$-ball of states around $\rho$ as the states at a distance of at most $\epsilon$ from $\rho$. For the purified distance, we write

$$\mathcal{B}_\epsilon(\rho_A) = \{ \sigma_A \in \mathcal{S}_\leq(A) \mid P(\rho_A, \sigma_A) \leq \epsilon \},$$
where $S_{\leq}(A)$ is the set of positive operators on $A$ with trace at most 1.

All the distances we introduced have the property that they cannot increase by applying a completely positive trace preserving map $F : S(A) \to S(C)$. For any $\rho, \sigma \in S_{\leq}(A)$, we have

$$P(\rho, \sigma) \geq P(F(\rho), F(\sigma)), \quad (2.8)$$

and

$$\|\rho - \sigma\|_1 \geq \|F(\rho) - F(\sigma)\|_1. \quad (2.9)$$

### 2.2.4 Information measures

The von Neumann entropy of $\rho \in S(A)$ is defined as $H(A)_{\rho} = -\text{tr}[\rho \log \rho]$. Note that for a classical state $\rho_X$ this is simply the Shannon entropy defined earlier. The conditional von Neumann entropy of $A$ given $B$ for $\rho_{AB} \in S(AB)$ is defined as

$$H(A|B)_{\rho} = H(AB)_{\rho} - H(B)_{\rho}. \quad (2.10)$$

There is an important difference with the classical case: $H(A|B)_{\rho}$ can be negative when the state $\rho$ is entangled between $A$ and $B$. The conditional min-entropy of a state $\rho_{AB} \in S(AB)$ defined as

$$H_{\text{min}}(A|B)_{\rho} = \max_{\sigma_B \in S(B)} H_{\text{min}}(A|B)_{\rho|\sigma}, \quad (2.10)$$

with

$$H_{\text{min}}(A|B)_{\rho|\sigma} = \max \{\lambda \in \mathbb{R} : \rho_{AB} \leq 2^{-\lambda} \cdot \text{id}_A \otimes \sigma_B\}. \quad (2.10)$$

For the special case where $B$ is trivial, we obtain $H_{\text{min}}(A)_{\rho} = -\log \|\rho_A\|_\infty$, where $\|\rho\|_\infty$ denotes the largest singular value of $\rho$. For the case where we are conditioning on classical side information, we can write the conditional min-entropy as:

$$H_{\text{min}}(X|QJ) = -\log \mathbb{E}_j \left\{2^{-H_{\text{min}}(X|Q,J=j)}\right\}. \quad (2.11)$$

The min-entropy is known to have interesting operational interpretations [König et al., 2009]. If $A$ is classical, then the min-entropy can be expressed as

$$H_{\text{min}}(A|B)_{\rho} = -\log P_{\text{guess}}(A|B), \quad (2.12)$$

where $P_{\text{guess}}(A|B)$ is the average probability of guessing the classical symbol $A = a$ maximized over all possible measurements on $B$. If $A$ is quantum, then $H_{\text{min}}(A|B)_{\rho}$ is

---

2 We write $\max$ instead of $\sup$ as we work with finite dimensional Hilbert spaces.
directly related to the maximal singlet fraction achievable by performing an operation on $B$:

$$H_{\min}(A|B)_{\rho} = -\log |A| \max_{A \to A'} F(\Phi_{AA'}, (\text{id}_A \otimes \Lambda)(\rho_{AB})) ,$$

(2.13)

where $\Phi_{AA'} = \frac{1}{|A|} \sum_{a,a' \in |A|} |aa\rangle \langle a'|$ is a maximally entangled state.

As the information theoretic tasks we wish to study usually allow for some error $\epsilon \geq 0$, the relevant entropy measures are often smoothed entropies. For the conditional min-entropy this takes the form

$$H_{\epsilon \min}(A|B)_{\rho} = \max_{\tilde{\rho}_{AB} \in S^r(\rho_{AB})} H_{\min}(A|B)_{\tilde{\rho}} .$$

(2.14)

**More technical properties of entropic quantities**

In this section, we state some additional entropic quantities that will be needed for some proofs.

It will sometimes be more convenient to work with a version of the min-entropy in which instead of maximizing over all states $\sigma_B$ on $B$, we simply take $\sigma_B = \rho_B$. The reason the standard definition of the conditional min-entropy involves a maximization as in equation (2.10) is to obtain the nice operational interpretation presented above. In particular, if the systems $A$ and $B$ are classical taking discrete values $\{a\}$ and $\{b\}$, then $H_{\min}(A|B)_{\rho|\rho} = -\log \max_{a,b} \langle ab|\rho|ab\rangle$, which is in general different from equation (2.11).

The smoothed version of this alternative definition becomes

$$H_{\epsilon \min}(A|B)_{\rho|\rho} = \max_{\tilde{\rho}_{AB} \in S^r(\rho_{AB})} H_{\min}(A|B)_{\tilde{\rho}|\tilde{\rho}} .$$

Tomamichel et al. [2011] showed that the smoothed versions of the two different definitions cannot be too far apart from each other.

**Lemma 2.2.2** ([Tomamichel et al., 2011] Lemma 18). Let $\epsilon' \geq 0$, $\epsilon' > 0$, and $\rho_{AB} \in S(AB)$. Then

$$H_{\min}(A|B)_{\rho|\rho} - \log \left( \frac{2}{\epsilon'^2} + \frac{1}{1 - \epsilon'} \right) \leq H_{\min}(A|B)_{\rho|\rho} \leq H_{\min}(A|B)_{\rho|\rho} .$$

The max-entropy is defined by

$$H_{\max}(A|B)_{\rho} = \max_{\sigma_B \in S(B)} \log F(\rho_{AB}, \text{id}_A \otimes \sigma_B) ,$$

(2.15)

and its smooth version

$$H_{\epsilon \max}(A|B)_{\rho} = \min_{\rho_{AB} \in S^r(\rho_{AB})} H_{\max}(A|B)_{\rho} .$$

(2.16)
The following lemma shows that the conditional min- and max-entropies are dual to one another.

**Lemma 2.2.3** ([Tomamichel et al. 2010]). Let $\rho_{AB} \in S(AB)$, $\epsilon \geq 0$, and $\rho_{ABC}$ be an arbitrary purification of $\rho_{AB}$. Then

$$H_{\min}^\epsilon(A|B)_{\rho} = -H_{\max}^\epsilon(A|C)_{\rho}.$$  

Finally, the quantum conditional collision entropy, which is closely related to the min-entropy, will be used in the proofs in Chapter 5. For a state $\rho_{AB} \in S(AB)$ relative to a state $\sigma_B \in S(B)$, it is defined as

$$H_2(A|B)_{\rho|\sigma} = -\log \text{tr} \left[ (\text{id}_A \otimes \sigma_B^{-1/2}) \rho_{AB} (\text{id}_A \otimes \sigma_B^{-1/2}) \right]^2,$$

(2.17)

where the inverses are generalized inverses. For $M \in \mathcal{L}(A)$, $M^{-1}$ is a generalized inverse of $M$ if $MM^{-1} = M^{-1}M = \Pi_S$, where $\Pi_S$ denotes the projector onto the support of $M$. In particular, if $M = \sum_i \alpha_i |v_i\rangle\langle v_i|$ and the vectors $|v_i\rangle$ are orthogonal with unit norm, then $M^{-1} = \sum_{i: \alpha_i \neq 0} \alpha_i^{-1} |v_i\rangle\langle v_i|$.

The following lemma relates the collision and the min-entropy.

**Lemma 2.2.4.** Let $\rho_{AB} \in S_{\leq}(AB)$ and $\sigma_B \in S(B)$ with $\text{supp}(\rho_{AB}) \subseteq \text{id}_A \otimes \text{supp}(\sigma_B)$, where $\text{supp}(.)$ denotes the support. Then

$$H_{\min}(A|B)_{\rho|\sigma} \leq H_2(A|B)_{\rho|\sigma}.$$  

**Proof** We have $\text{supp}(\rho_{AB}) \subseteq \text{id}_A \otimes \text{supp}(\rho_B)$ and hence by [Berta et al. 2011b. Lemma B.2]

$$H_{\min}(A|B)_{\rho|\sigma} = -\log \max_{\omega_{AB} \in S(AB)} \text{tr} \left[ \omega_{AB} \left( \text{id}_A \otimes \sigma_B^{-1/2} \right) \rho_{AB} \left( \text{id}_A \otimes \sigma_B^{-1/2} \right) \right],$$

where the inverses are generalized inverses. But for $\hat{\rho}_{AB} = \frac{\rho_{AB}}{\text{tr}[\rho_{AB}]} \in S(AB)$ we have,

$$H_2(A|B)_{\rho|\sigma} = -\log \left[ \rho_{AB} \left( \text{id}_A \otimes \sigma_B^{-1/2} \right) \rho_{AB} \left( \text{id}_A \otimes \sigma_B^{-1/2} \right) \right]$$

$$= -\log \left[ \rho_{AB} \right] - \log \text{tr} \left[ \hat{\rho}_{AB} \left( \text{id}_A \otimes \sigma_B^{-1/2} \right) \hat{\rho}_{AB} \left( \text{id}_A \otimes \sigma_B^{-1/2} \right) \right]$$

$$\geq -\log \max_{\omega_{AB} \in S(AB)} \text{tr} \left[ \omega_{AB} \left( \text{id}_A \otimes \sigma_B^{-1/2} \right) \rho_{AB} \left( \text{id}_A \otimes \sigma_B^{-1/2} \right) \right]$$

$$= H_{\min}(A|B)_{\rho|\sigma}.$$
We finish with three diverse lemmas that will be used several times. First the Alicki-Fannes inequality states that two states that are close in trace distance have von Neumann entropies that are close.

**Lemma 2.2.5** (Alicki and Fannes [2003]). For any states $\rho^{AB}$ and $\sigma^{AB}$ such that $\|\rho^{AB} - \sigma^{AB}\|_1 \leq \epsilon$ with $\epsilon \leq 1/2$, we have

$$|H(A|B)\rho - H(A|B)\sigma| \leq 4\epsilon \log d_A + 2h_2(\epsilon),$$

where $h_2(\epsilon) = -\epsilon \log \epsilon - (1 - \epsilon) \log(1 - \epsilon)$ is the binary entropy function.

For a reference, see [Wilde, 2011, Theorem 11.9.4]. Note that such a statement is not true of the min- and max-entropies, and it is for this reason that it is useful to define smoothed versions.

The next lemma says that if you discard a classical system, the min-entropy can only decrease.

**Lemma 2.2.6** ([Berta et al., 2011c, Lemma C.5]). Let $\rho^{AXB} \in S(AXB)$, $\epsilon \geq 0$, with $X$ classical. Then

$$H^\epsilon_{\text{min}}(A X | B) \rho \geq H^\epsilon_{\text{min}}(A | B) \rho.$$

The last lemma we present here states that for states of the form $\rho^\otimes n$, the smooth min-entropy converges to the von Neumann entropy when the number of copies $n$ grows. This is called the asymptotic equipartition property (AEP) for the smooth conditional min-entropy.

**Lemma 2.2.7** ([Tomamichel et al., 2009, Remark 10]). Let $\rho_{AB} \in S(AB)$, $\epsilon > 0$, and $n \geq 2 (1 - \epsilon^2)$. Then,

$$\frac{1}{n} H^\epsilon_{\text{min}}(A|B)_{\rho^\otimes n|\rho^\otimes n} \geq H(A|B)_{\rho} - \frac{4\sqrt{1 - 2\log \epsilon} \left(2 + \frac{\log |A|}{2}\right)}{\sqrt{n}}.$$

For a more detailed discussion of smooth entropies we refer to [Renner, 2008, Tomamichel, 2012].

### 2.3 Quantum computation

The most widely used model for quantum computation is the quantum circuit model. Let $U$ be a unitary acting on an $n$-qubit space. The objective is to implement $U$ with a small
number of fixed gates. The main measure of efficiency is then the size of the circuit, which is the number of elementary gates that are used to perform the unitary. We say that a circuit is efficient if the size of the circuit is polynomial in $n$.

There are many standard choices of sets of one and two-qubit gates that allow the approximation of all unitary transformations on $n$ qubits. This choice is not important here. The properties of quantum circuits we use here are the following. The Hadamard single-qubit gate defined by

$$H = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$$

is part of our elementary gates. And any reversible classical circuit on $n$ bits can be directly extended to a quantum circuit with the same size that acts on the computational basis elements in the same way as the classical circuit.
Chapter 3

Uncertainty relations for quantum measurements: Definition and constructions

Outline of the chapter  In this chapter, we start by introducing uncertainty relations and setting up some notation (Section 3.1). Then, we define metric uncertainty relations in Section 3.2. In Section 3.3, we prove the existence of strong metric uncertainty relations. Explicit constructions are given in Section 3.4.

3.1 Background

In quantum mechanics, an uncertainty relation is a statement about the relationship between measurements (or observables). Heisenberg’s uncertainty principle [Heisenberg, 1927] is one of the cornerstones of quantum mechanics. It states that the position and the momentum of a quantum particle cannot both have definite values. The uncertainty principle is a feature of quantum theory that makes it different from classical physics: having both a localized position and momentum is not a valid state according to quantum theory.

Heisenberg’s uncertainty relation was generalized in several ways. The most common way of presenting the uncertainty principle today is due to Robertson [1929]. It gives a lower bound on the product of the variances of two observables as a function of their commutator, which quantifies how compatible the two observables are. Later, Hirschman [1957] and

\footnote{In physics language, it probably makes more sense to use the word observable rather than measurement, but as we have not given a mathematical definition of an observable, we mostly use the word measurement.}
Bialynicki-Birula and Mycielski [1975] gave a formulation of an uncertainty relation in terms of the entropy of the measurement outcomes. Deutsch [1983] pointed out that using an entropy instead of the variance is a more desirable way of expressing uncertainty. He proved that for any state $|\psi\rangle$, we have

$$H(p_{B_1,|\psi\rangle}) + H(p_{B_2,|\psi\rangle}) \geq -\log \left( \frac{1+c(B_1,B_2)}{2} \right)$$

where $c(B_1,B_2) = \max_{b_1 \in B_1, b_2 \in B_2} |\langle b_1 | b_2 \rangle|$ and $B_1$ and $B_2$ are bases of the ambient Hilbert space. $p_{B,|\psi\rangle}$ denotes the outcome distribution when performing a measurement in $B$ on the state $|\psi\rangle$ and $H$ denotes the Shannon entropy. This uncertainty relation was later improved by Maassen and Uffink [1988] who showed that for all $|\psi\rangle$,

$$\frac{1}{2} \left( H(p_{B_1,|\psi\rangle}) + H(p_{B_2,|\psi\rangle}) \right) \geq -\log c(B_1,B_2). \quad (3.1)$$

Observe that by using the properties of the Shannon entropy, we can rewrite equation (3.1) as $H(X|K) \geq -\log c(B_1,B_2)$, where $K$ is uniformly distributed on $\{1, 2\}$ and $X$ is the outcome of a measurement in the computational basis for the state $U_K|\psi\rangle$. This says that even given the measurement $K$ that was performed, there is some uncertainty about the outcome. If $B_1$ and $B_2$ are mutually unbiased, i.e., $c(B_1,B_2) \leq 2^{-n/2}$ where $2^n$ is the dimension of the ambient Hilbert space, we obtain a lower bound of $\frac{n}{2}$ on the average measurement entropy. It is easy to see that such a lower bound cannot be improved: For any bases $B_1, B_2$, one can always choose a state $|\psi_1\rangle$ that is aligned with one of the vectors of $B_1$ so that $H(p_{B_1,|\psi_1\rangle}) = 0$, in which case $\frac{1}{2} \left( H(p_{B_1,|\psi_2\rangle}) + H(p_{B_2,|\psi_1\rangle}) \right) \leq \frac{n}{2}$. More generally when considering $t$ basis, the best lower bound on the average measurement entropy one can hope for is $(1 - 1/t)n$.

For many applications, an average measurement entropy of $\frac{n}{2}$ is not good enough. In this chapter, we want to find bases for which the average measurement entropy is larger than $\frac{n}{2}$ and close to the maximal value of $n$. As mentioned earlier, in order to achieve this, one has to consider a larger set of measurements. In this case, the natural candidate is a set of $t$ mutually unbiased bases, the defining property of which is a small inner product between any pair of vectors in different bases, more precisely $c(B_i,B_j) \leq 2^{-n/2}$ for all $i \neq j$. For $2^n + 1$ measurements, [Ivanovic [1992], Larsen [1990], Sanchez [1993] showed for $t = 2^n + 1$ mutually unbiased bases, the average entropy is at least $\log(2^n + 1) - 1$, which is close to the best possible. In fact, their result is stronger: it even holds for the collision entropy (Rényi entropy of order 2), which is in general smaller than the Shannon entropy. For $2 < t < 2^n + 1$, the behaviour of mutually unbiased bases is not well understood. The best general bound for an incomplete set of mutually unbiased bases was proved by
Damgård et al. [2004] and Azarchs [2004]:

$$\frac{1}{t} \sum_{k=1}^{t} H(p_{B_k|\psi}) \geq n + \log \left( \frac{t}{2^n + t - 1} \right).$$  (3.2)

Observe that this bound is not useful for \( t \leq 2^{n/2} \), because in this case the term \( \log(t/(2^n + t - 1)) \leq -n/2 \), which makes (3.2) at best as good as the uncertainty relation for two measurements in equation (3.1). Equation (3.2) is known to also hold for the collision entropy. A similar bound for the min-entropy was also proved in [Schaffner, 2007, Corollary 4.19]. Surprisingly, it was shown by Ballester and Wehner [2007] and Ambainis [2010] that there are up to \( t = \sqrt{2^n} \) mutually unbiased bases \( \{B_1, B_2, \ldots, B_t\} \) that only satisfy an average measurement entropy of \( \frac{n}{2} \), which is only as good as what can be achieved with two measurements (3.1). In other words, looking at the pairwise inner product between vectors in different bases is not enough to obtain uncertainty relations stronger than (3.1). To achieve an average measurement entropy of \((1 - \epsilon)n\) for small \( \epsilon \) while keeping the number of bases subexponential in \( n \), the only known constructions are probabilistic and computationally inefficient. Hayden et al. [2004] prove that random bases satisfy entropic uncertainty relations of the form (3.1) with \( n^4 \) measurements with an average measurement entropy of \( n - 3 \).

**Brief word on applications of uncertainty relations** Other than being one of the defining features of quantum mechanics, uncertainty relations have many applications particularly to proving the security of quantum cryptographic protocols. As an example, probably the simplest and most elegant proof of security for quantum key distribution known to date is based on a recently discovered uncertainty relation [Tomamichel and Renner, 2011]. Moreover, the proofs of the security of bit commitment and oblivious transfer in the bounded storage model are based on an uncertainty relation [Damgård et al., 2005, 2007, König et al., 2012]. We will describe several applications of uncertainty relations in Chapter 4 and Section 5.4. For more details on entropic uncertainty relations and their applications, see the survey [Wehner and Winter, 2010].

**Notation** Instead of talking about uncertainty relations for a set of bases, it is more convenient here to talk about uncertainty relations for a set of unitary transformations. Let \( \{|x\rangle^C\}_x \) be the computational basis of \( C \). We associate to the unitary transformation \( U \) the basis \( \{U|x\rangle\}_x \). On a state \(|\psi\rangle\), the outcome distribution is described by

$$p_{U|\psi}(x) = |\langle x|U|\psi\rangle|^2.$$  

As can be seen from this equation, we can equivalently talk about measuring the state \( U|\psi\rangle \) in the computational basis. An entropic uncertainty relation for \( U_1, \ldots, U_t \) can be written as
3.2 Metric uncertainty relations

Even though entropy is a good measure of randomness, it is usually easier to work with the distance to the uniform distribution when the distance is small. This will be our approach here: our measure of uncertainty will be the closeness in total variation distance to the uniform distribution. In other words, we are interested in sets of unitary transformations $U_1, \ldots, U_t$ that for all $|\psi\rangle \in C$ satisfy

$$\frac{1}{t} \sum_{k=1}^{t} \Delta(p_{U_k|\psi}), \text{unif}([d_C])) \leq \epsilon$$

for some $\epsilon \in (0, 1)$. $\Delta(p, q)$ refers to the total variation distance between distributions $p$ and $q$. This condition is very strong, in fact too strong for our purposes, and we will see that a weaker definition is sufficient to imply entropic uncertainty relations. Let $C = A \otimes B$. (For example, if $C$ consists of $n$ qubits, $A$ might represent the first $n - \log n$ qubits and $B$ the last $\log n$ qubits.) Moreover, let the computational basis for $C$ be of the form $\{|a\rangle_A \otimes |b\rangle_B\}_{a,b}$ where $\{|a\rangle\}$ and $\{|b\rangle\}$ are the computational bases of $A$ and $B$. Instead of asking for the outcome of the measurement on the computational basis of the whole space to be uniform, we only require that the outcome of a measurement of the $A$ system in its computational basis $\{|a\rangle\}$ be close to uniform. More precisely, we define for $a \in [d_A]$,

$$p_{U_k|\psi}^A(a) = \sum_{b=1}^{d_B} |\langle a |^A \langle b |^B U_k |\psi\rangle|^2.$$

We can then define a metric uncertainty relation. Naturally, the larger the $A$ system, the stronger the uncertainty relation for a fixed $B$ system.

**Definition 3.2.1** (Metric uncertainty relation). *Let $A$ and $B$ be Hilbert spaces. We say that a set $\{U_1, \ldots, U_t\}$ of unitary transformations on $AB$ satisfies an $\epsilon$-metric uncertainty relation on $A$ if for all states $|\psi\rangle \in AB$,

$$\frac{1}{t} \sum_{k=1}^{t} \Delta(p_{U_k|\psi}^A), \text{unif}([d_A])) \leq \epsilon.$$*
Remark. Observe that this implies that (3.4) also holds for mixed states: for any \( \psi \in \mathcal{S}(A \otimes B) \), \( \frac{1}{t} \sum_{k=1}^{t} \Delta \left( p_{U_k, \psi}^A, \text{unif}(\mathcal{d}_A) \right) \leq \epsilon \).

Note that there is a reason we are looking at the average over the different values of \( k \) rather than some other quantity. In fact we can rewrite the condition (3.4) as

\[
\Delta \left( q_{|\psi\rangle}, \text{unif}(\mathcal{d}_A) \times \text{unif}(\mathcal{t}) \right) \leq \epsilon,
\]

where \( q_{|\psi\rangle} \) is the distribution on \( \mathcal{d}_A \times \mathcal{t} \) of the random variable \( (X, K) \), where \( X \) refers to the outcome of the computational basis measurement when it is performed on state \( U_K |\psi\rangle \).

This means that even given the measurement \( K \) that was performed, the outcome of the measurement is still \( \epsilon \)-close to uniform.

**Metric uncertainty relations imply entropic uncertainty relations** In the next proposition, we show that a metric uncertainty relation implies an entropic uncertainty relation.

**Proposition 3.2.2.** Let \( \epsilon \in (0, 1/2) \) and \( \{U_1, \ldots, U_t\} \) be a set of unitaries on \( AB \) satisfying an \( \epsilon \)-metric uncertainty relation on \( A \):

\[
\frac{1}{t} \sum_{k=1}^{t} \Delta \left( p_{U_k, \psi}^A, \text{unif}(\mathcal{d}_A) \right) \leq \epsilon.
\]

Then

\[
\frac{1}{t} \sum_{k=1}^{t} H(p_{U_k, \psi}^A) \geq (1 - 8\epsilon) \log d_A - 2h_2(2\epsilon).
\]

where \( h_2(\epsilon) = -\epsilon \log \epsilon - (1 - \epsilon) \log(1 - \epsilon) \) is the binary entropy function.

**Proof** Recall that the distribution \( p_{U_k, \psi}^A \) (see equation (3.4) for a definition) on \( \mathcal{d}_A \) is a marginal of the distribution \( p_{U_k}^A |\psi\rangle \). Thus \( H(p_{U_k}^A |\psi\rangle) \geq H(p_{U_k, \psi}^A |\psi\rangle) \). Using Fannes’ inequality (a special case of the Alicki-Fannes inequality [2.2.5]), we have for all \( k \)

\[
H(p_{U_k, \psi}^A) \geq \log d_A - 8\Delta \left( p_{U_k, \psi}^A, \text{unif}(\mathcal{d}_A) \right) \log d_A - 2h_2 \left( 2\Delta \left( p_{U_k, \psi}^A, \text{unif}(\mathcal{d}_A) \right) \right).
\]

By averaging over \( k \), and using the concavity of \( h_2 \), we obtain the desired result. \qed

**Explicit link to low-distortion embeddings** Even though we do not explicitly use the link to low-distortion embeddings, we describe the connection as it might have other applications. In the definition of metric uncertainty relations, the distance between
distributions was computed using the trace distance. The connection to low-distortion metric embeddings is clearer when we measure closeness of distributions using the fidelity. We have

\[
F\left( p^A_{U_k|\psi}, \text{unif}(|d_A|) \right) = \frac{1}{\sqrt{d_A}} \sum_{a=1}^{d_A} \sqrt{p^A_{U_k|\psi}(a)}
\]

\[
= \frac{1}{\sqrt{d_A}} \sum_{a=1}^{d_A} \left| \sum_{b=1}^{d_B} |\langle a|A\rangle \langle b|B \rangle U_k|\psi \rangle|^2 \right|
\]

\[
= \frac{1}{\sqrt{d_A}} \| U_k|\psi \rangle \|_{\ell^A_1(\ell^B_2)}
\]

where the norm \( \ell^A_1(\ell^B_2) \) is defined by

**Definition 3.2.3** (\( \ell^1_1(\ell^2_2) \) norm). For a state \( |\psi\rangle = \sum_{a,b} \alpha_{a,b} |a\rangle^A |b\rangle^B \),

\[
\| |\psi\rangle \|_{\ell^A_1(\ell^B_2)} = \sum_a \| \alpha_{a,b} \|_b = \sum_a \sqrt{\sum_b |\alpha_{a,b}|^2}.
\]

We use \( \| \cdot \|_{12} \overset{\text{def}}{=} \| \cdot \|_{\ell^A_1(\ell^B_2)} \) when the systems \( A \) and \( B \) are clear from the context.

Observe that this definition of norm depends on the choice of the computational basis. The \( \ell^A_1(\ell^B_2) \) norm will always be taken with respect to the computational bases.

For \( \{ U_1, \ldots, U_t \} \) to satisfy an uncertainty relation, we want

\[
\frac{1}{t} \sum_k \frac{1}{\sqrt{d_A}} \| U_k|\psi \rangle \|_{\ell^A_1(\ell^B_2)} \geq 1 - \epsilon.
\]

This expression can be rewritten by introducing a new register \( K \) that holds the index \( k \). We get for all \( |\psi\rangle \)

\[
\left\| \frac{1}{\sqrt{t}} \sum_k U_k|\psi\rangle^C |k\rangle^K \right\|_{\ell^A_1(\ell^B_2)} \geq (1 - \epsilon) \sqrt{t \cdot d_A}.
\] (3.6)

Using the Cauchy-Schwarz inequality, we have that for all \( |\psi\rangle \),

\[
\left\| \frac{1}{\sqrt{t}} \sum_k U_k|\psi\rangle^C |k\rangle^K \right\|_{\ell^A_1(\ell^B_2)} \leq \sqrt{t \cdot d_A} \left\| \frac{1}{\sqrt{t}} \sum_k U_k|\psi\rangle^C |k\rangle^K \right\|_2 = \sqrt{t \cdot d_A}.
\] (3.7)

Rewriting (3.6) and (3.7) as

\[
(1 - \epsilon) \leq \frac{1}{\sqrt{t \cdot d_A}} \left\| \frac{1}{\sqrt{t}} \sum_k U_k|\psi\rangle^C |k\rangle^K \right\|_{\ell^A_1(\ell^B_2)} \leq 1,
\]
we see that the image of $C$ by the linear map $|\psi\rangle \mapsto \frac{1}{\sqrt{t}} \sum_k U_k |\psi\rangle \otimes |k\rangle$ is an almost Euclidean subspace of $(A \otimes K \otimes B, \ell_1^{AK}(\ell_2^B))$. In other words, as the map $|\psi\rangle \mapsto \frac{1}{\sqrt{t}} \sum_k U_k |\psi\rangle \otimes |k\rangle$ is an isometry (in the $\ell_2$ sense), it is an embedding of $(C, \ell_2)$ into $(AKB, \ell_1^{AK}(\ell_2^B))$ with distortion $1/(1 - \epsilon)$.

Observe that a general low-distortion embedding of $(C, \ell_2)$ into $(AKB, \ell_1^{AK}(\ell_2^B))$ does not necessarily give a metric uncertainty relation as it need not be of the form $|\psi\rangle \mapsto \frac{1}{\sqrt{t}} \sum_k U_k |\psi\rangle \otimes |k\rangle$. When $t = 2$, a metric uncertainty relation is related to the notion of Kashin decomposition [Kashin, 1977]; see also [Pisier, 1989, Szarek, 2006].

A remark on the composition of metric uncertainty relations There is a natural way of building an uncertainty relation for a Hilbert space from uncertainty relations on smaller Hilbert spaces. This composition property is also important for the cryptographic applications of metric uncertainty relations presented in Chapter 4 in which setting it ensures the security of parallel composition of locking schemes.

**Proposition 3.2.4.** Consider Hilbert spaces $A_1$, $A_2$, $B_1$, $B_2$. For $i \in \{1, 2\}$, let $\{U_{k_i}^{(i)}\}_{k_i \in [t_i]}$ be a set of unitary transformations of $A_i \otimes B_i$ satisfying an $\epsilon$-metric uncertainty relation on $A_i$. Then, $\{U_{k_1}^{(1)} \otimes U_{k_2}^{(2)}\}_{k_1, k_2 \in [t_1] \times [t_2]}$ satisfies a $2\epsilon$-metric uncertainty relation on $A_1 \otimes A_2$.

**Proof** Let $|\psi\rangle \in (A_1 \otimes B_1) \otimes (A_2 \otimes B_2)$ and let $p_{k_1, k_2}$ denote the distribution obtained by measuring $U_{k_1}^{(1)} \otimes U_{k_2}^{(2)} |\psi\rangle$ in the computational basis of $A_1 \otimes A_2$. Our objective is to show that

$$\frac{1}{t_1 t_2} \sum_{k_1 \in [t_1], k_2 \in [t_2]} \Delta(p_{k_1, k_2}, \text{unif}([d_{A_1}] \times [d_{A_2}])) \leq 2\epsilon. \quad (3.8)$$

We have

$$\Delta(p_{k_1, k_2}, \text{unif}([d_{A_1}] \times [d_{A_2}])) = \frac{1}{2} \sum_{a_1, a_2} \left| p_{k_1, k_2}(a_1, a_2) - \frac{1}{d_{A_1} d_{A_2}} \right| \quad (3.9)$$

$$\leq \frac{1}{2} \sum_{a_1, a_2} \left| p_{k_1, k_2}(a_1, a_2) - \frac{p_{A_2}^{A_1}(a_1)}{d_{A_2}} \right| + \frac{1}{2} \sum_{a_1, a_2} \left| \frac{p_{A_1}^{A_1}(a_1)}{d_{A_2}} - \frac{1}{d_{A_1} d_{A_2}} \right|$$

$$= \frac{1}{2} \sum_{a_1} p_{A_2}^{A_1}(a_1) \sum_{a_2} \left| p_{k_1, k_2}(a_1, a_2) - \frac{1}{d_{A_2}} \right| + \frac{1}{2} \sum_{a_1} \left| p_{A_1}^{A_1}(a_1) - \frac{1}{d_{A_1}} \right| \quad (3.10)$$

where $p_{k_1, k_2}^{A_1}(a_1) \overset{\text{def}}{=} \sum_{a_2} p_{k_1, k_2}(a_1, a_2)$ is the outcome distribution of measuring the $A_1$ system of $U_{k_1}^{(1)} \otimes U_{k_2}^{(2)} |\psi\rangle$. The distribution $p_{k_1, k_2}$ can also be seen as the outcome of
measuring the mixed state
\[ U_{k_1}^{(1)} \psi_{A_1 B_1} U_{k_1}^{(1)\dagger} \]
in the computational basis \( \{|a_1\rangle\} \). Thus, we have for any \( k_2 \in [t_2] \),
\[ \frac{1}{t_1} \sum_{k_1} \Delta \left( p_{k_1,k_2}^{A_1}, \text{unif}([d_{A_1}]) \right) \leq \epsilon. \]

Moreover, for \( a_1 \in [d_{A_1}] \), the distribution on \([d_{A_2}]\) defined by \( \frac{p_{k_1,k_2}^{A_1(a_1,a_2)}}{p_{k_1,k_2}^{A_1(a_1)}} \) is the outcome distribution of measuring in the computational basis of \( A_2 \) the state
\[ U_{k_2}^{(2)} \psi_{A_2 B_2} U_{k_2}^{(2)\dagger} \]
where \( \psi_{A_2 B_2}^{k_1,a_1} \) is the density operator describing the state of the system \( A_2 B_2 \) given that the outcome of the measurement of the \( A_1 \) system is \( a_1 \). We can now use the fact that \( \{U_{k_2}^{(2)}\} \) satisfies a metric uncertainty relation. Taking the average over \( k_1 \) and \( k_2 \) in equation (3.10), we get
\[ \frac{1}{t_1 t_2} \sum_{k_1,k_2} \Delta \left( p_{k_1,k_2}, \text{unif}([d_{A_1}] \times [d_{A_2}]) \right) \leq 2\epsilon. \]

\[ \Box\]

This observation is in the same spirit as [Indyk and Szarek, 2010, Proposition 1], and can in fact be used to build large almost Euclidean subspaces of \( \ell_1^A(\ell_2^B) \).

### 3.3 Metric uncertainty relations: existence

In this section, we prove the existence of families of unitary transformations satisfying strong uncertainty relations. The proof proceeds by showing that choosing random unitaries according to the Haar measure defines a metric uncertainty relation with positive probability. The techniques used are quite standard and date back to Milman’s proof of Dvoretzky’s theorem [Figiel et al., 1977] [Milman, 1971]. A version of Dvoretzky’s theorem states that for any norm \( \|\cdot\| \) over \( \mathbb{C}^d \), there exists a “large” subspace \( E \subseteq \mathbb{C}^d \) which is almost Euclidean, i.e., for all \( x \in E \), \((1 - \epsilon)\|x\|_2 \leq s\|x\| \leq (1 + \epsilon)\|x\|_2\) for some constant \( \epsilon > 0 \) and scaling factor \( s \). Using the connection between uncertainty relations and embeddings of \( \ell_2 \) into \( \ell_1(\ell_2) \) presented in the previous section, Theorem 3.3.2 can be viewed as a strengthening of Dvoretzky’s theorem for the \( \ell_1(\ell_2) \) norm [Milman and Schechtman, 1986].

General techniques from asymptotic geometric analysis have recently found many applications in quantum information theory. For example, [Aubrun et al., 2010] show that the
existence of large subspaces of highly entangled states follows from Dvoretzky’s theorem for the Schatten $p$-norm\footnote{The Schatten $p$-norm of a matrix $M$ is defined as the $\ell_p$ norm of a vector of singular values of $M$.} for $p > 2$. This in turns shows the existence of channels that violate additivity of minimum output $p$-Rényi entropy as was previously demonstrated by Hayden and Winter \cite{Hayden2008}. Using a more delicate argument, Aubrun et al. \cite{Aubrun2011} were also able to recover Hastings’ counterexample to the additivity conjecture \cite{Hastings2009}. The general strategy that is used to prove such results is to define a distribution over the set of objects one is looking for and use concentration of measure tools to prove that the desired properties can be satisfied with positive probability.

For Theorem 3.3.2, we need to introduce the Haar measure over the unitary group $\mathcal{U}(d)$. A natural way of defining a uniform measure over a group is to ask the measure of a subset to be invariant under multiplication by elements of the group. In particular, for the unitary group, consider measures $\mu$ on the unitary transformations of $\mathbb{C}^d$ that satisfy $\mu(S) = \mu(\{U \cdot M : M \in S\})$ for all measurable sets $S \subseteq \mathcal{U}(d)$ and unitaries $U \in \mathcal{U}(d)$. It follows from Haar’s theorem that there is a unique probability measure that satisfies this condition.

**Definition 3.3.1 (Haar measure).** The Haar measure $\mu_d$ on the set of unitary transformations on $\mathbb{C}^d$ is the unique probability measure that is invariant under multiplication by a unitary operation.

We can then define a rotation invariant probability measure on pure states of $\mathbb{C}^d$ by considering the distribution of $U|0\rangle$ where $U \sim \mu_d$ and $|0\rangle$ is any unit vector in $\mathbb{C}^d$. We say that $U|0\rangle$ is a random pure state.

We need another definition before stating the theorem. For some applications\footnote{Quantum hiding fingerprints studied in Section 4.1.5}, we require an additional property for $\{U_1, \ldots, U_t\}$. A set of unitary transformations $\{U_1, \ldots, U_t\}$ of $\mathbb{C}^d$ is said to define $\gamma$-approximately mutually unbiased bases ($\gamma$-MUBs) if for all elements $|x\rangle$ and $|y\rangle$ of the computational basis and all $k \neq k'$, we have

$$
|\langle x|U_k^\dagger U_{k'}|y\rangle| \leq \frac{1}{d^{1/2}}.
$$

(3.11)

1-MUBs correspond to the usual notion of mutually unbiased bases.

**Theorem 3.3.2 (Existence of metric uncertainty relations).** Let $c = 9\pi^2$ and $\epsilon \in (0, 1)$. Let $A$ and $B$ be Hilbert spaces with $\dim B \geq 9/\epsilon^2$ and $d \overset{\text{def}}{=} \dim A \otimes B \geq \frac{9\epsilon-16\pi}{\epsilon^2}$. Then, for all $t > \frac{4\cdot18\cdot\ln\left(9/\epsilon\right)}{\epsilon^2}$, there exists a set $\{U_1, \ldots, U_t\}$ of unitary transformations of $AB$ satisfying...
an $\epsilon$-metric uncertainty relation on $A$: for all states $|\psi\rangle \in AB$,

$$\frac{1}{t} \sum_{k=1}^{t} \Delta \left( p_{A|\psi}\right), \text{unif}([d_A]) \right) \leq \epsilon.$$ 

Moreover, for $\gamma \in (0, 1)$ and $d$ such that $4t^2d^2 \exp \left(-d^{1-\gamma}\right) < 1/2$, the unitaries \{U_1, \ldots, U_t\} can be chosen to also form $\gamma$-MUBs.

**Remark.** The proof proceeds by choosing a set of unitary transformations at random. See (3.15) and (3.16) for a precise bound on the probability that such a set does not form a metric uncertainty relation or a $\gamma$-MUB.

**Proof** The first step is to evaluate the expected value of $\Delta \left( p_{A|\psi}\right), \text{unif}([d_A]) \right)$ for a fixed state $|\psi\rangle$ when $U$ is a random unitary chosen according to the Haar measure. Then, we use a concentration of measure argument to show that with high probability, this distance is close to its expected value. After this step, we show that the additional averaging $\frac{1}{t} \sum_{k=1}^{t} \Delta \left( p_{A|\psi}\right), \text{unif}([d_A]) \right)$ of $t$ independent copies results in additional concentration at a rate that depends on $t$. We conclude by showing the existence of a family of unitaries that makes this expression small for all states $|\psi\rangle$ using a union bound over a $\delta$-net. The four main ingredients of the proof are precisely stated here but only proved in Appendix A.1.

We start by computing the expected value of the fidelity $\mathbb{E} \left\{ F \left( p_{A|\psi}, \text{unif}([d_A]) \right) \right\}$, which can be seen as an $\ell_1(\ell_2)$ norm.

**Lemma 3.3.3** (Expected value of $\ell_1(\ell_2)$ over the sphere). Let $|\varphi\rangle^{AB}$ be a random pure state on $AB$. Then,

$$\mathbb{E} \left\{ F \left( p_{A|\varphi}, \text{unif}([d_A]) \right) \right\} \geq \sqrt{1 - \frac{1}{d_B}}.$$ 

We then use the inequality $\Delta(\rho, \sigma) \leq \sqrt{1 - F(\rho, \sigma)^2}$ to get

$$\mathbb{E} \left\{ \Delta \left( p_{A|\varphi}\right), \text{unif}([d_A]) \right) \right\} \leq \mathbb{E} \left\{ \sqrt{1 - F \left( p_{A|\varphi}, \text{unif}([d_A]) \right)^2} \right\}.$$ 

By the concavity of the function $x \mapsto \sqrt{1 - x^2}$ on the interval $[0, 1]$,

$$\mathbb{E} \left\{ \Delta \left( p_{A|\varphi}\right), \text{unif}([d_A]) \right) \right\} \leq \sqrt{1 - \mathbb{E} \left\{ F \left( p_{A|\varphi}, \text{unif}([d_A]) \right) \right\}^2} \leq \sqrt{1 - \left(1 - \frac{1}{d_B}\right)} \leq \epsilon/3.$$
The last inequality comes from the hypothesis of the theorem that \( d_B \geq 9/\epsilon^2 \). In other words, for any fixed \( |\psi\rangle \), the average over \( U \) of the trace distance between \( p^A_U|\psi\rangle \) and the uniform distribution is at most \( \epsilon/3 \). The next step is to show that this trace distance is close to its expected value with high probability. For this, we use a version of Lévy’s lemma presented in [Milman and Schechtman [1986]].

**Lemma 3.3.4** (Lévy’s lemma). Let \( f : \mathbb{C}^d \to \mathbb{R} \) and \( \eta > 0 \) be such that for all pure states \( |\varphi_1\rangle, |\varphi_2\rangle \) in \( \mathbb{C}^d \),

\[
|f(|\varphi_1\rangle) - f(|\varphi_2\rangle)| \leq \eta |||\varphi_1\rangle - |\varphi_2\rangle||_2.
\]

Let \( |\varphi\rangle \) be a random pure state in dimension \( d \). Then for all \( 0 \leq \delta \leq \eta \),

\[
Pr \{|f(|\varphi\rangle) - E \{f(|\varphi\rangle)\} \geq \delta\} \leq 4 \exp \left( -\frac{\delta^2 d}{cn^2} \right)
\]

where \( c \) is a constant. We can take \( c = 9\pi^2 \).

We apply this concentration result to \( f : |\varphi\rangle^{AB} \mapsto \Delta(p^A_{|\varphi\rangle}, \text{unif}([d_A])) \). We start by finding an upper bound on the Lipshitz constant \( \eta \). For any pure states \( |\varphi_1\rangle^{AB} \) and \( |\varphi_2\rangle^{AB} \), we have

\[
|f(|\varphi_1\rangle) - f(|\varphi_2\rangle)| \leq \Delta(p^A_{|\varphi_1\rangle}, p^A_{|\varphi_2\rangle})
\]

\[
\leq \frac{1}{2} \sum_{a,b} |\langle a|A\langle b|B|\varphi_1\rangle|^2 - \sum_b |\langle a|A\langle b|B|\varphi_2\rangle|^2| = \Delta(p_{|\varphi_1\rangle}, p_{|\varphi_2\rangle})
\]

\[
\leq \sqrt{1 - F(p_{|\varphi_1\rangle}, p_{|\varphi_2\rangle})^2}
\]

\[
\leq \sqrt{2 \left( 1 - F(p_{|\varphi_1\rangle}, p_{|\varphi_2\rangle}) \right)}
\]

\[
= \sqrt{2 - 2 \sum_{a,b} |\langle a|\langle b||\varphi_1\rangle| \cdot |\langle a|\langle b||\varphi_2\rangle|}
\]

\[
= \sqrt{\sum_{a,b} \|\langle a|\langle b||\varphi_1\rangle| - |\langle a|\langle b||\varphi_2\rangle|\|^2}
\]

\[
\leq |||\varphi_1\rangle - |\varphi_2\rangle||_2.
\]

(3.12)

The first two inequalities follow from the triangle inequality. The third inequality is an application of (2.1). The fourth inequality follows from the fact that \( 1 - x^2 \leq 2(1 - x) \) for all \( x \in [0, 1] \). The last inequality follows again from the triangle inequality. Thus, applying
Lemma 3.3.4 we get for all $0 \leq \delta \leq 1$,

$$\Pr \{ |\Delta(p_{A|\varphi}), \text{unif}([d_A]) - \mu| \geq \delta \} \leq 4 \exp \left( -\frac{\delta^2 d}{c} \right)$$

(3.13)

where $\mu = \mathbb{E} \{ \Delta(p_{A|\varphi}), \text{unif}([d_A]) \}$. The following lemma bounds the tails of the average of independent copies of a random variable.

**Lemma 3.3.5 (Concentration of the average).** Let $a, b \geq 1$, $\delta \in (0, 1)$ and $t$ be a positive integer. Suppose $X$ is a random variable with 0 mean satisfying the tail bounds

$$\Pr \{ X \geq \eta \} \leq ae^{-b\eta^2} \quad \text{and} \quad \Pr \{ X \leq -\eta \} \leq ae^{-b\eta^2}. $$

Let $X_1, \ldots X_t$ be independent copies of $X$. Then if $\delta^2 b \geq 16a^2\pi$,

$$\Pr \left\{ \left| \frac{1}{t} \sum_{k=1}^{t} X_k \right| \geq \delta \right\} \leq \exp \left( -\frac{\delta^2 b t}{2} \right).$$

We apply the above Lemma with $X_k = \Delta(p_{U_k|\psi}), \text{unif}([d_A]) - \mu$ which satisfies the bound (3.13) in addition to being bounded in absolute value by 1. Taking $\delta = \epsilon/3$ and using Lemma 3.3.5 (which we can apply because we have $(\epsilon/3)^2 \cdot \frac{d}{c} \geq 16 \cdot 4^2 \cdot \pi$), we get

$$\Pr \left\{ \left| \frac{1}{t} \sum_{k=1}^{t} \Delta(p_{U_k|\psi}), \text{unif}([d_A]) - \mu \right| \geq \epsilon/3 \right\} \leq \exp \left( -\frac{1}{2} \frac{(\epsilon/3)^2 td}{c} \right).$$

Using this together with Lemma 3.3.3, we have

$$\Pr \left\{ \left| \frac{1}{t} \sum_{k=1}^{t} \Delta(p_{U_k|\psi}), \text{unif}([d_A]) \right| \geq 2\epsilon/3 \right\} \leq \exp \left( -\frac{\epsilon^2 td}{18c} \right).$$

(3.14)

We would like to have the event described in (3.14) hold for all $|\psi\rangle \in AB$. For this, we construct a finite set $\mathcal{N}$ of states (a $\delta$-net) for which we can ensure that $\frac{1}{t} \sum_{k=1}^{t} \Delta(p_{U_k|\psi}), \text{unif}([d_A]) < 2\epsilon/3$ for all $|\psi\rangle \in \mathcal{N}$ holds with high probability.

**Lemma 3.3.6 ($\delta$-net).** Let $\delta \in (0, 1)$. There exists a set $\mathcal{N}$ of pure states in $\mathbb{C}^d$ with $|\mathcal{N}| \leq (3/\delta)^{2d}$ such that for every pure state $|\psi\rangle \in \mathbb{C}^d$ (i.e., $\|\psi\|_2 = 1$), there exists $\tilde{\psi} \in \mathcal{N}$ such that

$$\|\psi - \tilde{\psi}\|_2 \leq \delta.$$
Let $\mathcal{N}$ be the $\epsilon/3$-net obtained by applying this lemma to the space $AB$ with $\delta = \epsilon/3$. We have

$$\Pr\left\{ \exists |\psi\rangle \in \mathcal{N} : \frac{1}{t} \sum_{k=1}^{t} \Delta(p_{U_k|\psi\rangle}^{A}, \text{unif}([d_A])) \geq 2\epsilon/3 \right\}$$

$$\leq |\mathcal{N}| \cdot \exp\left( -\frac{\epsilon^2 t d}{18c} \right)$$

Now for an arbitrary state $|\psi\rangle \in AB$, we know that there exists $|\tilde{\psi}\rangle \in N$ such that $\| |\psi\rangle - |\tilde{\psi}\rangle \|_2 \leq \epsilon/3$. As a consequence, for any unitary transformation $U$,

$$\Delta(p_{U|\psi\rangle}^{A}, \text{unif}([d_A])) \leq \Delta(p_{U|\psi\rangle}^{A}, \text{unif}([d_A])) + \Delta\left(p_{U|\psi\rangle}^{A} , p_{U|\tilde{\psi}\rangle}^{A}\right)$$

$$\leq \Delta\left(p_{U|\tilde{\psi}\rangle}^{A} , \text{unif}([d_A])\right) + \| U|\tilde{\psi}\rangle - U|\psi\rangle \|_2$$

$$\leq \Delta\left(p_{U|\tilde{\psi}\rangle}^{A} , \text{unif}([d_A])\right) + \epsilon/3.$$ 

In the first inequality, we used the triangle inequality and the second inequality can be derived as in (3.12). Thus,

$$\Pr\left\{ \exists |\psi\rangle \in AB : \frac{1}{t} \sum_{k=1}^{t} \Delta(p_{U_k|\psi\rangle}^{A}, \text{unif}([d_A])) \geq \epsilon \right\} \leq \exp\left( -d\left(\frac{\epsilon^2 t}{18c} - 2\ln(9/\epsilon)\right) \right).$$

(3.15)

If $t > \frac{4\cdot18c\cdot\ln(9/\epsilon)}{\epsilon^2}$, this bound is strictly smaller than $1/2$ and the result follows.

To prove that we can suppose that $\{U_1, \ldots, U_t\}$ define $\gamma$-MUBs, consider the function $f : |\varphi\rangle \mapsto \langle \psi|\varphi\rangle$ for some fixed vector $|\psi\rangle$. Then, if $|\varphi\rangle$ is a random pure state, we have $\mathbb{E}\{f(|\varphi\rangle)\} = 0$. Moreover, using Levy’s Lemma with $\delta = d^{-\gamma/2}$

$$\Pr\left\{ |\langle \psi|\varphi\rangle| \geq d^{-\gamma/2} \right\} \leq 4 \exp\left( -\frac{d^{1-\gamma}}{c} \right).$$

Thus,

$$\Pr\left\{ \exists k \neq k', x, y \in [d], |\langle x|U_k^+ U_{k'}|y\rangle| \geq d^{-\gamma} \right\} \leq 4t^2 d^2 \exp\left( -\frac{d^{1-\gamma}}{c} \right)$$

(3.16)

which completes the proof.

\[ \square \]

**Corollary 3.3.7 (Existence of entropic uncertainty relations).** Let $C$ be a Hilbert space of dimension $d > 2$. There exists a constant $c' \geq 1$ such that for any integer $t > 2$ such that
3.3. Metric uncertainty relations: existence

\[ \frac{9 \cdot 16^2}{5 \cdot 18 \log t} \leq d, \text{ there exists a set } \{U_1, \ldots, U_t\} \text{ of unitary transformations of } C \text{ satisfying the following entropic uncertainty relation: for any state } |\psi\rangle, \]

\[ \frac{1}{t} \sum_{k=1}^{t} H(p_{U_k|\psi}) \geq \left( 1 - 8 \sqrt{\frac{c' \log t}{t}} \right) \log d \log \left( \frac{18t}{c' \log t} \right) - 2h_2 \left( 2 \sqrt{\frac{c' \log t}{t}} \right). \]

In particular, in the limit \( d \to \infty \), we obtain the existence of a sequence of sets of \( t \) bases satisfying

\[ \lim_{d \to \infty} \frac{1}{t} \sum_{k=1}^{t} H(p_{U_k|\psi}) \geq 1 - \sqrt{\frac{c' \log t}{t}}. \]

Remark. Recall that the bases (or measurements) that constitute the uncertainty relation are defined as the images of the computational basis by \( U_k^\dagger \). Note that for any set of unitaries \( \{U_1, \ldots, U_t\} \), we have

\[ \frac{1}{t} \sum_{k=1}^{t} H(p_{U_k|\psi}) \leq \left( 1 - \frac{1}{t} \right) \log d. \]

It is an open question whether there exists uncertainty relations matching this bound, even asymptotically as \( d \to \infty \) [Wehner and Winter, 2010]. Wehner and Winter [2010] ask whether there even exists a growing function \( f \) such that

\[ \lim_{d \to \infty} \frac{1}{t} \sum_{k=1}^{t} H(p_{U_k|\psi}) \geq 1 - \frac{1}{f(t)}. \]

The corollary answers this question in the affirmative with \( f(t) = \sqrt{\frac{t}{c' \log t}} \).

Proof Define \( c' = 5 \cdot 18c \) where \( c \) comes from Lévy’s Lemma 3.3.4 \[ \epsilon = \sqrt{\frac{c' \log t}{t}} \] and decompose \( C = A \otimes B \) with \( dB = [9/e^2] \). As \( d \geq \frac{9e^2}{\epsilon^2} \leq \frac{9e^2}{16^2} \) and

\[ \frac{4 \cdot 18c \log (9/e^2)}{\epsilon^2} = 4 \cdot 18c \log \left( \sqrt{\frac{t}{c' \log t}} \right) \cdot \frac{t}{5 \cdot 18c \log t} \leq t, \]

we get a family \( U_1, \ldots, U_t \) of unitary transformations that satisfies

\[ \frac{1}{t} \sum_{k=1}^{t} \Delta(p_{U_k|\psi}, \text{unif}([d_A])) \leq \epsilon. \]

By Proposition 3.2.2, these unitary transformations also satisfy an entropic uncertainty relation:

\[ \frac{1}{t} \sum_{k=1}^{t} H(p_{U_k|\psi}) \geq (1 - 8\epsilon) \log \left( \frac{d}{9/e^2} \right) - 2h_2(2\epsilon) \geq (1 - 8\epsilon) \log d - \log(18/e^2) - 2h_2(2\epsilon). \]

\[ \square \]
3.4 Metric uncertainty relations: explicit construction

In this section, we are interested in obtaining families \( \{U_1, \ldots, U_t\} \) of unitaries satisfying metric uncertainty relations where \( U_1, \ldots, U_t \) are explicit and efficiently computable using a quantum computer. For this section, we consider for simplicity a Hilbert space composed of qubits, i.e., of dimension \( d = 2^n \) for some integer \( n \). This Hilbert space is of the form \( A \otimes B \) where \( A \) describes the states of the first \( \log d_A \) qubits and \( B \) the last \( \log d_B \) qubits. Note that we assume that both \( d_A \) and \( d_B \) are powers of two.

We construct a set of unitaries by adapting an explicit low-distortion embedding of \((\mathbb{R}^d, \ell_2)\) into \((\mathbb{R}^{d'}, \ell_1)\) with \( d' = d^{1+o(1)} \) by Indyk [2007]. Indyk’s construction has two main ingredients: a set of mutually unbiased bases and an extractor. Our construction uses the same paradigm while requiring additional properties of both the mutually unbiased bases and the extractor.

In order to obtain a locking scheme that only needs simple quantum operations, we construct sets of approximately mutually unbiased bases from a restricted set of unitaries that can be implemented with single-qubit Hadamard gates. Moreover, we impose three additional properties on the extractor: we need our extractor to be strong, to define a permutation and to be efficiently invertible. We want the extractor to be strong because we are constructing metric uncertainty relations as opposed to a norm embedding. The property of being a permutation extractor is needed to ensure that the induced transformation on \((\mathbb{C}^2)^\otimes n\) preserves the \( \ell_2 \) norm. We also require the efficient invertibility condition to be able to build an efficient quantum circuit for the permutation. See Definition 3.4.4 for a precise formulation.

The intuition behind Indyk’s idea is as follows. Let \( V_1, \ldots, V_r \) be unitaries defining (approximately) mutually unbiased bases (see equation (3.17)) and let \( \{P_y\}_{y \in S} \) be a permutation extractor (Definition 3.4.4). The role of the mutually unbiased bases is to guarantee that for all states \( |\psi\rangle \) and for most values of \( j \in [r] \), most of the mass of the state \( V_j |\psi\rangle \) is “well spread” in the computational basis. This spread is measured in terms of the min-entropy of the distribution \( p_{V_j|\psi} \). Then, the extractor \( \{P_y\}_y \) will ensure that on average over \( y \in S \), the masses \( \sum_b |\langle a | b \rangle |^2 P_y V_j |\psi\rangle \) are almost equal for all \( a \in [d_A] \). More precisely, the distribution \( p^A_{P_y V_j |\psi} \) is close to uniform.

We start by recalling the definition of mutually unbiased bases. A set of unitary transformations \( V_1, \ldots, V_r \) is said to define \( \gamma \)-approximately mutually unbiased bases (or \( \gamma \)-MUBs) if for \( i \neq j \) and any elements \( |x\rangle \) and \( |y\rangle \) of the computational basis, we have

\[
|\langle x| V_i \dagger V_j |y\rangle | \leq \frac{1}{d^{\gamma/2}}. \tag{3.17}
\]
As shown in the following lemma, there is a construction of mutually unbiased bases that can be efficiently implemented [Wootters and Fields 1989].

**Lemma 3.4.1 (Quantum circuits for MUBs).** Let \( n \) be a positive integer and \( d = 2^n \). For any integer \( r \leq d+1 \), there exists a family \( V_1, \ldots, V_r \) of unitary transformations of \( \mathbb{C}^d \) that define mutually unbiased bases. Moreover, there is a randomized classical algorithm with runtime \( O(n^2 \text{ polylog } n) \) that takes as input \( j \in [r] \) and outputs a binary vector \( \alpha_j \in \{0, 1\}^{2n-1} \), and a quantum circuit of size \( O(n \text{ polylog } n) \) that when given as input the vector \( \alpha_j \) (classical input) and a quantum state \( |\psi\rangle \in \mathbb{C}^d \) outputs \( V_j |\psi\rangle \).

**Remark.** The randomization in the algorithm is used to find an irreducible polynomial of degree \( n \) over \( \mathbb{F}_2[X] \). It could be replaced by a deterministic algorithm that runs in time \( O(n^4 \text{ polylog } n) \). Observe that if \( n \) is odd and \( r \leq (d+1)/2 \), it is possible to choose the unitary transformations to be real (see Heath et al. [2006]).

**Proof** We define \( V_1 = \text{id} \), and the remaining unitaries are indexed by binary vectors \( u \in \{0, 1\}^n \), for example the binary representations of integers from 0 to \( r - 2 \). The construction is based on operations in the finite field \( \mathbb{F}_{2^n} \). The field \( \mathbb{F}_{2^n} \) can be seen as an \( n \)-dimensional vector space over \( \mathbb{F}_2 \). Choose \( \theta \in \mathbb{F}_{2^n} \) such that \( 1, \theta, \ldots, \theta^{n-1} \) form a basis of \( \mathbb{F}_{2^n} \). For any \( x, y \in [n] \), \( \theta^x \cdot \theta^y \in \mathbb{F}_{2^n} \) can be decomposed in our chosen basis as \( \theta^x \cdot \theta^y = \sum_{\ell=0}^{n-1} m_{\ell}(x, y) \theta^\ell \) for some \( m_{\ell}(x, y) \in \mathbb{F}_2 \). We can thus define the matrices \( M_0, M_1, \ldots, M_{n-1} \) from the multiplication table

\[
\begin{pmatrix}
1 \\
\theta \\
\vdots \\
\theta^{n-1}
\end{pmatrix}
\begin{pmatrix}
1 & \theta & \ldots & \theta^{n-1}
\end{pmatrix}
= M_0 + M_1 \theta + \cdots + M_{n-1} \theta^{n-1}.
\]

where \( M_\ell = (m_{\ell}(x, y))_{x,y \in [n]} \). For a given \( u \in \{0, 1\}^n \), we define the matrix

\[
N_u = \sum_{\ell=0}^{n-1} u_\ell M_\ell.
\]

Notice that as \( \theta^x \cdot \theta^y = \theta^{x+y} \), the entry \( N_u(x, y) \) of \( N_u \) only depends on \( x + y \), i.e., \( N_u(x, y) = N_u(x', y') \) if \( x + y = x' + y' \). So we can represent this matrix by a vector \( \alpha_u(x + y) = N_u(x, y) \) of length \( 2n - 1 \). We then define a \( \mathbb{Z}_4 \)-valued quadratic form by: for \( v \in \{0, 1\}^n \),

\[
T_u(v) = v^T N_u v \mod 4.
\]
Note that the operations $v^T N_u v$ are not performed in $\mathbb{F}_2$ but rather in $\mathbb{Z}$. Using the vector $\alpha_u$, we can write

$$T_u(v) = \sum_{x,y \in [n]} v_x N_u(x,y) v_y \mod 4 = \sum_{z=0}^{2n-2} \left( \sum_{x=0}^{z} v_x v_{z-x} \right) \alpha_u(z) \mod 4$$

if we define $v_x = 0$ for $x \geq n$. We then define the diagonal matrix $D_u = \text{diag} \left( i^{T_u(v)} \right)_{v \in \mathbb{F}_2^n}$.

Finally, we define for $2 \leq j \leq r$,

$$V_j = D_{\text{bin}(j-2)} H^\otimes n$$

where $\text{bin}(j) \in \{0,1\}^n$ is the binary representation of length $n$ of the integer $j$.

The fact that these unitaries define mutually unbiased bases was proved in Wootters and Fields [1989]. We now analyse how fast these unitary transformations can be implemented.

Given the index $j$ as input, we show it is possible to compute $u = \text{bin}(j-2)$ and compute the vector $\alpha_j \overset{\text{def}}{=} \alpha_u$ in time $O(n^2 \text{ polylog } n)$. In fact, we start by computing a representation of the field $\mathbb{F}_{2^n}$ by finding an irreducible polynomial $Q$ of degree $n$ in $\mathbb{F}_2[X]$, so that $\mathbb{F}_{2^n} = \mathbb{F}_2[X]/Q$. This can be done in expected time $O(n^2 \text{ polylog } n)$ (Corollary 14.43 in the book von zur Gathen and Gerhard [1999]). There also exists a deterministic algorithm for finding an irreducible polynomial in time $O(n^4 \text{ polylog } n)$ [Shoup, 1990]. We then take $\theta = X$. Computing the polynomial $X^x \cdot X^y = X^{x+y} \mod Q$ can be done in time $O(n \text{ polylog } n)$ using the fast Euclidean algorithm (see Corollary 11.8 in von zur Gathen and Gerhard [1999]). As $x+y \in [0, 2n-2]$, we can explicitly represent all the polynomials $X^z$ for $0 \leq z \leq 2n-2$ in time $O(n^2 \text{ polylog } n)$. It is then simple to compute the vector $\alpha_u$ using the vector $u$ in time $O(n^2)$.

To build the quantum circuit, we first observe that applying a Hadamard transform only takes $n$ single-qubit Hadamard gates. Then, to design a circuit performing the unitary transformation $D_{\text{bin}(j-2)}$, we start by building a classical circuit that computes

$$T_u(v) = \sum_{z=0}^{2n-2} \left( \sum_{x=0}^{z} v_x v_{z-x} \right) \alpha_u(z) \mod 4$$

on inputs $v$ and $\alpha_u$. Observing that $\sum_{x=0}^{z} v_x v_{z-x}$ is the coefficient of $Y^z$ in the polynomial $(\sum_{x=0}^{n-1} v_x Y_x)^2$, we can use fast polynomial multiplication to compute $T_u(v)$ in time $O(n \text{ polylog } n)$ (Corollary 8.27 in von zur Gathen and Gerhard [1999]). This circuit can be transformed into a reversible circuit with the same size (up to some multiplicative constant).
that takes as input \((v, \alpha_j, g)\) where \(v \in \{0, 1\}^n\), \(\alpha_j \in \{0, 1\}^{2n-1}\) and \(g \in \mathbb{Z}_4\), and outputs \((v, \alpha_j, g + T_u(v) \mod 4)\).

This reversible classical circuit can be readily transformed into a quantum circuit that computes the unitary transformation defined by
\[
W : |v\rangle|g\rangle \mapsto |v\rangle|g + T_u(v)\mod 4\rangle.
\]
Recall that we want to implement the transformation \(D_u: |v\rangle \mapsto i^{T_u(v)}|v\rangle\) efficiently. This is simple to obtain using the quantum circuit for \(W\). In fact, if we use a catalyst state \(|\phi\rangle = |0\rangle - i|1\rangle - |2\rangle + i|3\rangle\), we have
\[
W|v\rangle|\phi\rangle = i^{T_u(v)}|v\rangle|\phi\rangle = D_{\text{bin}(j-2)}|v\rangle|\phi\rangle.
\]
Finally, \(D_{\text{bin}(j-2)}H^\otimes n\) can be implemented by a quantum circuit of size \(O(n \polylog n)\).

It is also possible to obtain approximately mutually unbiased bases that use smaller circuits. In fact, the following lemma shows that we can construct large sets of approximately mutually unbiased bases defined by unitaries in the restricted set
\[
\mathcal{H} = \{H^v \overset{\text{def}}{=} H^{v_1} \otimes \cdots \otimes H^{v_n}, v \in \{0, 1\}^n\},
\]
where \(H\) is the Hadamard transform on \(\mathbb{C}^2\) defined by
\[
H = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}.
\]
In our construction of metric uncertainty relations (Theorem 3.4.6), we could use the 1-MUBs of Lemma 3.4.1 or the \((1/2 - \delta)\)-MUBs of Lemma 3.4.2. As the construction of approximate MUBs is simpler and can be implemented with simpler circuits, we will mostly be using Lemma 3.4.2.

**Lemma 3.4.2 (Approximate MUBs in \(\mathcal{H}\)).** Let \(n'\) be a positive integer and \(n = 2^{n'}\).

1. For any integer \(r \leq n\), there exists a family \(V_1, \ldots, V_r \in \mathcal{H}\) that define 1/2-MUBs.

2. For any \(\delta \in (0, 1/2)\), there exists a constant \(c > 0\) independent of \(n\) such that for any \(r \leq 2^n\) there exists a family \(V_1, \ldots, V_r\) of unitary transformations in \(\mathcal{H}\) that define \((1/2 - \delta)\)-MUBs.

Moreover, in both cases, given an index \(j \in [r]\), there is a polynomial time (classical) algorithm that computes the vector \(v \in \{0, 1\}^n\) that defines the unitary \(V_j = H^v\).

**Proof** Observe that for any \(v \in \{0, 1\}^n\) and any \(y \in \{0, 1\}^n\), we have
\[
H^v(|y_1\rangle \otimes \cdots \otimes |y_n\rangle) = H^{v_1} |y_1\rangle \otimes \cdots \otimes H^{v_n} |y_n\rangle = \sum_{y' \in \{0, 1\}^n} \frac{(-1)^{v \cdot y'}}{\sqrt{2}^{w(v)}} |y'_1 \cdots y'_n\rangle.
\]
where \( w(v) \) is the number of non-zero components of \( v \). Thus,

\[
|\langle x | H^v H^{v'} | y \rangle| = |\langle x | H^{v+v'} | y \rangle| \leq \frac{1}{2d_H(v,v')/2},
\]

(3.18)

where \( d_H(v,v') = \{ i : v_i \neq v_i' \} \) is the Hamming distance between the two vectors \( v \) and \( v' \). Using this observation, we see that a binary code \( C \subseteq \{0,1\}^n \) with minimum distance \( \gamma \) defines a set of \( \gamma \)-MUBs in \( \mathcal{H} \). It is now sufficient to find binary codes with minimum distance as large as possible. For the first construction, we use the Hadamard code that has minimum distance \( n/2 \). The Hadamard codewords are indexed by \( x \in \{0,1\}^n \); the codeword corresponding to \( x \) is the vector \( v \in \{0,1\}^n \) whose coordinates are \( v_z = x \cdot z \) for all \( z \in \{0,1\}^n \). This code has the largest possible minimum distance for a non-trivial binary code but its shortcoming is that the number of codewords is only \( n \). For our applications, it is sometimes desirable to have \( r \) larger than \( n \) (this is useful to allow the error parameter \( \epsilon \) of our metric uncertainty relation to be smaller than \( n^{-1/2} \)).

For the second construction, we use families of linear codes with minimum distance \((1/2 - \delta)n\) with a number of codewords that is exponential in \( n \). For this, we can use Reed-Solomon codes concatenated with linear codes on \( \{0,1\}^{\Theta(n')} \) that match the performance of random linear codes; see for example Appendix E in Goldreich [2008]. For a simpler construction, note that we can also get \( 2^{\Omega(\sqrt{n})} \) codewords by using a Reed-Solomon code concatenated with a Hadamard code.

The next lemma shows that for any state \( |\psi\rangle \), for most values of \( j \), the distribution \( p_{V_j|\psi} \) is close to a distribution with large min-entropy provided \( \{V_j\} \) define \( \gamma \)-MUBs. This result might be of independent interest. In fact, Damgård et al. [2007] prove a lower bound close to \( n/2 \) on the min-entropy of a measurement in the computational basis of the state \( U|\psi\rangle \) where \( U \) is chosen uniformly from the full set of the \( 2^n \) unitaries of \( \mathcal{H} \). They leave as an open question the existence of small subsets of \( \mathcal{H} \) that satisfy the same uncertainty relation. When used with the \( \gamma \)-MUBs of Lemma 3.4.2, the following lemma partially answers this question by exhibiting such sets of size polynomial in \( n \) but with a min-entropy lower bound close to \( n/4 \) instead. This can be used to reduce the amount of randomness needed for many protocols in the bounded and noisy quantum storage models.

**Lemma 3.4.3.** Let \( n \geq 1, d = 2^n \) and \( \epsilon \in (0,1) \) and consider a set of \( r = \lceil \frac{2}{\epsilon^2} \rceil \) unitary transformations \( V_1, \ldots, V_r \) of \( \mathbb{C}^d \) defining \( \gamma \)-MUBs. For all \( |\psi\rangle \in \mathbb{C}^d \),

\[
\left\{ j \in [r] : \exists q_j, \Delta(p_{V_j|\psi}), q_j) \leq \epsilon \text{ and } H_{\text{min}}(q_j) \geq \frac{\gamma n}{2} - \log(8/\epsilon^2) \right\} \geq (1 - \epsilon)r.
\]

**Proof** This proof proceeds along the lines of [Indyk 2007, Lemma 4.2]. Similar results can also be found in the sparse approximation literature; see [Tropp 2004, Proposition 4.3] and references therein.
Consider the \( rd \times d \) matrix \( V \) obtained by concatenating the rows of the matrices \( V_1, \ldots, V_r \). For \( S \subseteq [rd] \), \( V_S \) denotes the submatrix of \( V \) obtained by selecting the rows in \( S \). The coordinates of the vector \( V|\psi\rangle \in \mathbb{C}^r \) are indexed by \( z \in [rd] \) and denoted by \((V|\psi\rangle)_z\).

**Claim.** We have for any set \( S \subseteq [rd] \) of size at most \( d^{\gamma/2} \) and any unit vector \(|\psi\rangle\),
\[
\| (V|\psi\rangle)_S \|_2^2 \leq 1 + \frac{|S|}{d^{\gamma/2}}. \tag{3.19}
\]

To prove the claim, we want an upper bound on the operator 2-norm of the matrix \((V_S)\), which is the square root of the largest eigenvalue of \( G = V_S^\dagger V_S \). As two distinct rows of \( V \) have an inner product bounded by \( \frac{1}{d^{\gamma/2}} \), the non-diagonal entries of \( G \) are bounded by \( \frac{1}{d^{\gamma/2}} \). Moreover, the diagonal entries of \( G \) are all 1. By the Gershgorin circle theorem, all the eigenvalues of \( G \) lie in the disc centered at 1 of radius \( \frac{|S|-1}{d^{\gamma/2}} \). We conclude that (3.19) holds.

Now pick \( S \) to be the set of indices of the \( d^{\gamma/2} \) largest entries of the vector \( \{(V|\psi\rangle)_z \} \). Using the previous claim, we have \( \| (V|\psi\rangle)_S \|_2^2 \leq 2 \). Moreover, since \( S \) contains the \( d^{\gamma/2} \) largest entries of \( \{(V|\psi\rangle)_z \} \), we have that for all \( z \not\in S \),
\[
| (V|\psi\rangle)_z |^2 d^{\gamma/2} \leq \| V|\psi\rangle \|_2^2 = \sum_{j=1}^{r} \| V_j|\psi\rangle \|_2^2 = r. \]
Thus, for all \( z \not\in S \), \( | (V|\psi\rangle)_z |^2 \leq \frac{r}{d^{\gamma/2}} \).

We now build the distributions \( q_j \). For every \( j \in [r] \), define
\[
w_j = \sum_{z \in S \cap \{ (j-1)d+1, \ldots, jd \} } | (V|\psi\rangle)_z |^2,
\]
which is the total weight in \( S \) of \( V_j|\psi\rangle \). Defining \( T_\epsilon = \{ j : w_j > \epsilon \} \), we have \( |T_\epsilon| \epsilon \leq \| (V|\psi\rangle)_S \|_2^2 \leq 2 \). Thus,
\[
|T_\epsilon| \leq 2/\epsilon \leq 2r.
\]

We define the distribution \( q_j \) for \( j \in [r] \) by
\[
q_j(x) = \begin{cases} 
\frac{w_j}{d} & \text{if } (j-1)d + x \not\in S \\
\frac{w_j}{d} & \text{if } (j-1)d + x \in S.
\end{cases}
\]

Since
\[
\sum_x q_j(x) = w_j + \sum_{x \in [d]: (j-1)d + x \not\in S} |\langle x|V_j|\psi\rangle|^2 = \sum_{x \in [d]} |\langle x|V_j|\psi\rangle|^2 = 1,
\]
\( q_j \) is a probability distribution. Moreover, we have that for \( j \not\in T_\epsilon \)
\[
\Delta (p|v_j\rangle, q_j) \leq \frac{1}{2} \left( \sum_{x:(j-1)d+x \not\in S} \frac{w_j}{d} + \sum_{x:(j-1)d+x \in S} \left( \frac{w_j}{d} + |\langle x|V_j|\psi\rangle|^2 \right) \right) = w_j \leq \epsilon.
\]
The distribution $q_j$ also has the property that for all $x \in [d]$, $q_j(x) \leq \frac{x}{d^{\ell/2}} + \frac{1}{d} \leq \frac{2x}{d^{\ell/2}}$. In other words, $H_{\min}(q_j) \geq \frac{2n}{2} - \log(8/\epsilon^2)$. □

We now move to the second building block in Indyk’s construction: randomness extractors. Randomness extractors are functions that extract uniform random bits from weak sources of randomness.

**Definition 3.4.4** (Strong permutation extractor). Let $n$ and $m \leq n$ be positive integers, $\ell \in [0,n]$ and $\epsilon \in (0,1)$. A family of permutations $\{P_y\}_{y \in S}$ of $\{0,1\}^n$ where each permutation $P_y$ is described by two functions $P_E^y : \{0,1\}^n \rightarrow \{0,1\}^m$ (the first $m$ output bits of $P_y$) and $P_R^y : \{0,1\}^n \rightarrow \{0,1\}^{n-m}$ (the last $n-m$ output bits of $P_y$) is said to be an explicit $(n,\ell) \rightarrow \epsilon$ strong permutation extractor if:

- For any random variable $X$ on $\{0,1\}^n$ such that $H_{\min}(X) \geq \ell$, and an independent seed $U_S$ uniformly distributed over $S$, we have
  
  $$\Delta\left(p_{\{U_S,P_E^y(X)\}}, unif(S \times \{0,1\}^m)\right) \leq \epsilon,$$

  which is equivalent to

  $$\frac{1}{|S|} \sum_{y \in S} \Delta\left(p_{P_E^y(X)}, unif(\{0,1\}^m)\right) \leq \epsilon. \quad (3.20)$$

- For all $y \in S$, both the function $P_y$ and its inverse $P_y^{-1}$ are computable in time polynomial in $n$.

**Remark.** A similar definition of permutation extractors was used in Reingold et al. [2000] in order to avoid some entropy loss in an extractor construction. Here, the reason we use permutation extractors is different; it is because we want the induced transformation $P_y$ on $\mathbb{C}^{2n}$ to preserve the $\ell_2$ norm.

We can adapt an extractor construction of Guruswami et al. [2009] to obtain a permutation extractor with the following parameters. The details of the construction are presented in Appendix A.2.

**Theorem 3.4.5** (Explicit strong permutation extractors). For all (constant) $\delta \in (0,1)$, all positive integers $n$, all $\ell \in [c \log(n/\epsilon), n]$ ($c$ is a constant independent of $n$ and $\epsilon$), and all $\epsilon \in (0, 1/2)$, there is an explicit $(n, \ell) \rightarrow (1-\delta)\ell$ strong permutation extractor $\{P_y\}_{y \in S}$ with $\log |S| \leq O(\log(n/\epsilon))$. Moreover, the functions $(x,y) \mapsto P_y(x)$ and $(x,y) \mapsto P_y^{-1}(x)$ can be computed by circuits of size $O(n \text{ polylog}(n/\epsilon))$. 

A permutation $P$ on $\{0, 1\}^n$ defines a unitary transformation on $(\mathbb{C}^2)^\otimes n$ that we also call $P$. The permutation extractor $\{P_y\}$ will be seen as a family of unitary transformations over $n$ qubits. Moreover, just as we decomposed the space $\{0, 1\}^n$ into the first $m$ bits and the last $n - m$ bits, we decompose the space $(\mathbb{C}^2)^\otimes n$ into $A \otimes B$, where $A$ represents the first $m$ qubits and $B$ represents the last $n - m$ qubits. The properties of $\{P^E_y\}$ will then be reflected in the system $A$.

Combining Theorem 3.4.5 and Lemma 3.4.3, we obtain a set of unitaries satisfying a metric uncertainty relation.

**Theorem 3.4.6 (Explicit uncertainty relations: key optimized).** Let $\delta > 0$ be a constant, $n$ be a positive integer, $\epsilon \in (2^{-c'n}, 1)$ ($c'$ is a constant independent of $n$). Then, there exist $t \leq \left(\frac{2}{\epsilon}^n\right)^c$ (for some constant $c$ independent of $n$ and $\epsilon$) unitary transformations $U_1, \ldots, U_t$ acting on $n$ qubits such that: if $A$ represents the first $(1 - \delta)n/4 - O(\log(1/\epsilon))$ qubits and $B$ represents the remaining qubits, then for all $|\psi\rangle \in AB$,

$$\frac{1}{t} \sum_{k=1}^{t} \Delta(p_{U_k|\psi\rangle}^A, \text{unif}([d_A])) \leq \epsilon.$$

Moreover, the mapping that takes the index $k \in [t]$ and a state $|\psi\rangle$ as inputs and outputs the state $U_k|\psi\rangle$ can be performed by a classical computation with polynomial runtime and a quantum circuit that consists of single-qubit Hadamard gates on a subset of the qubits followed by a permutation in the computational basis. This permutation can be computed by (classical or quantum) circuits of size $O(n \text{ polylog}(n/\epsilon))$.

**Remark.** Observe that in terms of the dimension $d$ of the Hilbert space, the number of unitaries $t$ is polylogarithmic.

**Proof** Let $\epsilon' = \epsilon/6$. Lemma 3.4.2 gives $r = \lceil 2/\epsilon^2 \rceil$ unitary transformations $V_1, \ldots, V_r$ that define $\gamma$-mutually unbiased bases with $\gamma = 1/2 - \delta/4$. Moreover, all these unitaries can be performed by a quantum circuit that consists of single-qubit Hadamard gates on a subset of the qubits. Theorem 3.4.5 with $\ell = (1 - \delta/2)n/4 - \log(8/\epsilon^2)$ and error $\epsilon'$ gives $|S| \leq 2^{c \log(n/\epsilon')}$ permutations $\{P_y\}_{y \in S}$ of $\{0, 1\}^n$ that define an $(n, \ell) \mapsto_{\epsilon'} (1 - \delta/2)\ell$ extractor and are computable by classical circuits of size $O(n \text{ polylog}(n/\epsilon))$. We now argue that this classical circuit can be used to build a quantum circuit of size $O(n \text{ polylog}(n/\epsilon))$ that computes the unitaries $P_y$.

Given classical circuits that compute $P$ and $P^{-1}$, we can construct reversible circuits $C_P$ and $C_{P^{-1}}$ for $P$ and $P^{-1}$. The circuit $C_P$ when given input $(x, 0)$ outputs the binary string $(x, P(x))$, so that it keeps the input $x$. Such a circuit can readily be transformed into
a quantum circuit that acts on the computational basis states as the classical circuit. We also call these circuits $C_P$ and $C_{P^{-1}}$. Observe that we want to compute the unitary $P$, so we have to erase the input $x$. For this, we combine the circuits $C_P$ and $C_{P^{-1}}$ as described in Figure 3.1. Note that the size of this quantum circuit is the same as the size of the original classical circuit up to some multiplicative constant. Thus, this quantum circuit has size $O(n \log(8/\epsilon^2))$.

![Figure 3.1: Quantum circuit to compute the permutation $P$ using quantum circuits $C_P$ for $P$ and $C_{P^{-1}}$ for $P^{-1}$. ($C_{P^{-1}})^{-1}$ is simply the circuit $C_{P^{-1}}$ taken backwards. The bottom register is an ancilla register.](image)

The unitaries $\{U_1, \ldots, U_t\}$ are obtained by taking all the possible products $P_y V_j$ for $j \in [r], y \in S$. Note that $t = r|S|$. We now show that the set $\{U_1, \ldots, U_t\}$ satisfies the uncertainty relation property. Using Lemma 3.4.3 for any state $|\psi\rangle$, the set

$$T_{|\psi\rangle} \triangleq \{ j : \exists q_j, \Delta(p_{V_j|\psi}; q_j) \leq \epsilon' \text{ and } H_{\min}(q_j) \geq (1 - \delta/2)n/4 - \log(8/\epsilon^2) \}$$

has size at least $(1 - \epsilon')r$. Moreover, for all $a \in [d_A]$, $p_{P_y V_i|\psi}^A(a) = \sum_b |\langle a | \langle b | P_y V_i | \psi \rangle|^2 = \Pr \{ P_y^E(X) = a \}$ where $X$ has distribution $p_{V_i|\psi}$. By definition, for $i \in T_{|\psi\rangle}$, we have $\Delta(p_{V_i|\psi}, q_i) \leq \epsilon'$ with $H_{\min}(q_i) \geq (1 - \delta/2)n/4 - \log(8/\epsilon^2)$. Using the fact that $\{P_y^E\}$ is a strong extractor (see [3.20]) for min-entropy $(1 - \delta/2)n/4 - \log(8/\epsilon^2)$, it follows that

$$\frac{1}{|S|} \sum_{y \in S} \Delta(p_{P_y V_i|\psi}, \text{unif}([d_A])) \leq 2\epsilon'$$

for all $i \in T_{|\psi\rangle}$. As $|T_{|\psi\rangle}| \geq (1 - \epsilon')r$, we obtain

$$\frac{1}{t} \sum_{k=1}^{t} \Delta(p_{U_k|\psi}, \text{unif}([d_A])) \leq 3\epsilon' = \epsilon/2.$$ 

To conclude, we show that $t$ can be taken to be a power of two at the cost of multiplying the error by at most two. In fact, let $p$ be the smallest integer satisfying $t \leq 2^p$, so that $2^p \leq 2t$. 


By repeating $2^p - t$ unitaries, it is easily seen that we obtain an $\epsilon$-metric uncertainty relation with $2^p$ unitaries from an $\epsilon/2$-metric uncertainty relation with $t$ unitaries. □

Note that the $B$ system we obtain is quite large and to get strong uncertainty relations, we want the system $B$ to be as small as possible. For this, it is possible to repeat the construction of the previous theorem on the $B$ system. The next theorem gives a construction where the $A$ system is composed of $n - O(\log \log n) - O(\log(1/\epsilon))$ qubits. Of course, this is at the expense of increasing the number of unitaries in the uncertainty relation.

**Theorem 3.4.7** (Explicit uncertainty relation: message length optimized). Let $n$ be a positive integer and $\epsilon \in (2^{-c'n}, 1)$ where $c'$ is a constant independent of $n$. Then, there exist $t \leq (\frac{n}{2})^c \log n$ (for some constant $c$ independent of $n$ and $\epsilon$) unitary transformations $U_1, \ldots, U_t$ acting on $n$ qubits that are all computable by quantum circuits of size $O(n \text{ polylog}(n/\epsilon))$ such that: if $A$ represents the first $n - O(\log \log n) - O(\log(1/\epsilon))$ qubits and $B$ represents the remaining qubits, then for all $|\psi\rangle \in AB$,

$$\frac{1}{t} \sum_{k=1}^{t} \Delta(p_{U_k|\psi\rangle}^{A}, \text{unif}([d_A])) \leq \epsilon.$$  

(3.21)

Moreover, the mapping that takes the index $k \in [t]$ and a state $|\psi\rangle$ as inputs and outputs the state $U_k|\psi\rangle$ can be performed by a classical precomputation with polynomial runtime and a quantum circuit of size $O(n \text{ polylog}(n/\epsilon))$. The number of unitaries $t$ can be taken to be a power of two.

**Proof** Using the construction of Theorem 3.4.6, we obtain a system $A$ over which we have some uncertainty relation and a system $B$ that we do not control. In order to decrease the dimension of the system $B$, we can apply the same construction to that system. The system $B$ then gets decomposed into $A_2B_2$, and we know that the distribution of the measurement outcomes of system $A_2$ in the computational basis is close to uniform. As a result, we obtain an uncertainty relation on the system $AA_2$ (see Figure 3.2).

More precisely, we start by demonstrating a simple property about the composition of metric uncertainty relations. Note that this composition is different from the one described in (3.8), but the proof is quite similar.

**Claim.** Suppose the set $\{U_1^{(1)}, \ldots, U_t^{(1)}\}$ of unitaries on $A_1B_1$ satisfies a $(t_1, \epsilon_1)$-metric uncertainty relation on system $A_1$ and the $\{U_1^{(2)}, \ldots, U_t^{(2)}\}$ of unitaries on $B_1 = A_2B_2$ satisfies a $(t_2, \epsilon_2)$-metric uncertainty relation on $A_2$. Then the set of unitaries
In order to reduce the dimension of the $B$ system, we can re-apply the uncertainty relation to the $B$ system.

$$\left\{ (\text{id}^{A_1} \otimes U^{(2)}_{k_2}) \cdot U^{(1)}_{k_1} \right\}_{k_1,k_2 \in [t_1] \times [t_2]}$$

satisfies a $(t_1 t_2, \epsilon_1 + \epsilon_2)$-metric uncertainty relation on $A_1 A_2$: for all $|\psi\rangle \in A_1 A_2 B_2$,

$$\frac{1}{t_1 t_2} \sum_{k_1,k_2 \in [t_1] \times [t_2]} \Delta\left( p_{U^{(2)}_{k_2} U^{(1)}_{k_1} |\psi\rangle}, \text{unif}([d_{A_1} d_{A_2}]) \right) \leq \epsilon_1 + \epsilon_2.$$

For a fixed value of $k_1 \in [t_1]$ and $a_1 \in [d_{A_1}]$, we can apply the second uncertainty relation to the state $\langle a_1|^{A_1} U_{k_1} |\psi\rangle_{||} = \frac{1}{\sqrt{p_{U_{k_1} |\psi\rangle}(a_1)}} \sum_{b_2} \langle a_1| (b_1|U_{k_1} |\psi\rangle) |b_1\rangle \in B_1 = A_2 B_2$.

As $\{|b_1\rangle\}_{b_1} = \{|a_2\rangle |b_2\rangle\}_{a_2,b_2}$, we have

$$\frac{1}{t_2} \sum_{k_2} \sum_{a_2} \left| \frac{1}{p_{U_{k_1} |\psi\rangle}(a_1)} \sum_{b_2} \langle a_1| A_1 A_2 B_2 (\text{id}^{A_1} \otimes U_{k_2}) U_{k_1} |\psi\rangle |b_2\rangle - \frac{1}{d_{A_2}} \right| \leq \epsilon_2.$$

We can then calculate, in the same vein as (3.10)

$$\frac{1}{t_1 t_2} \sum_{k_1,k_2,a_1,a_2} \sum_{b_2} \left| \langle a_1| A_1 A_2 B_2 (\text{id}^{A_1} \otimes U_{k_2}) U_{k_1} |\psi\rangle |b_2\rangle - \frac{1}{d_{A_1} d_{A_2}} \right|^2 \leq \frac{1}{t_1 t_2} \sum_{k_1,k_2,a_1,a_2} \sum_{b_2} \left| \langle a_1| A_1 A_2 B_2 (\text{id}^{A_1} \otimes U_{k_2}) U_{k_1} |\psi\rangle |b_2\rangle - \frac{1}{d_{A_1} d_{A_2}} \right|^2 \leq \frac{1}{t_1} \sum_{k_1} \sum_{a_1,a_2} \left| \frac{p_{U_{k_1} |\psi\rangle}(a_1)}{d_{A_2}} - \frac{1}{d_{A_1} d_{A_2}} \right| \leq \epsilon_2 + \epsilon_1.$$

This completes the proof of the claim.

To obtain the claimed dimensions, we compose the construction of Theorem 3.4.6 $h$ times with an error parameter $\epsilon' = \epsilon/h$ and $\delta = 1/8$. Starting with a space of $n$ qubits, the
dimension of the $B$ system (after one step) can be bounded by
\[
\frac{7}{8} n - O(\log(1/\epsilon')) \leq \log d_B \leq \frac{7}{8} n.
\]
So after $h$ steps, we have
\[
(\frac{7}{8})^h n - O(\log(1/\epsilon')) \cdot 8(1 - (\frac{7}{8})^h) \leq \log d_{B_h} \leq (\frac{7}{8})^h n.
\]
Thus,
\[
(\frac{7}{8})^h n - O(\log(1/\epsilon')) \leq \log d_{B_h} \leq (\frac{7}{8})^h n.
\]
Note that $h$ cannot be arbitrarily large: in order to apply the construction of Theorem 3.4.6 on a system of $m$ qubits with error $\epsilon'$, we should have $\epsilon' \geq 2^{-c'm}$. In other words, if
\[
\log d_{B_h} \geq \frac{1}{c'} \log(h/\epsilon), \tag{3.22}
\]
then we can apply the construction $h$ times. Let $c''$ be a constant to be chosen later and
\[
h = \left\lceil \frac{1}{\log(8/7)} \left( \log n - \log(c'' \log \log n + c'' \log(1/\epsilon)) \right) \right\rceil.
\]
This choice of $h$ satisfies (3.22). In fact,
\[
\log d_{B_h} \geq c'' \log \log n + c'' \log(1/\epsilon) - O(\log(h/\epsilon)) \\
\geq \frac{1}{c'} \log(h/\epsilon)
\]
if $c''$ is chosen large enough. Moreover, we get
\[
\log d_{B_h} = 2^{-\log n} \cdot 2^{\log O(\log \log n + \log(1/\epsilon))} \cdot n = O(\log \log n + \log(1/\epsilon))
\]
as stated in the theorem.

Each unitary of the obtained uncertainty relation is a product of $h$ unitaries each obtained from Theorem 3.4.6. The overall number of unitaries is the product of the number of unitaries for each of the $h$ steps. As a result, we have $t \leq \left( \frac{n}{\epsilon} \right)^{c \log n}$ for some constant $c$. $t$ can be taken to be a power of two as the number of unitaries at each step can be taken to be a power of two. As for the running time, every unitary transformation of the uncertainty relation is a product of $O(\log n)$ unitaries each computed by a quantum circuit of size $O(n \text{polylog}(n/\epsilon))$ and can thus be computed by a quantum circuit of size $O(n \text{polylog}(n/\epsilon))$. \hfill \square

It is of course possible to obtain a trade-off between the key size and the dimension of the $B$ system by choosing the number of times the construction of Theorem 3.4.6 is applied. In the next corollary, we show how to obtain an explicit entropic uncertainty relation whose average entropy is $(1 - \epsilon)n$. 

Corollary 3.4.8 (Explicit entropic uncertainty relations). Let $n \geq 100$ be an integer, and $\epsilon \in (10n^{-1/2}, 1)$. Then, there exists $t \leq \left(\frac{n}{\epsilon}\right)^{c\log(1/\epsilon)}$ (for some constant $c$ independent of $n$ and $\epsilon$) unitary transformations $U_1, \ldots, U_t$ acting on $n$ qubits that are all computable by quantum circuits of size $O(n \text{ polylog } n)$ satisfying an entropic uncertainty relation: for all pure states $|\psi\rangle \in (\mathbb{C}^2)^\otimes n$,

$$\frac{1}{t} \sum_{k=1}^{t} H(p_{U_k|\psi\rangle}) \geq (1 - 9\epsilon)n - 2h_2(2\epsilon).$$

Moreover, the mapping that takes the index $k \in [t]$ and a state $|\psi\rangle$ as inputs and outputs the state $U_k|\psi\rangle$ can be performed by a classical precomputation with polynomial runtime and a quantum circuit of size $O(n \text{ polylog } n)$. The number of unitaries $t$ can be taken to be a power of two.

Proof The proof is basically the same as the proof of Theorem 3.4.7 except that we repeat the construction $h = \lceil \log(1/\epsilon) / \log(8/7) \rceil$ times. We thus have

$$\log d_{B_h} \leq (7/8)^h n \leq \epsilon n.$$ 

We obtain a set of $t \leq \left(\frac{n}{\epsilon}\right)^{c\log(1/\epsilon)}$ unitary transformations. Applying Proposition 3.2.2 we get

$$\frac{1}{t} \sum_{i=1}^{t} H(p_{U_k|\psi\rangle}) \geq (1 - 8\epsilon)(1 - \epsilon)n - 2h_2(2\epsilon)$$

$$\geq (1 - 9\epsilon)n - 2h_2(2\epsilon).$$

$\Box$
Chapter 4

Uncertainty relations for quantum measurements: Applications

Outline of the chapter In this chapter, we give several applications of uncertainty relations. We start in Section 4.1 with applications related to information locking which all have a cryptographic flavour. In Section 4.2 we consider the communication problem called quantum identification.

4.1 Locking classical information in quantum states

Outline of the section We apply the results on metric uncertainty relations of the previous chapter to obtain locking schemes. After an introductory section on locking classical correlations (Section 4.1.1), we show how to obtain a locking scheme using a metric uncertainty relation in Section 4.1.2. Using the constructions of the previous chapter, this leads to locking schemes presented in Corollaries 4.1.5 and 4.1.7. Section 4.1.4 discusses the existence of error tolerant locking schemes. In Section 4.1.5 we show how to construct quantum hiding fingerprints by locking a classical fingerprint. In Section 4.1.6 we observe that these locking schemes can be used to construct efficient string commitment protocols. Section 4.1.7 discusses the link to locking entanglement of formation.

4.1.1 Background

Locking of classical correlations was first described by DiVincenzo et al. [2004] as a violation of the incremental proportionality of the maximal classical mutual information that can be obtained by local measurements on a bipartite state. More precisely, for a bipartite
state $\omega^{AB}$, the maximum classical mutual information $I_c$ is defined by

$$I_c(A; B)_{\omega} = \max_{\{M_A^i\}, \{M_B^i\}} I(I_A; I_B),$$

where $\{M_A^i\}$ and $\{M_B^i\}$ are measurements on $A$ and $B$, and $I_A, I_B$ are the (random) outcomes of these measurements on the state $\omega^{AB}$. Incremental proportionality is the intuitive property that $\ell$ bits of communication between two parties can increase their mutual information by at most $\ell$ bits. DiVincenzo et al. [2004] considered the states

$$\omega^{XKC} = \frac{1}{2d} \sum_{k=1}^{d} \sum_{x=1}^{t} |x\rangle\langle x|^X \otimes |k\rangle\langle k|^K \otimes (U_k |x\rangle\langle x| U_k^\dagger)^C$$

(4.1)

where $U_1 = \text{id}$ and $U_2$ is the Hadamard transform. It was shown by DiVincenzo et al. [2004] that the classical mutual information $I_c(XK; C)_{\omega} = \frac{1}{2} \log d$. However, if the holder of the $C$ system also knows the value of $k$, then we can represent the global state by the following density operator

$$\omega^{XKC K'} = \frac{1}{2d} \sum_{k=1}^{d} \sum_{x=1}^{t} |x\rangle\langle x|^X \otimes |k\rangle\langle k|^K \otimes (U_k |x\rangle\langle x| U_k^\dagger)^C \otimes |k\rangle\langle k|^K'$$

It is easy to see that $I_c(XK; CK')_{\omega} = 1 + \log d$. This means that with only one bit of communication (represented by the register $K'$), the classical mutual information between systems $XK$ and $C$ jumped from $\frac{1}{2} \log d$ to $1 + \log d$. In other words, it is possible to unlock $\frac{1}{2} \log d$ bits of information (about $X$) from the quantum system $C$ using a single bit.

Hayden et al. [2004] proved an even stronger locking result. They generalize the state in equation (4.1) to

$$\omega^{XKC K'} = \frac{1}{td} \sum_{x=1}^{d} \sum_{k=1}^{t} |x\rangle\langle x|^X \otimes |k\rangle\langle k|^K \otimes (U_k |x\rangle\langle x| U_k^\dagger)^C \otimes |k\rangle\langle k|^K'$$

(4.2)

where $U_k$ are chosen independently at random according to the Haar measure. They show that for any $\epsilon > 0$, by taking $t = (\log d)^3$ and if $d$ is large enough,

$$I_c(X; C)_{\omega} \leq \epsilon \log d \quad \text{and} \quad I_c(XK; CK')_{\omega} = \log d + \log t$$

with high probability. Note that the size of the key measured in bits is only $\log t = O(\log \log d)$ and it should be compared to the $(1 - \epsilon) \log d$ bits of unlocked (classical) information. It should be noted that their argument is probabilistic, and it does not say how to construct the unitary transformations $U_k$. Standard derandomization techniques are
4.1. Locking classical information in quantum states

not known to work in this setting. For example, unitary $t$-designs use far too many bits of randomness [Dankert et al., 2009]. Moreover, using a $\delta$-biased subset of the set of Pauli matrices fails to produce a locking scheme unless the subset has a size of the order of the dimension $d$ [Ambainis and Smith, 2004; Desrosiers and Dupuis, 2010] (see Section 4.1.3).

Here, we view locking as a cryptographic task in which a message is encoded into a quantum state using a key whose size is much smaller than the message. Having access to the key, one can decode the message. However, an eavesdropper who does not have access to the key and has complete uncertainty about the message can extract almost no classical information about the message.

**Definition 4.1.1** ($\epsilon$-locking scheme). Let $n$ be a positive integer, $\ell \in [0, n]$ and $\epsilon \in [0, 1]$. An encoding $E : [2^n] \times [t] \rightarrow S(C)$ is said to be $(\ell, \epsilon)$-locking for the quantum system $C$ if:

- For all $x \neq x' \in [2^n]$ and all $k \in [t]$, $\Delta(E(x, k), E(x', k)) = 1$.
- Let $X$ (the message) be a random variable on $[2^n]$ with min-entropy $H_{\text{min}}(X) \geq \ell$, and $K$ (the key) be an independent uniform random variable on $[t]$. For any measurement $\{M_i\}$ on $C$ and any outcome $i$,

$$\Delta(p_{X|I=i}, p_X) \leq \epsilon,$$

where $I$ is the outcome of measurement $\{M_i\}$ on the (random) quantum state $E(X, K)$.

When the min-entropy bound $\ell$ is not specified, it should be understood that $\ell = n$ meaning that $X$ is uniformly distributed on $[2^n]$. The state $E(x, k)$ for $x \in [2^n]$ and $k \in [t]$ is referred to as the ciphertext.

**Remark.** The relevant parameters of a locking scheme are: the number of bits $n$ of the (classical) message, the dimension $d$ of the (quantum) ciphertext, the number $t$ of possible values of the key and the error $\epsilon$. Strictly speaking, a classical one-time pad encryption, for which $t = 2^n$, is $(0, 0)$-locking according to this definition. However, here we seek locking schemes for which $t$ is much smaller than $2^n$, say polynomial in $n$. This cannot be achieved using a classical encryption scheme.

In the remainder of this section, we comment on the definition. We should stress first that this is not a composable cryptographic task, namely because an eavesdropper could choose to store quantum information about the message instead of measuring. In fact, as shown by [König et al., 2007], using the communicated message $X$ as a key for a one-time pad encryption might not be secure; see also [Dupuis et al., 2010b].
Thus, a locking map destroys almost all classical correlations with the message, but it is impossible to erase all quantum correlations with a key significantly smaller than the message. For example, consider a map \(E : \{0, 1\}^n \times \mathbb{N} \rightarrow S(C)\) such that the requirement (4.3) is replaced by \(I(X; C) \leq \delta\), where \(I\) is the quantum mutual information computed for the state \(\omega^{XKC} = \frac{1}{t^n} \sum_{x,k} |x\rangle_x \langle x| \otimes |k\rangle^K \otimes E(x, k)^C\). We have

\[
H(X) = H(X) + H(CK) - H(CK)
\leq H(X) + H(C) + H(K) - H(CK).
\]

Now we use the fact for all \(k\), the states \(\{E(x, k)\}_x\) are perfectly distinguishable. Thus, there exists an isometry that maps \(\omega^{CK} K\) to \(\omega^{CKX}\). Hence, \(H(CK) \omega = H(CKX)\). As a result,

\[
H(X) \leq H(X) + H(C) + H(K) - H(CKX)
\leq H(K) + H(X) + H(C) - H(CX)
= H(K) + I(X; C).
\]

This argument shows that if the key is much smaller than the message, then the quantum mutual information between the message and the ciphertext is large, it is in fact at least the size of the message minus the size of the key. It is basically the same argument that Shannon used to prove that any perfect encryption scheme has to use a key of size at least the message size [Shannon, 1949]. The reason this argument fails for the classical mutual information \(I_c\) is that the measurement to be made on the ciphertext to decode correctly depends on the value taken by the key. So replacing the system \(C\) by the outcome \(I\) of some fixed measurement on \(C\), the inequality \(H(IK) \geq H(IX)\) does not hold.

One could compare a locking scheme to an entropically secure encryption scheme [Dodis and Smith, 2005, Russell and Wang, 2002]. These two schemes achieve the same task of encrypting a high entropy message using a small key. The security definition of a locking scheme is strictly stronger. In fact, for a classical eavesdropper (i.e., an eavesdropper that can only measure) an \(\epsilon\)-locking scheme is secure in a strong sense. This additional security guarantee comes at the cost of upgrading classical communication to quantum communication. With respect to quantum entropically secure encryption [Desrosiers, 2009, Desrosiers and Dupuis, 2010], the security condition of a locking scheme is also more stringent (see Section 4.1.3 for an example of an entropically secure encryption scheme that is not \(\epsilon\)-locking). However, a quantum entropically secure scheme allows the encryption of quantum states.

We mentioned that if the adversary has no quantum storage, then a message that is transmitted using a locking scheme can be used in subsequent protocols. In the following
4.1. Locking classical information in quantum states

proposition, we show that it is still safe to re-use the transmitted message provided the
adversary is only allowed to have a small quantum memory. We follow the same technique
as in [Hallgren et al. 2010, Corollary 2].

**Proposition 4.1.2.** Let $E$ be an $\epsilon$-locking scheme and $F : C \rightarrow YQ$ be a (eavesdropping)
completely positive trace preserving map that sends all states $E(x, k)$ to states on $YQ$ that
are classical on $Y$. Then, we have

$$\Delta(\omega^{XYQ}, \omega^X \otimes \omega^{YQ}) \leq \epsilon \cdot c \sqrt{d},$$

where $\omega^{XYQ} = \frac{1}{d_X d_K} \sum_{x,k} |x\rangle\langle x| \otimes |k\rangle\langle k| \otimes F(E(x, k))$ and $c$ is a constant.

**Proof** The idea is to use the fact that there exists a measurement that can be used
to distinguish any pair of states reasonably well. More precisely, we use a result of
Ambainis and Emerson [2007, Theorem 4] that states that there exists a measurement map
$N : L(Q) \rightarrow L(Z)$ such that for any $\omega_1, \omega_2$,

$$\|N(\omega_1) - N(\omega_2)\|_1 \geq c d_Q^{-1/2} \|\omega_1 - \omega_2\|_1,$$

for some constant $c$; see also Radhakrishnan et al. [2009]. We will need a slightly more
general statement that applies to non-normalized states: for any $p_1 \geq p_2 \geq 0$,

$$\|p_1 N(\omega_1) - p_2 N(\omega_2)\|_1 \geq c/4 \cdot d_Q^{-1/2} \|\omega_1 - \omega_2\|_1. \tag{4.6}$$

In order to prove this, we proceed as in [Hallgren et al. 2010, Corollary 2]. We denote by
$\{\mu_1(z)\}_z$ and $\{\mu_2(z)\}_z$ the outcome distributions for measurement $N$ on the states $\omega_1$ and
$\omega_2$. We have

$$\|p_1 N(\omega_1) - p_2 N(\omega_2)\|_1 = \sum_z |p_1 \mu_1(z) - p_2 \mu_2(z)|$$

$$= \sum_z |p_1 (\mu_1(z) - \mu_2(z)) + (p_1 - p_2) \mu_2(z)|.$$

We now lower bound this expression in two different ways. First, we have

$$\sum_z |p_1 (\mu_1(z) - \mu_2(z)) + (p_1 - p_2) \mu_2(z)| \geq \sum_{z: \mu_1(z) \geq \mu_2(z)} |p_1 (\mu_1(z) - \mu_2(z))|$$

$$= p_1 \frac{\|\mu_1 - \mu_2\|_1}{2},$$
using the fact that $\mu_1$ and $\mu_2$ are probability distributions. Second, we have

$$
\sum_z |p_1(\mu_1(z) - \mu_2(z)) + (p_1 - p_2)\mu_2(z)| \geq \left| \sum_z p_1(\mu_1(z) - \mu_2(z)) + (p_1 - p_2)\mu_2(z) \right| = |p_1 - p_2|.
$$

Thus,

$$
\|p_1\mathcal{N}(\omega_1) - p_2\mathcal{N}(\omega_2)\|_1 \geq \frac{|p_1 - p_2|}{2} + \frac{p_1\|\mathcal{N}(\omega_1) - \mathcal{N}(\omega_2)\|_1}{4} \geq \frac{|(p_1 - p_2)\omega_2|}{2} + \frac{p_1c\ell^{-1/2}\omega_1 - \omega_2\|_1}{4} \geq c\ell^{-1/2}/4 \cdot \|p_1\omega_1 - p_2\omega_2\|_1.
$$

In the second inequality, we used the property (4.5) of the measurement $\mathcal{N}$ and in the third inequality, we used the triangle inequality. This proves property (4.6).

We are now in a position to prove the desired result. Let $\omega_{XYQ} = \frac{1}{d_X} \sum_{x,y} |x\rangle\langle x| \otimes p_{Y|x}(y|y) \otimes \omega_{x,y}^Q$. We have

$$
\Delta(\omega_{XYQ}, \omega^X \otimes \omega^Q_Y) = \Delta \left( \sum_{x,y} \frac{|x\rangle\langle x|}{d_X} \otimes |y\rangle\langle y| \otimes p_{Y|x}(y|y) \omega_{x,y}, \frac{id^X}{d_X} \otimes \sum_{x,y} \frac{p_{Y|x}(y|y)}{d_X} |y\rangle\langle y| \otimes \omega_{x,y} \right).
$$

Letting $\omega_y = \frac{1}{p_{Y|(y)}(y)} \sum_{x} \frac{p_{Y|x}(y|y)}{d_X} \omega_{x,y}$, we can write

$$
\Delta(\omega_{XYQ}, \omega^X \otimes \omega^Q_Y) = \frac{1}{d_X} \sum_{x,y} \frac{4d^{1/2}_Q}{c} \Delta \left( p_{Y|x}(y|y) \mathcal{N}(\omega_{x,y}), p_Y(y)\omega_y \right) \leq \frac{4d^{1/2}_Q}{c} \Delta \left( \sum_{x,y} \frac{|x\rangle\langle x|}{d_X} \otimes p_{Y|x}(y|y) \mathcal{N}(\omega_{x,y}), \frac{id^X}{d_X} \otimes \sum_{y} |y\rangle\langle y| \otimes p_Y(y)\mathcal{N}(\omega_y) \right) \leq \frac{4d^{1/2}_Q}{c} \Delta(p_{XYZ}, p_X \times p_{YZ}),
$$

where $Y, Z$ are obtained by performing the measurement defined by $(id^Y \otimes \mathcal{N}^{Q\rightarrow Z}) \circ \mathcal{F}$ on the state $\mathcal{E}(X, K)$. We conclude by using the fact that $\mathcal{E}$ is $\epsilon$-locking so that $\Delta(p_{XYZ}, p_X \times p_{YZ}) \leq \epsilon$. \qed
Proposition 4.1.2 is interesting for the scheme presented in Corollary 4.1.7 below, for which the sender and the receiver do not use any quantum memory. One could then use such a scheme for key distribution in the bounded quantum storage model, where the adversary is only allowed to have a quantum memory of logarithmic size in \( n \) but can have an arbitrarily large classical memory. Note that even though this is a strong assumption compared to the unconditional security of BB84 [Bennett and Brassard, 1984], one advantage of such a protocol for key distribution is that it only uses one-way communication between the two parties. In contrast, the BB84 quantum key distribution protocol needs interaction between the two parties.

Another remark about Definition 4.1.1 is that we used the statistical distance between \( p_{X|I=i} \) and \( p_X \) instead of the mutual information between \( X \) and \( I \) to measure the information gained about \( X \) from a measurement. Using the trace distance is a stronger requirement as demonstrated by the following proposition.

**Proposition 4.1.3.** Let \( \epsilon \in [0, 1/2] \) and \( \mathcal{E} : [2^n] \times [t] \rightarrow S(C) \) be an \( \epsilon \)-locking scheme. Define the state

\[
\omega^{XKCK'} = \frac{1}{td} \sum_{k=1}^t \sum_{x=1}^{2^n} |x\rangle\langle x|_X \otimes |k\rangle\langle k|_K \otimes \mathcal{E}(x,k)_C \otimes |k\rangle\langle k|_K'.
\]

Then,

\[
I_c(X;C)_\omega \leq 8\epsilon n + 2h_2(2\epsilon) \quad \text{and} \quad I_c(XK;CK')_\omega = n + \log t.
\]

**Proof** First, we can suppose that the measurement performed on the system \( X \) is in the basis \( \{|x\rangle\}_x \). In fact, the outcome distribution of any measurement on the \( X \) system can be simulated classically using the values of the random variables \( X \).

Now let \( I \) be the outcome of a measurement performed on the \( C \) system. Using Fannes’ inequality (a special case of Lemma 2.2.5), we have for any \( i \)

\[
H(X) - H(X|I = i) \leq 8\Delta(p_X, p_X|I=i) - 2h_2 \left( 2\Delta(p_X, p_X|I=i) \right)
\]

\[
\leq 8\epsilon n + 2h_2(2\epsilon)
\]

using the fact that \( \mathcal{E} \) defines an \( \epsilon \)-locking scheme. Thus,

\[
I(X;I) = H(X) - \sum_i \Pr \{ I = i \} H(X|I = i)
\]

\[
\leq 8\epsilon n + 2h_2(2\epsilon).
\]
As this holds for any measurement, we get $I_c(X; C) \leq 8\epsilon n + 2h_2(2\epsilon)$.

The trace distance was also used in Dupuis [2010], Dupuis et al. [2010b] to define a locking scheme. To measure the leakage of information about $X$ caused by a measurement, they used the probably more natural trace distance between the joint distribution of $p_{(X, I)}$ and the product distribution $p_X \times p_I$. Note that our definition is stronger, in that for all outcomes of the measurement $i$, $\Delta(p_{X[I=i]}, p_X) \leq \epsilon$ whereas the definition of Dupuis et al. [2010b] says that this only holds on average over $i$. The condition of Dupuis et al. [2010b] is probably sufficient for most applications but our techniques naturally achieve the stronger form without degrading the parameters. We should finally note that the trace distance condition cannot be much stronger than the condition on the classical mutual information. In fact, using Pinsker’s inequality, we can upper bound the trace distance using the mutual information:

$$\Delta(p_{(X, I)}, p_X \times p_I) \leq \sqrt{I(X; I)/2}.$$ 

For a survey on locking classical correlations, see Leung [2009].

**Other related work**

In a cryptographic setting, Damgård et al. [2004] used ideas related to locking to develop quantum ciphers that have the property that the key used for encryption can be recycled. In Damgård et al. [2005], they construct a quantum key recycling scheme (see also Oppenheim and Horodecki [2005]) with near optimal parameters by encoding the message together with its authentication tag using a full set of mutually unbiased bases.

### 4.1.2 Locking using a metric uncertainty relation

The following theorem shows that a locking scheme can easily be constructed using a metric uncertainty relation.

**Theorem 4.1.4.** Let $\epsilon \in (0, 1)$ and $\{U_1, \ldots, U_t\}$ be a set of unitary transformations of $A \otimes B$ that satisfies an $\epsilon$-metric uncertainty relation on $A$, i.e., for all states $|\psi\rangle \in AB$,

$$\frac{1}{t} \sum_{k=1}^{t} \Delta(p_{U_k|\psi}, \text{unif}([d_A])) \leq \epsilon.$$

Assume $d_A = 2^n$. Then, the mapping $E : [2^n] \times [t] \rightarrow S(AB)$ defined by

$$E(x, k) = \frac{1}{d_B} \sum_{b=1}^{d_B} U_k^\dagger (|x\rangle^A \otimes |b\rangle^B) U_k,$$
is $\epsilon$-locking. Moreover, for all $\ell \in [0, n]$ such that $2^{\ell-n} > \epsilon$, it is $(\ell, \frac{2\epsilon}{2^{\ell-n}-\epsilon})$-locking.

**Remark.** Figure 4.1 illustrates the locking scheme. The state that the encoder inputs in the $B$ system is simply private randomness. The encoder chooses a uniformly random $b \in [d_B]$ and sends the quantum state $U_k^\dagger|x\rangle^A|b\rangle^B$. Note that $b$ does not need to be part of the key (i.e., shared with the receiver). This makes the dimension $d = d_Ad_B$ of the ciphertext larger than the number of possible messages $2^n$. If one insists on having a ciphertext of the same size as the message, it suffices to consider $b$ as part of the message and apply a one-time pad encryption to $b$. The number of possible values taken by the key increases to $t \cdot d_B$.

![Diagram](image_url)

**Figure 4.1:** Illustration of the locking scheme described in Theorem 4.1.4.

**Proof**  
First, it is clear that different messages are distinguishable. In fact, for $x \neq x'$ and any $k$,

$$
\Delta(\mathcal{E}(x, k), \mathcal{E}(x', k)) = \frac{1}{2} \text{tr} \left[ \sqrt{|x\rangle \langle x|^A \otimes \frac{\text{id}_B}{d_B}} - |x'\rangle \langle x'|^A \otimes \frac{\text{id}_B}{d_B} \right] = 1.
$$

We now prove the locking property. Let $X$ be the random variable representing the message. Assume that $X$ is uniformly distributed over some set $S \subseteq [d_A]$ of size $|S| \geq 2^\ell$. Let $K$ be a uniformly random key in $[t]$ that is independent of $X$. Consider a POVM $\{M_i\}$ on the system $AB$. Without loss of generality, we can suppose that the POVM elements $M_i$ have rank 1. Otherwise, by writing $M_i$ in its eigenbasis, we could decompose outcome $i$ into more outcomes that can only reveal more information. So we can write the elements as weighted rank one projectors: $M_i = \xi_i |e_i\rangle \langle e_i|$ where $\xi_i > 0$. Our objective is to show that the outcome $I$ of this measurement on the state $\mathcal{E}(X, K)$ is almost independent of $X$. More precisely, for a fixed measurement outcome $I = i$, we want to compare the conditional
distribution \( p_{X|i} \) with \( p_X \). The trace distance between these distributions can be written as

\[
\frac{1}{2} \sum_{x=1}^{d_A} \left| \Pr \{ X = x | I = i \} - \Pr \{ X = x \} \right|.
\] (4.7)

Towards this objective, we start by computing the distribution of the measurement outcome \( I \), given the value of the message \( X = x \) (note that the receiver does not know the key):

\[
\Pr \{ I = i | X = x \} = \frac{\xi_i}{td_B} \sum_{k=1}^{t} \sum_{b=1}^{d_B} \text{tr} \left[ U_k |e_i\rangle \langle e_i| U_k^\dagger \cdot \! |x\rangle \langle x| \otimes |b\rangle \langle b| \right]
\]

\[
= \frac{\xi_i}{td_B} \sum_{k=1}^{t} \sum_{b=1}^{d_B} \langle x|^{(A)} \langle b|^{(B)} U_k |e_i\rangle \langle e_i| U_k^\dagger \rangle \langle x|^{(A)} |b\rangle^{(B)}
\]

\[
= \frac{\xi_i}{td_B} \sum_{k=1}^{t} \sum_{b=1}^{d_B} |\langle x|^{(A)} \langle b|^{(B)} U_k |e_i\rangle |^2
\]

\[
= \frac{\xi_i}{d_B} \frac{1}{t} \sum_{k=1}^{t} p_{A|e_i}^U(x).
\]

Since \( X \) is uniformly distributed over \( S \), we have that for all \( x \in S \)

\[
\Pr \{ X = x | I = i \} = \frac{\Pr \{ X = x \} \Pr \{ I = i | X = x \}}{\sum_{x' \in S} \Pr \{ X = x' \} \Pr \{ I = i | X = x' \}}
\]

\[
= \frac{(1/t) \cdot \sum_k p_{A|e_i}^U(x)}{(1/t) \cdot \sum_{x' \in S} \sum_k p_{A|e_i}^U(x')}.
\] (4.8)

Observe that in the case where \( X \) is uniformly distributed over \([2^n] (S = [2^n])\), it is simple to obtain directly that

\[
\Delta(p_{X|i}, p_X) = \frac{1}{2} \sum_{x=1}^{d_A} \left| \frac{1}{t} \sum_{k=1}^{t} p_{A|e_i}^U(x) - \frac{1}{2^n} \right| \leq \epsilon
\]
4.1. Locking classical information in quantum states

using the fact that \( \{U_k\} \) satisfies a metric uncertainty relation on \( A \). Now let \( S \) be any set of size at least \( 2^\ell \) and let \( \alpha = \frac{1}{t} \sum_{x' \in S} \sum_k p_{U_k[e_i]}^A(x') \). We then bound

\[
\frac{1}{2} \sum_{x=0}^{d_A-1} |\Pr \{X = x | I = i\} - \Pr \{X = x\}| = \frac{1}{2} \sum_{x \in S} \left| \frac{1}{t} \sum_{k=1}^{t} p_{U_k[e_i]}^A(x) - \frac{\alpha}{|S|} \right| \\
\leq \frac{1}{2\alpha} \cdot \frac{1}{t} \sum_{x \in S} \left( \sum_{x' \in S} \left| \sum_{k=1}^{t} p_{U_k[e_i]}^A(x') - \frac{1}{2^n} \right| + \left| \frac{1}{2^n} - \frac{\alpha}{|S|} \right| \right).
\]

We now use the fact that \( \{U_k\} \) satisfies a metric uncertainty relation on \( A \): we get

\[
\frac{1}{t} \sum_{k} \frac{1}{2} \sum_{x \in S} \left| p_{U_k[e_i]}^A(x) - \frac{1}{2^n} \right| \leq \frac{1}{t} \sum_{k} \frac{1}{2} \sum_{x' \in S} \left| p_{U_k[e_i]}^A(x') - \frac{1}{2^n} \right| \leq \epsilon
\]

and

\[
\frac{1}{2} \left| \frac{|S|}{2^n} - \alpha \right| = \frac{1}{2} \left| \frac{|S|}{2^n} - \frac{1}{t} \sum_{x' \in S} \sum_{k=1}^{t} p_{U_k[e_i]}^A(x') \right| \leq \epsilon. \tag{4.9}
\]

As a result, we have

\[
\Delta(p_{X[I=i]}, p_X) \leq \frac{2\epsilon}{\alpha}.
\]

Using (4.9), we have \( \alpha \geq |S|2^{-n} - \epsilon \geq 2^{\ell-n} - \epsilon \). If \( \epsilon < 2^{\ell-n} \), we get

\[
\Delta(p_{X[I=i]}, p_X) \leq \frac{2\epsilon}{2^{\ell-n} - \epsilon}.
\]

In the general case when \( X \) has min-entropy \( \ell \), the distribution of \( X \) can be seen as a mixture of uniform distributions over sets of size at least \( 2^\ell \). So there exist independent random variables \( J \in \mathbb{N} \) and \( \{X_j\} \) uniformly distributed on sets of size at least \( 2^\ell \) such that \( X = X_J \). One can then write

\[
\frac{1}{2} \sum_{x} |\Pr \{X = x | I = i\} - \Pr \{X = x\}| = \frac{1}{2} \sum_{x,j} |\Pr \{J = j\} \left( \Pr \{X_j = x | I = i, J = j\} - \Pr \{X_j = x | J = j\} \right) | \leq \frac{2\epsilon}{2^{\ell-n} - \epsilon}.
\]
Using Theorem 4.1.4 together with the existence of metric uncertainty relations (Theorem 3.3.2), we show the existence of ε-locking schemes whose key size depends only on ε and not on the size of the encoded message.

**Corollary 4.1.5** (Existence of locking schemes). Let \( n \) be a large enough integer and \( \epsilon \in (0, 1) \). Then there exists an \( \epsilon \)-locking scheme encoding an \( n \)-bit message using a key of at most \( 2 \log(1/\epsilon) + O(\log \log(1/\epsilon)) \) bits into at most \( n + 2 \log(18/\epsilon) \) qubits.

**Remark.** Observe that in terms of number of bits, the size of the key is only a factor of two larger (up to smaller order terms) than the lower bound of \( \log(1/(\epsilon + 2^{-n})) \) bits that can be obtained by guessing the key. In fact, consider the strategy of performing the decoding operation corresponding to the key value 0. In this case, we have \( \Pr \{ X = i | I = i \} \geq \Pr \{ K = 0 \} = 1/t \). Thus, \( \Delta(p_{X|I=i},p_X) \geq 1/t - 2^{-n} \).

Recall that we can increase the size of the message to be equal to the number of qubits of the ciphertext at the cost of increasing the key size to at most \( 4 \log(1/\epsilon) + O(\log(\log(1/\epsilon)) \).

**Proof** Use the construction of Theorem 3.3.2 with \( d_A = 2^n \) and \( d_B = 2^q \) such that \( 2^{q-1} < 9/\epsilon^2 \leq 2^q \) and \( d = d_A d_B \). Take \( t = 2^p \) to be the power of two with \( 2^{p-1} \leq 4 \cdot \frac{18 \cdot \log(9/\epsilon)}{\epsilon^2} < 2^p \).

To construct \( (\ell, \epsilon) \)-locking schemes with \( \ell < n \), it suffices to use Theorem 3.3.2 with say \( \epsilon' = 2^{\ell-n} \epsilon/4 \). In this case, we obtain a key of size \( O((n - \ell) + \log(1/\epsilon)) \). We note that this increase in the key size is unavoidable because of the following proposition.

**Proposition 4.1.6.** Assume \( \mathcal{E} \) defines an \( (\ell, \epsilon) \)-locking scheme with \( \epsilon \leq 1/4 \) and a key of size \( \log t \). Then \( \log t \geq n - \ell - 2 \).

**Proof** We proceed as in the proofs of lower bounds on the key size in entropic security [Dodis and Smith, 2005]. The idea is to show that if \( \mathcal{E} \) is an \( (\ell, \epsilon) \)-locking scheme, then it can be used to build an encryption scheme for messages of \( n - \ell \) bits that has the following properties. Given the secret key, the encryption of \( w \) and \( w' \) are perfectly distinguishable, but without the key, the encryption of \( w \) and \( w' \) are almost indistinguishable. For such a scheme, we show in Proposition A.3.1 that the key size is at least the size of the message \( \log t \geq n - \ell - 2 \).

Define the random variables \( X_w \) for \( w \in \{0,1\}^{n-\ell} \) which are uniformly distributed on \( \{w\} \times \{0,1\}^\ell \). Our encryption scheme encrypts \( w \) using the key \( k \) into \( \mathcal{E}(X_w,k) \). First, clearly a decoder having the key can determine \( w \) using \( \mathcal{E}(X_w,k) \). Second, we show that
the ciphertexts corresponding to \( w \) and \( w' \neq w \) are almost indistinguishable:

\[
\Delta(\rho_w, \rho_{w'}) \leq 2\epsilon, \tag{4.10}
\]

where \( \rho_w = \frac{1}{k^2} \sum_{k \in \mathbb{E}[t], y \in \{0,1\}^\ell} \mathcal{E}(w \cdot y, k) \). To show this we let \( \Lambda \) be a positive operator such that \( \Delta(\rho_w, \rho_{w'}) = \text{tr}[\Lambda(\rho_w - \rho_{w'})] \) (see equation (2.5)).

We then have

\[
\text{tr}[\Lambda(\rho_w - \rho_{w'})] \leq \left| \text{tr} \left[ \Lambda \left( \frac{\rho_w + \rho_{w'}}{2} - \rho_w \right) \right] \right| + \left| \text{tr} \left[ \Lambda \left( \frac{\rho_w + \rho_{w'}}{2} - \rho_{w'} \right) \right] \right|
\]

\[
\leq \frac{1}{2^\ell} \sum_y \left| \text{tr} \left[ \Lambda \left( \frac{\rho_w + \rho_{w'}}{2} - \frac{1}{k} \sum_{k \in \mathbb{E}[t]} \mathcal{E}(w' \cdot y, k) \right) \right] \right|
\]

\[
+ \frac{1}{2^\ell} \sum_y \left| \text{tr} \left[ \Lambda \left( \frac{\rho_w + \rho_{w'}}{2} - \frac{1}{k} \sum_{k \in \mathbb{E}[t]} \mathcal{E}(w \cdot y, k) \right) \right] \right|
\]

\[
= 2 \frac{1}{2^{\ell+1}} \sum_{z \in \{w,w'\} \times \{0,1\}^\ell} \left| \text{tr} \left[ \Lambda \left( \frac{\rho_w + \rho_{w'}}{2} - \frac{1}{k} \sum_{k \in \mathbb{E}[t]} \mathcal{E}(z, k) \right) \right] \right|
\]

\[
= 2 \sum_{z \in \{w,w'\} \times \{0,1\}^\ell} \Pr \{Z = z\} |\Pr \{I = 0\} - \Pr \{I = 0|Z = z\}|
\]

\[
= 2 \Delta(p_{ZI}, p_Z \times p_I),
\]

where \( Z \) is uniformly distributed on \( \{w,w'\} \times \{0,1\}^\ell \) and \( I \) is the outcome of the measurement \( \{\Lambda, \text{id} - \Lambda\} \) performed on the state \( \mathcal{E}(Z,K) \). Inequality (4.10) follows from the fact that \( \mathcal{E} \) is an \((\ell,\epsilon)\)-locking scheme.

Using Proposition [A.3.1] we conclude that \( \log t \geq n - \ell - 2 \). \( \square \)

The following corollary gives explicit locking schemes. We mention the constructions based on Theorems 3.4.6 and 3.4.7. Of course, one could obtain a tradeoff between the key size and the dimension of the quantum system.

**Corollary 4.1.7 (Explicit locking schemes).** Let \( \delta > 0 \) be a constant, \( n \) be a positive integer, \( \epsilon \in (2^{-c'n}, 1) \) (\( c' \) is a constant independent of \( n \)).

- Then, there exists an efficient \( \epsilon \)-locking scheme encoding an \( n \)-bit message in a quantum state of \( n' \leq (4 + \delta)n + O(\log(1/\epsilon)) \) qubits using a key of size \( O(\log(n/\epsilon)) \) bits. In fact, both the encoding and decoding operations are computable using a classical computation with polynomial running time and a quantum circuit with only Hadamard gates and preparations and measurements in the computational basis.
• There also exists an efficient \(\epsilon\)-locking scheme \(\mathcal{E}'\) encoding an \(n\)-bit message in a quantum state of \(n\) qubits using a key of size \(O(\log(n/\epsilon) \cdot \log n)\) bits. \(\mathcal{E}'\) is computable by a classical algorithm with polynomial runtime and a quantum circuit of size \(O(n \text{ polylog}(n/\epsilon))\).

**Proof** For the first result, we observe that the construction of Theorem 4.1.4 encodes the message in the computational basis. Recall that the unitaries \(U_k\) of Theorem 3.4.6 are of the form \(U_k = P_k V_k\) where \(P_k\) is a permutation of the computational basis. Hence, it is possible to classically compute the label of the computational basis element \(P_k^\dagger |x\rangle |b\rangle\). One can then prepare the state \(P_k^\dagger |x\rangle |b\rangle\) and apply the unitary \(V_k^\dagger\) to obtain the ciphertext. The decoding is performed in a similar way. One first applies the unitary \(V_k\), measures in the computational basis and then applies the permutation \(P_k\) to the \(n\)-bit string corresponding to the outcome.

For the second construction, we apply Theorem 3.4.7 with \(n' = n + c' \lceil \log \log n + \log(1/\epsilon) \rceil\) for some large enough constant \(c'\). We can then use a one-time pad encryption on the input to the \(B\) system. This increases the size of the key by only \(c' \lceil \log \log n + \log(1/\epsilon) \rceil\) bits. \(\Box\)

As mentioned earlier (see equation (4.1)), explicit states that exhibit locking behaviour have been presented in DiVincenzo et al. [2004]. However, this is the first explicit construction of states \(\omega\) that achieves the following strong locking behaviour: for any \(\delta > 0\), for \(n\) large enough, the state \(\omega^{XCK}\) verifies \(I_c(X; C)_{\omega} \leq \delta\) and \(I_c(X; CK)_{\omega} = n + \log d_K\) where \(K\) is a classical \(O(\log(n/\delta))\)-bit system. This is a direct consequence of Corollary 4.1.7 taking \(\epsilon = \delta/(20n)\), and Proposition 4.1.3. We should also mention that König et al. [2007] explicitly construct a state exhibiting some weak locking behaviour. We summarize the different locking schemes in Table 4.1.

### 4.1.3 Impossibility of locking using Pauli operators

The objective of this section is to give an example of a construction that is not a locking scheme to illustrate what is needed to obtain a locking scheme. The \(2 \times 2\) Pauli matrices are the four matrices \(\{I, \sigma_x, \sigma_z, \sigma_x \sigma_z\}\) where

\[
\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \text{and} \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.
\]

For bit strings \(u, v \in \{0, 1\}^n\), we define the unitary operation \(\sigma_x^u \sigma_z^v\) on \((\mathbb{C}^2)^{\otimes n}\) by

\[
\sigma_x^u \sigma_z^v = \sigma_x^{u_1} \sigma_z^{v_1} \otimes \cdots \otimes \sigma_x^{u_n} \sigma_z^{v_n}.
\]

It was shown by Ambainis et al. [2000] that one can encrypt an \(n\)-qubit state \(|\psi\rangle\) perfectly using a key \((U, V)\) of \(2n\) bits. To encrypt \(|\psi\rangle\), one simply applies \(\sigma_x^U \sigma_z^V\) to \(|\psi\rangle\), where \(U\)
Table 4.1: Comparison of different locking schemes. \( n \) is the number of bits of the message. The information leakage and the size of the key are measured in bits and the size of the ciphertext in qubits. Efficient locking schemes have encoding and decoding quantum circuits of size polynomial in \( n \). The locking schemes of the first and next to last actually have encoding circuits that are in principle implementable with current technology; they only use classical computations and simple single-qubit transformations. It should be noted that our locking definition is stronger than all the previous definitions. Note that the variable \( \epsilon \) can depend on \( n \). For example, one can take \( \epsilon = \eta/n \) to make the information leakage arbitrarily small. The symbol \( O(\cdot) \) refers to constants independent of \( \epsilon \) and \( n \), but there is a dependence on \( \delta \) for the next to last row. The symbol \( \ll(\cdot) \) refers to \( O(\log \log(\cdot)) \).

|                      | Inf. leak. | Key          | Ciphertext | Efficient? |
|----------------------|------------|--------------|------------|------------|
| DiVincenzo et al.    | \( n/2 \)  | 1            | \( n \)    | yes        |
| Hayden et al.        | 3          | \( 4 \log(n) \) | \( n \)    | no         |
| Dupuis et al.        | \( \epsilon n \) | \( 2 \log(n/\epsilon^2) + O(1) \) | \( n \)    | no         |
| Corollary 4.1.5      | \( \epsilon n \) | \( 2 \log(1/\epsilon) + \ll(1/\epsilon) \) | \( n + 2 \left\lceil \log(9/\epsilon) \right\rceil \) | no         |
| Corollary 4.1.5      | \( \epsilon n \) | \( 4 \log(1/\epsilon) + \ll(1/\epsilon) \) | \( n \)    | no         |
| Corollary 4.1.7      | \( \epsilon n \) | \( O_\delta(\log(n/\epsilon)) \) | \( (4 + \delta) \cdot n \) | yes        |
| Corollary 4.1.7      | \( \epsilon n \) | \( O(\log(n/\epsilon) \log(n)) \) | \( n \)    | yes        |
and $V$ are uniformly distributed on $\{0,1\}^n$. This can be thought of as a quantum version of one-time pad encryption. Of course, this encryption scheme also defines a $(0,0)$-locking scheme, but the size of the key is $2n$ bits. Recall that we want to use the assumption that the message is random to reduce the key size to $O(\text{polylog}(n))$ bits.

Ambainis and Smith [2004] showed that to achieve approximate encryption, it is sufficient to choose the key uniformly at random from a well-chosen subset $S \subseteq \{0,1\}^{2n}$ of size only $O(n^2 \cdot 2^n)$. Such pseudorandom subsets are called $\delta$-biased sets and have also been used to construct entropically secure encryption schemes [Desrosiers and Dupuis, 2010, Dodis and Smith, 2005]. For example, Desrosiers and Dupuis [2010] showed that it is possible to encrypt a uniformly random state by applying $\sigma^U_x \sigma^V_z$ where $(U,V)$ is chosen uniformly from a set $S \subseteq \{0,1\}^n$ of size $O(n^2)$ (see [Desrosiers and Dupuis, 2010, Dodis and Smith, 2005] for a precise definition of entropic security). Such a set of transformations can seem like a good candidate for a locking scheme. The following proposition shows that this scheme is far from being $\epsilon$-locking. Note that this also shows that the notion of entropic security defined in [Desrosiers, 2009, Desrosiers and Dupuis, 2010] is weaker than the definition of locking.

**Proposition 4.1.8.** Consider an $\epsilon$-locking scheme $E$ of the form $E(x,k = (u,v)) = \sigma^u_x \sigma^v_y|x$ where the message $x \in \{0,1\}^n$ and the key $u,v \in \{0,1\}^n$ (see Definition 4.1.1). Suppose the secret key $K$ is chosen uniformly from a set $S \subseteq \{0,1\}^{2n}$. Then $|S| \geq (1 - \epsilon)2^n$.

**Proof.** Let $X$ be the message ($X$ is uniformly distributed over $\{0,1\}^n$) and $(U,V)$ be the key. The key is uniformly distributed on $S$. We show that a measurement in the computational basis gives a lot of information about $X$. Let $I$ be the outcome of measuring $E(X,K)$ in the computational basis. We have for $x,i \in \{0,1\}^n$,

$$\Pr\{X = x|I = i\} = \Pr\{I = i|X = x\} = \frac{1}{|S|} \sum_{(u,v) \in S} |\langle i| \sigma^u_x \sigma^v_y|x\rangle|^2.$$

Observing that the term $|\langle i| \sigma^u_x \sigma^v_y|x\rangle|^2 \in \{0,1\}$, we have that for any fixed $i$, there are at most $|S|$ different values of $x$ for which $\Pr\{X = x|I = i\} > 0$. Thus, defining $T = \{x \in \{0,1\}^n : \Pr\{X = x|I = i\} = 0\}$, we have

$$\Delta(p_X|I=i), p_X) \geq \Pr\{X \in T\} - \Pr\{X \in T|I = i\} = \frac{|T|}{2^n} \geq 1 - \frac{|S|}{2^n}.$$

By the definition of a locking scheme, we should have

$$\Delta(p_X|I=i), p_X) \leq \epsilon.$$
which concludes the proof. □

4.1.4 Error-tolerant information locking

The first protocol in Corollary 4.1.7 can in principle be implemented using current technology. We say “in principle” because a locking scheme as defined here does not allow for any error in the transmission of the ciphertext. Can we construct a locking scheme that can tolerate a reasonable rate of errors?

One simple approach to build a protocol that tolerates errors is to use a quantum error correcting code (QECC) to encode the ciphertext. Depending on the properties of the code, this would allow the receiver to correct some fraction of errors. Moreover, the security is preserved because an eavesdropper could perform an encoding into a QECC as part of his attack. Thus, it is possible to make any locking scheme error-tolerant provided we can perform encoding and decoding operations for a good QECC. Unfortunately, the encoding and decoding maps of interesting quantum error correcting codes are beyond the reach of current technology. But note that our objective is not necessarily to recover the quantum ciphertext correctly, we only want to be able to recover the classical message. Can we construct a locking scheme that can tolerate a reasonable rate of errors and that can be implemented with current technology?

In the remainder of this section, we show that some natural class of error tolerant protocols cannot be good locking schemes. Consider a locking scheme of the following form. The key is written as $k \in [t]$ and the message $x \in \{0, 1\}^n$ is locked in the following way:

- A classical (possibly randomized) function determined by the key $k$ is applied to $x$. $x$ is mapped to $P_k(x, r) \in \{0, 1\}^{n'}$, where $r$ is a random string private to Alice.

- The bitstring $P_k(x, r)$ is then encoded in a code $C_k$ possibly depending on the key $k$. The codes $C_k$ are assumed to have minimum distance $\alpha$ for all $k$. This bitstring is denoted $z = C_k(P_k(x, r)) \in \{0, 1\}^m$. We denote by $C_k$ the set of bitstrings in this code.

- A quantum encoding of the form $H^{v_k}$ where $v_k \in \{0, 1\}^m$ is performed on the computational basis element $|C_k(P_k(x, z))\rangle$.

We start with a lemma that says that given a set of vectors that are almost orthogonal, they can be well approximated by orthogonal vectors. It was first proven by Schönenmann [1966]; see also [Kempe and Vidick, 2010] Claim 20.
Lemma 4.1.9. Let \( s \leq d \) and \(|u_1\), \ldots, |u_s\rangle \in \mathbb{C}^d\) be unit vectors such that \( \frac{1}{s} \sum_{i \neq j} |\langle u_i | u_j \rangle|^2 \leq \epsilon \). Then there exist orthogonal unit vectors \(|v_1\), \ldots, |v_s\rangle\) such that \( \frac{1}{s} \sum_i \| |u_i\rangle - |v_i\rangle \|^2 \leq \epsilon \).

Proof. We start by fixing a set of orthonormal vectors \(|w_1\), \ldots, |w_s\rangle\) such that the span of \(|u_1\), \ldots, |u_s\rangle\) is included in the span of \(|w_1\), \ldots, |w_s\rangle\). We then define the matrix \( X \) whose columns represent the vectors \(|u_i\rangle\) using the vectors \(|w_1\), \ldots, |w_s\rangle\). Write the SVD decomposition of \( X \) as \( X = UV^\dagger \) and let the singular values of \( X \) be \( \sigma_1, \ldots, \sigma_s \). We have

\[
\sum_{i \in [s]} (1 - \sigma_i^2)^2 = \| \text{id} - \Sigma \Sigma^\dagger \|^2 = \| \text{id} - X^\dagger X \|^2 = \sum_{i \neq j} |\langle u_i | u_j \rangle|^2 + \sum_{i \in [s]} (1 - |\langle u_i | u_i \rangle|)^2 \\
\leq s \epsilon
\]

by assumption. We now define \( Y = UV^\dagger \), look at the columns of this matrix and call these vectors \(|v_i\rangle\) (of course the underlying basis used is still \(|w_i\rangle\)). We have by writing the desired expression in the basis \(|w_i\rangle\) and then multiplying by \( U^\dagger \) on the left and \( V \) on the right:

\[
\sum_i \| |u_i\rangle - |v_i\rangle \|^2 = \| X - Y \|^2 = \| \Sigma - \text{id} \|^2 = \sum_i (\sigma_i - 1)^2 \\
\leq \sum_i (1 - \sigma_i)^2 (1 + \sigma_i)^2 \leq s \epsilon.
\]

\( \square \)

Proposition 4.1.10. For any encoding of the form above such that the random variable \((Z, V) = (P_K(X, R), v_K)\) is uniformly distributed on its support, there exists a measurement on the ciphertext that gives an outcome \( I \) such that

\[
I(I; X) \geq (1 - 16 \cdot t2^{-\alpha+1})n - \log t - 2,
\]

provided \( t2^{-\alpha+1} \leq 1 \).

Remark. Recall that we want \( t \) to be sub-exponential (even polynomial) in \( n \). Moreover, to be able to correct a constant fraction of errors, we want the minimum distance \( \alpha \) to be linear in \( n' \). In this case, \( t2^{-\alpha+1} \ll 1 \) and the measurement given by the proposition completely breaks the locking scheme. \( \square \)

Proof. We will apply Lemma 4.1.9 to the set of vectors \( \{H^v|z\rangle : (z, v) \in S\} \) with \( S = \{(z, v) : \exists k : z \in C_k, v = v_k\} \). Note that a pair \((z, v)\) that appears for different
values of $k$ is counted only once. Fix any $z'$ and $v'$ (not necessarily in $S$). We have
\[
\sum_{(z,v) \in S, (z,v) \neq (z',v')} |\langle z' | H^{v'} H^{v_k} | z \rangle|^2 = \sum_{v_k = v', z \neq z'} |\langle z' | z \rangle|^2 + \sum_{v_k \neq v', z \in C_k} |\langle z' | H^{v'} H^{v_k} | z \rangle|^2
\]
\[
= 0 + \sum_{v_k \neq v', z \in C_k \cap B_{v_k+v'}(z')} |\langle z' | H^{v'} H^{v_k} | z \rangle|^2
\]
(4.11)
where $B_{v_k+v'}(z') = \{ y \in \{0, 1\}^m : y_i = z_i \text{ whenever } (v_k + v')_i = 0 \}$. Observe that for all $z$ and $v$, $|B_v(z)| \leq 2^w(v)$ and in this case $w(v_k + v') = d_H(v_k, v')$. We now fix some $k$ such that $v_k \neq v'$ and bound the size of $C_k \cap B_{v_k+v'}$ assuming that $C_k$ is a code of minimum distance $\alpha$. The strings of $C_k \cap B_{v_k+v'}$ agree on $2^{n-d_H(v_k,v')}$ bits. This means it induces a code of minimum distance $\alpha$ on strings of length $d_H(v_k, v')$. Using the Singleton bound, we get
\[
|C_k \cap B_{v_k+v'}| \leq 2^{d_H(v_k,v')-\alpha+1}.
\]
To bound the expression in (4.11), we observe that $|\langle z' | H^{v'} H^{v_k} | z \rangle|^2 \leq 2^{-d_H(v_k,v')}$. Thus, for a fixed $v_k$,
\[
\sum_{z \in C_k \cap B_{v_k+v'}(z')} |\langle z' | H^{v'} H^{v_k} | z \rangle|^2 \leq 2^{-d_H(v_k,v')} \cdot 2^{d_H(v_k,v')-\alpha+1}
\]
\[
\leq 2^{-\alpha+1}.
\]
As a result, we can bound the average over $(z', v') \in S$
\[
\frac{1}{|S|} \sum_{(z',v') \in S} \sum_{(z,v) \in S, (z,v) \neq (z',v')} |\langle z' | H^{v'} H^{v_k} | z \rangle|^2 \leq t \cdot 2^{-\alpha+1}.
\]
Using Lemma 4.1.9 for the set of vector $\{H^v | z\} : (z, v) \in S$}, we obtain a set of orthonormal vectors $|w_{z,v}\rangle$ for $(z, v) \in S$ such that $\frac{1}{|S|} \sum_{(z,v) \in S} \|H^v | z\rangle - |w_{z,v}\rangle\|^2 \leq t 2^{-\alpha+1}$.

We can rewrite this inequality as
\[
\frac{1}{|S|} \sum_{(z,v) \in S} \text{Re}\langle w_{z,v} | H^v | z \rangle \geq 1 - t 2^{-\alpha+1}.
\]

Using the Cauchy-Schwarz inequality, we get $|\sum_{z,v} \text{Re}\langle w_{z,v} | H^v | z \rangle|^2 \leq |S| \cdot \sum_{z,v} (\text{Re}\langle w_{z,v} | H^v | z \rangle)^2$. It follows that
\[
\frac{1}{|S|} \sum_{(z,v) \in S} |\langle w_{z,v} | H^v | z \rangle|^2 \geq \frac{1}{|S|} \cdot \left( \frac{1}{|S|} \sum_{(z,v) \in S} \text{Re}\langle w_{z,v} | H^v | z \rangle \right)^2
\]
\[
\geq (1 - t 2^{-\alpha+1})^2.
\]
(4.12)
The attack on the locking scheme is defined by the projective measurement of the orthonormal set \( \{ |w_z, v \rangle \}_{(z,v) \in S} \). Note that this is a valid attack because this set does not depend on the private randomness \( r \) and only uses the description of the protocol (the codes \( C_k \) and the bitstrings \( v_k \)). Let \( I \) be the outcome of the measurement (only the \( z \) part, so more precisely we perform the projective measurement whose elements are \( \sum_{v'} |w_z, v \rangle |H^v|z\rangle \).

We have

\[
\Pr \{ I = Z \} = \mathbb{E}_{x,r,k} \left\{ \sum_{v':(C_k(P_k(x)),v',h)} \right. \\
\left. \Pr \{ Z = z, V = v \} \sum_{v':(z,v')} \right| \langle w_z,v \rangle |H^v|z\rangle |^2 \\
\geq \sum_{z,v} \Pr \{ Z = z, V = v \} |\langle w_z,v \rangle |H^v|z\rangle |^2 \\
= \frac{1}{|S|} \sum_{(z,v) \in S} |\langle w_z,v \rangle |H^v|z\rangle |^2 \\
\geq (1 - t^{2-\alpha+1})^2.
\]

Here, we used the assumption that \( Z,V \) is uniformly distributed on its support so that \( \Pr \{ Z = z, V = v \} = 1/|S| \). This condition is satisfied for example by the scheme of Corollary 4.1.7. It follows that \( \frac{1}{2} \| p_{XI} - p_{XZ} \|_1 \leq 2t^{2-\alpha+1} \).

\[
\mathbf{I}(I; X) = \mathbf{H}(X) - \mathbf{H}(X|I) \\
\geq \mathbf{H}(X) - \mathbf{H}(X|Z) - 4 \cdot 4t^{2-\alpha+1}n - 2h_2(4t^{2-\alpha+1}) \\
\geq n - 16t2^{-\alpha+1}n - (\mathbf{H}(X|ZK) + \mathbf{H}(K)) - 2h_2(4t^{2-\alpha+1}) \\
= n - 16t2^{-\alpha+1}n - \log t - 2h_2(4t^{2-\alpha+1}).
\]

In the first inequality, we used the Alicki-Fannes inequality (Lemma 2.2.5). For the second inequality, we used the fact that \( \mathbf{H}(X|Z) \leq \mathbf{H}(X|ZK) + \mathbf{H}(K) \) and for the last equality, we used the fact that given \( Z \) and \( K \), we can decode \( X \) so that \( \mathbf{H}(X|ZK) = 0 \).

This result says that if we want to build a locking scheme with a small key that can be implemented with current technology and that tolerates some errors, we should look for schemes that do not lie in the class described above. However, if a key of size \( cn \) for some constant \( c < 1 \) is acceptable (where \( n \) is the size of the message), it is possible to construct a locking scheme that is tolerant to errors. In fact, we could for instance after applying the permutation to the \( n \)-bit message together with the private random string obtaining a bit string \( z \) of length roughly \( 4n \), compute some parities of \( z \). Let \( y \) denote the string of parities.
obtained. As in the first protocol of Corollary 4.1.7, Alice then encodes \( z \) by performing some Hadamard gates according to the key and sends this ciphertext to Bob. In addition, Alice uses another part of the key to encrypt \( y \) (using a one-time pad) and sends it to Bob. Using the key, Bob can recover \( z' \) a noisy version of \( z \) and can recover \( y \) perfectly (as it is sent through a classical channel). Using \( z' \) and \( y \), Bob can recover \( z \) provided \( z' \) didn’t have too many errors.

To obtain a smaller key size, one idea is to use the method described above as a key expansion procedure and repeat it many times. Typically the number of times we would like to expand the key is \( O(\log n) \), and it is thus possible to choose independently the permutation and Hadamard that are be applied at each step. But it seems difficult to analyse how these protocols compose. It is not clear what kind of information can be obtained from a measurement that acts on the big protocol.

### 4.1.5 Quantum hiding fingerprints

In this section, we show that the locking scheme of Corollary 4.1.5 can be used to build mixed state quantum hiding fingerprints as defined by Gavinsky and Ito [2010]. A quantum fingerprint encodes an \( n \)-bit string into a quantum state \( \rho_x \) of \( n' \ll n \) qubits such that given \( y \in \{0, 1\}^n \) and the fingerprint \( \rho_x \), it is possible to decide with small error probability whether \( x = y \) [Buhrman et al., 2001]. The additional hiding property ensures that measuring \( \rho_x \) leaks very little information about \( x \). Gavinsky and Ito [2010] used the accessible information\(^1\) as a measure of the hiding property. Here, we strengthen this definition by imposing a bound on the total variation distance instead (see Proposition 3.2.2).

**Definition 4.1.11 (Quantum hiding fingerprint).** Let \( n \) be a positive integer, \( \delta, \epsilon \in (0, 1) \) and \( C \) be a Hilbert space. An encoding \( f : \{0, 1\}^n \to S(C) \) together with a set of measurements \( \{M^x, \text{id} - M^y\} \) for each \( y \in \{0, 1\}^n \) is a \((\delta, \epsilon)\)-hiding fingerprint if

1. (Fingerprint property) For all \( x \in \{0, 1\}^n \), \( \text{tr} [M^x f(x)] = 1 \) and for \( y \neq x \), \( \text{tr} [M^y f(x)] \leq \delta \).

2. (Hiding property) Let \( X \) be uniformly distributed on \( \{0, 1\}^n \). Then, for any POVM \( \{N_i\} \) on the system \( C \) whose outcome on \( f(X) \) is denoted \( I \), we have for all possible outcomes \( i \),

\[
\Delta(p_{X|I=i}, p_X) \leq \epsilon.
\]

\(^1\)The accessible information about \( X \) in a quantum system \( C \) refers to the maximum over all measurements of the system \( C \) of \( I(X; I) \) where \( I \) is the outcome of that measurement.
We usually want the Hilbert space $C$ to be composed of $O(\log n)$ qubits. Gavinsky and Ito [2010] proved that for any constant $c$, there exist efficient quantum hiding fingerprinting schemes for which the number of the qubits in system $C$ is $O(\log n)$ and both the error probability $\delta$ and the accessible information are bounded by $1/n^c$. Here, we prove that the same result can be obtained by locking a classical fingerprint. The general structure of our quantum hiding fingerprint for parameters $n, \delta$ and $\epsilon$ is as follows:

1. Choose a random prime $p \in \mathcal{P}_{n,\epsilon}$ uniformly from the set $\mathcal{P}_{n,\epsilon}$. 
2. Set $t = \lceil c \log(1/\epsilon) \rceil$ and generate $t$ random unitaries $U^p_1, \ldots, U^p_t$ acting on $A \otimes B$.

3. The fingerprint consists of the random prime $p$ and the state $(U^p_k)^\dagger |x \mod p\rangle^A |b\rangle^B$ where $k \in [t]$ and $b \in [d_B]$ are chosen uniformly and independently. The density operator representing this state is denoted $f(x) = \sum_{k,b} (U^p_k)^\dagger |x \mod p\rangle |x \mod p\rangle^A \otimes |b\rangle^B U^p_k$.

Observe that even though this protocol is randomized because the unitaries are chosen at random, it is possible to implement it with resources polynomial in $n$ as the size of the message to be locked is $O(\log n)$ bits. In fact, one can approximately sample a random unitary in dimension $2^{O(\log n)}$ using a polynomial number of public random bits. The mixed state protocol of Gavinsky and Ito [2010] achieves roughly the same parameters. Their construction is also randomized but it uses random codes instead of random unitaries. For this reason, the protocol of Gavinsky and Ito [2010] would probably be more efficient in practice.

**Theorem 4.1.12.** There exist constants $c, c'$ and $c''$, such that for all positive integer $n$, $\delta, \epsilon \in (0, 1/4)$ if we define $\mathcal{P}_{n,\delta,\epsilon}$ to be the set of primes in the interval $[l, u]$ where

$$l = \left( \frac{c''}{\delta} \cdot \frac{\log^2(1/\epsilon)}{e^8} \right)^{1/0.9} + 10n \quad \text{and} \quad u = l + (2n/\delta)^2$$

and provided $u \leq 2n^{2-2}$, the scheme described above is a $(\delta, \epsilon)$-hiding fingerprint with probability $1 - 2^{-\Omega(n)}$ over the choice of random unitaries.

The proof of this result involves two parts. First, we need to show that the fingerprint of a uniformly distributed $X \in \{0, 1\}^n$ does not give away much information about $X$. This follows easily from Theorem 3.3.2 and Theorem 4.1.4. We also need to show that for every $y \in \{0, 1\}^n$, there is a measurement that Bob can apply to the fingerprint to determine with
high confidence whether it corresponds to a fingerprint of $y$ or not. In order to prove this we use Lemma 4.1.9 that gives a way of approximating a set of almost orthogonal vectors by a set of orthogonal vectors.

**Lemma 4.1.13.** Let $\{U_1, \ldots, U_t\}$ be a set of unitary transformations on $AB$ that define $\gamma$-MUBs and $d \doteq d_A d_B$. Define for $y \in [d_A]$ the subspace $F_y = \text{span}\{U_k^\dagger|y\rangle|b\rangle, k \in [t], b \in [d_B]\}$. Then for any $x \in [d_A], y \neq x, k_0 \in [t]$ and $b_0 \in [d_B], \quad \text{tr} \left[ \Pi_{F_y} U_{k_0}^\dagger x\rangle\langle b_0| \right] \leq 3(td_B)^2d^{-\gamma},$

where $\Pi_F$ is the projector on the subspace $F$.

**Proof** Consider the set of vectors $\{U_k^\dagger|y\rangle|b\rangle\}_{k \in [t], b \in [d_B]}$. We have for all $(k, b) \neq (k', b'),$

$$|\langle y|\langle b'|U_k U_k^\dagger|y\rangle|b\rangle| \leq d^{-\gamma/2},$$

and as a result,

$$\frac{1}{td_B} \sum_{(k, b) \neq (k', b')} |\langle y|\langle b'|U_k U_k^\dagger|y\rangle|b\rangle|^2 \leq td_B d^{-\gamma}.$$

Using Lemma 4.1.9 we obtain a set of orthonormal vectors $\{|e_{k, b}(y)\rangle\}_{k, b}$ such that

$$\frac{1}{td_B} \sum_{k, b} \|\|e_{k, b}(y)\rangle - U_k^\dagger|y\rangle|b\rangle\|^2 \leq td_B d^{-\gamma}.$$

Note that $\{|e_{k, b}(y)\rangle\}_{k, b}$ is an orthonormal basis for $F_y$ so we can write $\Pi_{F_y} = \sum_{k, b} |e_{k, b}(y)\rangle\langle e_{k, b}(y)|$. Now observe that, using the Cauchy Schwarz inequality and the fact that the vectors have unit norm, we have $|\langle e_{k, b}(y)|U_{k_0}^\dagger|x\rangle|b_0\rangle| \leq |\langle y|b\rangle U_k U_k^\dagger|y\rangle|b_0\rangle| + \|\|e_{k, b}(y)\rangle - U_k^\dagger|y\rangle|b\rangle\|_2$. As a result, we have

$$\text{tr} \left[ \Pi_{F_y} U_{k_0}^\dagger x\rangle\langle b_0| \right] = \sum_{k, b} |\langle e_{k, b}(y)|U_{k_0}^\dagger|x\rangle|b_0\rangle|^2 \leq \sum_{k, b} |\langle y|\langle b|U_k U_k^\dagger|y\rangle|b\rangle| + \|\|e_{k, b}(y)\rangle - U_k^\dagger|y\rangle|b\rangle\|_2|^2 \leq t d_B d^{-\gamma} + (td_B)^2 d^{-\gamma} + \sum_{k, b} |\langle y|b\rangle U_k U_k^\dagger|y\rangle|b_0\rangle| \cdot \|\|e_{k, b}(y)\rangle - U_k^\dagger|y\rangle|b\rangle\|_2 \leq 2(td_B)^2 d^{-\gamma} + d^{-\gamma/2} \sqrt{td_B} \sqrt{\sum_{k, b} \|\|e_{k, b}(y)\rangle - U_k^\dagger|y\rangle|b\rangle\|_2^2} \leq 3(td_B)^2 d^{-\gamma}.$$

Proof [of Theorem 4.1.12] We start by proving the hiding property. For any fixed $p$, the random variable $Z \overset{\text{def}}{=} X \mod p$ is almost uniformly distributed on $\{0, \ldots, p-1\}$. In fact, we have for any $z \in \{0, \ldots, p-1\}$, $\Pr \{Z = z\} \leq \frac{2^n/p+1}{2^n}$. In other words, $H_{\text{min}}(Z) \geq \log p - \log(1 + p2^{-n})$. Thus, using Theorem 3.3.2 and Theorem 4.1.4, we have that except with probability exponentially small in $n$ (on the choice of the random unitary), the fingerprinting scheme satisfies for any measurement outcome $i$

$$\Delta(p_Z|I=i, p_Z) \leq \frac{2\epsilon}{1+p2^{-n} - \epsilon} \leq 4\epsilon$$

where $I$ denotes the outcome of a measurement on the state $f(X)$. Recall that we are interested in the information leakage about $X$ not $Z$. For this, we note that the random variables $X, Z, I$ form a Markov chain. Thus,

$$\Delta(p_X|I=i, p_X) \leq \frac{2\epsilon}{1+p2^{-n} - \epsilon} \leq 4\epsilon.$$

This proves the hiding property.

We now analyse the fingerprint property. Let $x, y \in [2^n]$ and $p$ be the random prime of the fingerprint. We define the measurements by $M^y = \Pi_{F_y}$ for all $y \in \{0, 1\}^n$ where $\Pi_{F_y}$ is the projector onto the subspace $F_y = \text{span}\{U_k^{p\dagger}|y \mod p\}|b\rangle, k \in [t], b \in [d_B]\}$. If $x = y$, then $f(x)$ is a mixture of states in $\text{span}\{U_k^{p\dagger}|y \mod p\}|b\rangle, k \in [t], b \in [d_B]\}$. Thus $\text{tr}[M^y f(x)] = 1$.

We now suppose that $x \neq y$. First, we have for a random choice of prime $p \in \mathcal{P}_{n,\epsilon,\delta}$, $\Pr \{x \mod p = y \mod p\} = \Pr \{x - y \mod p = 0\} \leq \delta/2$ as the number of distinct prime divisors of $x - y$ is at most $n$ and the number of primes in $[l, u]$ is at least $2n/\delta$ for $n$ large enough. Then, whenever $x \mod p \neq y \mod p$, Lemma 4.1.13 with $\gamma = 0.9$
4.1. Locking classical information in quantum states

4.1.6 String commitment

In this section, we show how to use a locking scheme to obtain a weak form of bit commitment [Buhrman et al., 2006]. Bit commitment is an important two-party cryptographic primitive defined as follows. Consider two mutually distrustful parties Alice and Bob who are only allowed to communicate over some channel. The objective is to be able to achieve the following: Alice secretly chooses a bit $x$ and communicates with Bob to convince him that she fixed her choice, without revealing the actual bit $x$. This is the commit stage. At the reveal stage, Alice reveals the secret $x$ and enables Bob to open the commitment. Bob can then check whether Alice was honest.

Using classical or quantum communication, unconditionally secure bit commitment is known to be impossible [Lo and Chau, 1997, Mayers, 1997]. However, commitment protocols with weaker security guarantees do exist [Buhrman et al., 2006, 2008, Damgård et al., 2005, Spekkens and Rudolph, 2001]. Here, we consider the string commitment scenario studied in [Buhrman et al., 2008, Section III]. In a string commitment protocol, Alice commits to an $n$-bit string. Alice’s ability to cheat is quantified by the number of strings she can reveal successfully. The ability of Bob to cheat is quantified by the information he can obtain about the string to be committed. One can formalize these notions in many ways. We use a security criterion that is similar to the one of [Buhrman et al., 2008] except that we use the statistical distance between the outcome distribution and the uniform distribution, instead of the accessible information. Our definition is slightly stronger by virtue of Proposition 4.1.3. For a detailed study of string commitment in a more general setting, see [Buhrman et al., 2008].

**Definition 4.1.14.** An $(n, \alpha, \beta)$-quantum bit string commitment is a quantum communication protocol between Alice (the committer) and Bob (the receiver) which has two phases.
CHAPTER 4. Uncertainty relations: Applications

When both players are honest the protocol takes the following form.

- **(Commit phase)** Alice chooses a string $X \in \{0, 1\}^n$ uniformly. Alice and Bob communicate, after which Bob holds a state $\rho_X$.

- **(Reveal phase)** Alice and Bob communicate and Bob learns $X$.

The parameters $\alpha$ and $\beta$ are security parameters.

- If Alice is honest, then for any measurement performed by Bob on her state $\rho_X$, we have $\Delta(p_X, p_X|I=i) \leq \frac{\beta}{n}$ where $I$ is the outcome of the measurement.

- If Bob is honest, then for all commitments of Alice: $\sum_{x \in \{0, 1\}^n} p_x \leq 2^\alpha$, where $p_x$ is the probability that Alice successfully reveals $x$.

Following the strategy of Buhrman et al. [2008], the following protocol for string commitment can be defined using a locking scheme $\mathcal{E}$.

- **Commit phase**: Alice has the string $X \in \{0, 1\}^n$ and chooses a key $K \in [t]$ uniformly at random. She sends the state $\mathcal{E}(X, K)$ to Bob.

- **Reveal phase**: Alice announces both the string $X$ and the key $K$. Using the key, Bob decodes some value $X'$. He accepts if $X = X'$.

A protocol is said to be efficient if both the communication (in terms of the number of qubits exchanged) is polynomial in $n$ and the computations performed by Alice and Bob can be done in polynomial time on a quantum computer. The protocol presented in Buhrman et al. [2008] is not efficient in terms of computation and is efficient in terms of communication only if the cost of communicating a (random) unitary in dimension $2^n$ is disregarded. Using the efficient locking scheme of Corollary 4.1.7, we get

**Corollary 4.1.15.** Let $n$ be a positive integer and $\beta \in (n2^{-cn}, n)$ ($c$ is a constant independent of $n$). There exists an efficient $(n, c\log(n^2/\beta), \beta)$-quantum bit string commitment protocol for some constant $c$ independent of $n$ and $\beta$.

**Proof** We use the first construction of Corollary 4.1.7 with $\epsilon = \beta/n$. If Bob is honest, the security analysis is exactly the same as in Buhrman et al. [2008]. If Alice is honest, the security follows directly from the definition of the locking scheme. \qed
4.1. Locking classical information in quantum states

4.1.7 Locking entanglement of formation

The entanglement of formation is a measure of the entanglement in a bipartite quantum state that attempts to quantify the number of singlets required to produce the state in question using only local operations and classical communication [Bennett et al., 1996]. For a bipartite state \( \rho^{XY} \), the entanglement of formation is defined as

\[
E_f(X; Y)_{\rho} = \min_{\{p_i, |\psi_i\rangle\}} \sum_i p_i H(X)_{\psi_i}.
\]

where the minimization is taken over all possible ways to write \( \rho^{XY} = \sum_i p_i |\psi_i\rangle\langle\psi_i| \) with \( \sum_i p_i = 1 \). Entanglement of formation is related to the following quantity:

\[
I^{\leftarrow}(X; Y')_{\rho} = \max_{\{M_i\}} I(X; I)
\]

where the maximization is taken over all measurements \( \{M_i\} \) performed on the system \( Y' \) and \( I \) is the outcome of this measurement. This quantity is sometimes referred to as a classical correlations between \( X \) and \( Y \) [Henderson and Vedral, 2001]. As mentioned previously, when the system \( X \) is classical, this correlation measure is called accessible information. Koashi and Winter [2004] showed that for a pure state \( |\rho\rangle^{XYY'} \), a simple identity holds:

\[
E_f(X; Y)_{\rho} + I^{\leftarrow}(X; Y')_{\rho} = H(X)_{\rho}.
\]

Let \( \{U_1, \ldots, U_t\} \) be a set of unitary transformations of \( A \otimes B \simeq C \) and define

\[
|\rho\rangle^{ABC'AK} = \frac{1}{\sqrt{d_A d_B}} \sum_{k \in [t], a \in [d_A], b \in [d_B]} |a\rangle^A |b\rangle^B \left(U_k^\dagger |a\rangle \otimes |b\rangle \right)^C |a\rangle^{A'} |k\rangle^K.
\]

If \( \{U_1, \ldots, U_t\} \) satisfies an \( \epsilon \)-metric uncertainty relation, then we get a locking effect using Theorem 4.1.4 and Proposition 4.1.3. In fact, we have \( I^{\leftarrow}(A; C)_{\rho} \leq 8\epsilon \log d_A + 2h_2(2\epsilon) \) and \( I^{\leftarrow}(A; CK) = \log d_A \). Thus, using (4.14), we get

\[
E_f(A; A'BK)_{\rho} = H(A)_{\rho} - I^{\leftarrow}(A; C)_{\rho} \geq (1 - 8\epsilon) \log d_A - 2h_2(2\epsilon)
\]

and discarding the system \( K \) of dimension \( t \) we obtain a separable state

\[
E_f(A; A'B)_{\rho} = 0.
\]

Explicit states exhibiting weak locking behaviour of the entanglement of formation have been presented in Horodecki et al. [2005a]. Strong but non-explicit instances of locking the entanglement of formation were derived in Hayden et al. [2006]. Here, using Theorem 3.4.6, we obtain explicit examples of strong locking behaviour.
One could also consider other quantities related to classical correlations, such as the popular quantum discord [Ollivier and Zurek, 2001], and they would exhibit a similar locking behaviour.

### 4.2 Quantum identification codes

Consider the following quantum analogue of the equality testing communication problem. Alice is given an $n$-qubit state $|\psi\rangle \in C$ and Bob is given $|\varphi\rangle \in C$. Namely, Bob wants to output 1 with probability in the interval $[|\langle \psi | \varphi \rangle|^2 - \epsilon, |\langle \psi | \varphi \rangle|^2 + \epsilon]$ and 0 with probability in the interval $[1 - |\langle \psi | \varphi \rangle|^2 - \epsilon, 1 - |\langle \psi | \varphi \rangle|^2 + \epsilon]$. This task is referred to as quantum identification [Winter, 2004]. Note that communication only goes from Alice to Bob. There are many possible variations to this problem. One of the interesting models is when Alice receives the quantum state $|\psi\rangle$ and Bob gets a classical description of $|\varphi\rangle$. An $\epsilon$-quantum-ID code is defined by an encoder, which is a quantum operation that maps Alice’s quantum state $|\psi\rangle$ to another quantum state which is transmitted to Bob, and a family of decoding POVMs $\{D_\varphi, \text{id} - D_\varphi\}$ for all $|\varphi\rangle$ that Bob performs on the state he receives from Alice.

**Definition 4.2.1 (Quantum identification [Winter, 2004]).** Let $\mathcal{H}_1, \mathcal{H}_2, C$ be Hilbert spaces and $\epsilon \in (0, 1)$. An $\epsilon$-quantum-ID code for the space $C$ using the channel $\mathcal{N} : \mathcal{S}(\mathcal{H}_1) \rightarrow \mathcal{S}(\mathcal{H}_2)$ consists of an encoding map $\mathcal{E} : \mathcal{S}(C) \rightarrow \mathcal{S}(\mathcal{H}_1)$ and a set of POVMs $\{D_\varphi, \text{id} - D_\varphi\}$ acting on $\mathcal{S}(\mathcal{H}_2)$, one for each pure state $|\varphi\rangle$ such that

$$\forall |\psi\rangle, |\varphi\rangle \in C, \quad \left| \text{tr} [D_\varphi \mathcal{N}(\mathcal{E}(\psi))] - |\langle \varphi | \psi \rangle|^2 \right| \leq \epsilon.$$ 

Here we consider channels $\mathcal{N}$ transmitting noiseless qubits and noiseless classical bits. We also say that $\epsilon$-quantum identification of $n$-qubit states can be performed using $\ell$ bits and $m$ qubits when there exists an $\epsilon$-quantum-ID code for the space $C = (C^2)^\otimes n$ using the channel $\mathcal{N}' = \text{id}_2^\otimes \ell \otimes \text{id}_2^\otimes m$, where $\text{id}_2$ and $\text{id}_2$ are the noiseless bit and qubit channels. Hayden and Winter [2012] showed that classical communication alone cannot be used for quantum identification. However, a small amount of quantum communication makes classical communication useful. Using our metric uncertainty relations, we prove better bounds on the number of qubits of communication and give an efficient encoder for this problem.

Our protocol is based on a duality between quantum identification and approximate forgetfulness of a quantum channel demonstrated in [Hayden and Winter, 2012, Theorem 7]. Specialized to our setting, the direction of the duality we use states that if $V : C \rightarrow A \otimes B$ defines a low-distortion embedding of $(C, \ell_2)$ into $(AB, \ell_1^A(\ell_2^B))$, then the maps $\Gamma_a : C \rightarrow B$...
for $a \in [d_A]$ defined by $|\psi\rangle \mapsto \sum_{b \in [d_A]} (\langle a|b|V|\psi\rangle |b\rangle$ approximately preserve inner products on average. The following lemma gives a precise statement. We give an elementary proof in the interest of making the presentation self-contained.

**Lemma 4.2.2.** Let $V : C \to A \otimes B$ be an isometry, i.e., for all $|\psi\rangle \in C$, $\|V|\psi\rangle\|_2 = \||\psi\rangle\|_2$. For any vector $|\psi\rangle \in C$, we define the vectors $|\psi_a\rangle \in B$ by $V|\psi\rangle = \sum_{a \in [d_A]} |a\rangle |\psi_a\rangle$. Assume that $V$ satisfies the following property:

$$\forall |\psi\rangle \in C \quad \sum_{a \in [d_A]} \left(\|\psi_a\|_2^2 - \frac{\||\psi\rangle\|_2^2}{d_A}\right) \leq \epsilon \|\psi\|_2^2.$$  \hspace{1cm} \text{(4.15)}

Then we have for all unit vectors $|\psi\rangle, |\varphi\rangle \in C$ with $V|\psi\rangle = \sum_{a \in [d_A]} |a\rangle |\psi_a\rangle$ and $V|\varphi\rangle = \sum_{a \in [d_A]} |a\rangle |\varphi_a\rangle$

$$\leq \sum_{a \in [d_A]} \left(\frac{\|\psi_a\|_2^2}{d_A} - \|\psi\|_2 \|\varphi_a\|_2\right) + \sum_{a \in [d_A]} \frac{|\langle \psi_a | \varphi_a \rangle|^2}{d_A \|\psi_a\|_2 \|\varphi_a\|_2} \leq 12\epsilon + 2\epsilon. \hspace{1cm} \text{(4.16)}$$

**Proof.** Let $|\psi\rangle$ and $|\varphi\rangle$ be unit vectors in $C$. We use the triangle inequality to get

$$\leq \sum_{a \in [d_A]} \left|\frac{|\langle \psi_a | \varphi_a \rangle|^2}{d_A} - \|\psi\|_2 \|\varphi_a\|_2\right| + \sum_{a \in [d_A]} \left|\frac{|\langle \psi_a | \varphi_a \rangle|^2}{d_A \|\psi_a\|_2 \|\varphi_a\|_2} \right|.$$  \hspace{1cm} \text{(4.17)}

We start by dealing with the first term in (4.17). Observe that

$$\left|\frac{|\langle \psi_a | \varphi_a \rangle|^2}{d_A} - \|\psi\|_2 \|\varphi_a\|_2\right| \leq \left|\text{Re}(\langle \psi_a | \varphi_a \rangle)^2 - \frac{\text{Re}(\langle \psi | \varphi \rangle)^2}{d_A}\right| + \left|\text{Im}(\langle \psi_a | \varphi_a \rangle)^2 - \frac{\text{Im}(\langle \psi | \varphi \rangle)^2}{d_A}\right| \leq 2 \left|\text{Re}(\langle \psi_a | \varphi_a \rangle) - \frac{\text{Re}(\langle \psi | \varphi \rangle)}{d_A}\right| + 2 \left|\text{Im}(\langle \psi_a | \varphi_a \rangle) - \frac{\text{Im}(\langle \psi | \varphi \rangle)}{d_A}\right|. \hspace{1cm} \text{(4.18)}$$

In the last inequality, we used the fact that $|x^2 - y^2| \leq 2|x - y|$ whenever $|x + y| \leq 2$. To bound these terms, we apply the assumption about $V$ (equation (4.15)) to the vector $|\psi\rangle - |\varphi\rangle$:

$$\sum_{a \in [d_A]} \left|\|\psi_a\|_2^2 - \frac{|\langle \psi_a | \varphi_a \rangle|^2}{d_A}\right| \leq \epsilon \|\psi\|_2^2 - \|\varphi\|_2^2 \leq 4\epsilon.$$
By expanding \( \| \psi_a - \varphi_a \|_2^2 \) and \( \| \psi - \varphi \|_2^2 \), we obtain using the triangle inequality

\[
\sum_{a \in [d_A]} \left| 2 \text{Re} \langle \psi_a | \varphi_a \rangle - \frac{2 \text{Re} \langle \psi | \varphi \rangle}{d_A} \right| \leq 4 \epsilon + \sum_{a \in [d_A]} \left( \| \psi_a \|_2 \| \varphi_a \|_2 - \frac{\| \psi \|_2^2}{d_A} \right) + \sum_{a \in [d_A]} \left( \| \varphi_a \|_2 - \frac{\| \varphi \|_2^2}{d_A} \right) \leq 6 \epsilon.
\]

In the last inequality, we used equation (4.15) for \( |\psi \rangle \) and \( |\varphi \rangle \). The same argument can be applied to \( i |\psi \rangle \) and \( |\varphi \rangle \) to get

\[
2 \sum_{a \in [d_A]} \left| \text{Im} \langle \psi_a | \varphi_a \rangle - \frac{\text{Im} \langle \psi | \varphi \rangle}{d_A} \right| \leq 6 \epsilon.
\]

Thus, substituting in equation (4.18) we obtain

\[
\left| \langle \psi_a | \varphi_a \rangle^2 - \frac{|\langle \psi | \varphi \rangle|^2}{d_A} \right| \leq 12 \epsilon.
\]

We now consider the second term in (4.17). We have, using the Cauchy-Schwarz inequality,

\[
\sum_{a \in [d_A]} \left| \langle \psi_a | \varphi_a \rangle \right|^2 - \frac{\left| \langle \psi_a | \varphi_a \rangle \right|^2}{d_A} \leq \sum_{a \in [d_A]} \left( \| \psi_a \|_2 \| \varphi_a \|_2 - \frac{1}{d_A} \right) \leq \sum_{a \in [d_A]} \left( \| \psi_a \|_2 \| \varphi_a \|_2 - \frac{1}{\sqrt{d_A}} \right) + \sum_{a \in [d_A]} \left( \| \psi_a \|_2 - \frac{1}{\sqrt{d_A}} \right) \leq \sum_{a \in [d_A]} \left( \| \psi_a \|_2 \right)^2 - \frac{1}{\sqrt{d_A}} \leq 2 \sqrt{\epsilon}.
\]

For the third inequality, we used once again the Cauchy-Schwarz inequality and for the fourth inequality, we used the fact that \( \sum_{a \in [d_A]} \| \psi_a \|_2^2 = \| V \psi \|_2^2 = 1 \) and the inequality \( |x - y|^2 \leq |x - y| |x + y| = |x^2 - y^2| \) for all nonnegative \( x, y \). Plugging this bound into equation (4.17), we obtain the desired result.
Theorem 4.2.3 (Quantum identification using classical communication). Let \( n \) be a positive integer and \( \epsilon \in (2^{-c'n}, 1) \) where \( c' \) is a constant independent of \( n \). Then for some \( m = O(\log(1/\epsilon)) \), \( \epsilon \)-quantum identification of \( n \)-qubit states can be performed using a single message of \( n \) bits and \( m \) qubits.

Moreover, for some \( m = O(\log(n/\epsilon) \cdot \log(n)) \), \( \epsilon \)-quantum identification of \( n \)-qubit states can be performed using a single message of \( n \) bits and \( m \) qubits with an encoding quantum circuit of polynomial size.

Proof Let \( \{U_1, \ldots, U_t\} \) be a set of unitaries on \( n \) qubits verifying an \( \epsilon' \)-metric uncertainty relation with \( \epsilon' = 1/2 \cdot (\epsilon/28)^2 \). We start by preparing the uniform superposition state \( \frac{1}{\sqrt{t}} \sum_k |k\rangle^K \) and apply the unitary \( U_k \) on system \( C = A \otimes B \) controlled by the register \( K \). We get the state \( \frac{1}{\sqrt{t}} \sum_k |k\rangle^K (U_k |\psi\rangle)^{AB} = \sum_{k,a} |k\rangle^K |a\rangle^A |\psi_{k,a}\rangle^B \) for some non-normalized vectors \( |\psi_{k,a}\rangle \in B \). Alice then measures the system \( KA \) in the computational basis obtaining an outcome \( k, a \) and sends \( k, a \) and \( |\hat{\psi}_{k,a}\rangle \) to Bob, where \( |\hat{\psi}_{k,a}\rangle = |\psi_{k,a}\rangle/\| |\psi_{k,a}\rangle \|_2 \). Observe that \( \sum_{k,a} \| |\psi_{k,a}\rangle \|_2^2 = 1 \) and \( \| |\psi_{k,a}\rangle \|_2^2 = \frac{1}{t} \cdot p_{U_k |\psi}(a) \) so that the metric uncertainty relation property can be written as

\[
\frac{1}{2} \sum_{k,a} \left( \| |\psi_{k,a}\rangle \|_2^2 - \frac{1}{t d_A} \right) \leq \epsilon'.
\] (4.19)
This shows that the isometry $|\psi\rangle \rightarrow \frac{1}{\sqrt{t}} \sum_k |k\rangle^K (U_k|\psi\rangle)^{AB}$ satisfies the condition (4.15) of Lemma 4.2.2.

The decoding POVMs for received classical information $k, a$ and state $|\varphi\rangle$ are defined by $D^k_{\varphi,a} = |\hat{\varphi}_{k,a}\rangle\langle \hat{\varphi}_{k,a}|$ where $\frac{1}{\sqrt{t}} \sum_k |k\rangle^K (U_k|\varphi\rangle)^{AB} = \sum_{k,a} |k\rangle^K |a\rangle^A |\varphi_{k,a}\rangle^B$ and $|\hat{\varphi}_{k,a}\rangle = |\varphi_{k,a}\rangle / \|\varphi_{k,a}\|_2$. The protocol is illustrated in Figure 4.2.

We now analyse the probability that Bob outputs 1. Recall that outcome 1 corresponds to the projector $|\varphi\rangle\langle \varphi|$. The probability that the protocol in Figure 4.2 outputs 1 is

$$\sum_{k,a} \|\psi_{k,a}\|_2^2 \cdot \text{tr} \left[ D_{\varphi,a}^k |\hat{\psi}_{k,a}\rangle\langle \hat{\psi}_{k,a}| \right] = \sum_{k,a} \|\psi_{k,a}\|_2^2 \langle \hat{\psi}_{k,a} | \hat{\varphi}_{k,a} \rangle^2.$$

Applying Lemma 4.2.2, we have

$$\frac{1}{td_A} \sum_{k,a} \left| \langle \hat{\psi}_{k,a} | \hat{\varphi}_{k,a} \rangle \right|^2 - \left| \langle \psi | \varphi \rangle \right|^2 \leq 14 \sqrt{2}\epsilon' = \epsilon/2. \quad (4.20)$$

Using the triangle inequality, equations (4.20) and (4.19), we obtain

$$\sum_{k,a} \|\psi_{k,a}\|_2^2 \left| \langle \hat{\psi}_{k,a} | \hat{\varphi}_{k,a} \rangle \right|^2 - \left| \langle \psi | \varphi \rangle \right|^2 \leq \sum_{k,a} \left| \langle \hat{\psi}_{k,a} | \hat{\varphi}_{k,a} \rangle \right|^2 - \left| \langle \psi | \varphi \rangle \right|^2 + \sum_{k,a} \|\psi_{k,a}\|_2^2 - \frac{1}{td_A} \cdot 2$$

$$\leq \epsilon/2 + 4\epsilon' \leq \epsilon.$$

Thus, the probability of obtaining outcome 1 is in the interval $[\left| \langle \psi | \varphi \rangle \right|^2 - \epsilon, \left| \langle \psi | \varphi \rangle \right|^2 + \epsilon]$.

We conclude by using the metric uncertainty relations of Theorems 3.3.2 and 3.4.7. For the explicit construction, we still need to argue that the encoding can be computed by a quantum circuit of size $O(n^2 \text{polylog}(n/\epsilon))$ and depth $O(n \text{polylog}(n/\epsilon))$ using classical precomputations. To obtain this running time, we actually use the 1-MUBs of Lemma 3.4.1 in the construction of Theorem 3.4.7. The only thing we need to precompute is an irreducible polynomial of degree $n$ over $\mathbb{F}_2[X]$. Then, using the same argument as in the proof of Lemma 3.4.1, we can compute the unitary operation that takes as input the state $|j\rangle \otimes |\psi\rangle$ and outputs the state $|j\rangle \otimes V_j|\psi\rangle$ using a circuit of size $O(n^2 \text{polylog} n)$ and depth $O(n \text{polylog} n)$. Since the permutation extractor we use can be implemented by a quantum circuit of size $O(n \text{polylog}(n/\epsilon))$, the unitary transformation $|k\rangle \otimes |\psi\rangle \rightarrow |k\rangle \otimes U_k|\psi\rangle$ can be computed by a quantum circuit of size $O(n^2 \text{polylog}(n/\epsilon))$ and depth $O(n \text{polylog}(n/\epsilon))$. □

This result can be thought of as an analogue of the well-known fact that the public-coin randomized communication complexity of equality is $O(\log(1/\epsilon))$ for an error probability $\epsilon$. 

\[\square\]
4.2. Quantum identification codes

[Kushilevitz and Nisan, 1997]. Quantum communication replaces classical communication and classical communication replaces public random bits. Classical communication can be thought of as an extra resource because on its own it is useless for quantum identification [Hayden and Winter, 2012, Theorem 11].
Chapter 5

Uncertainty relations in the presence of quantum side information

Outline of the chapter. In the previous chapters, it was assumed that the adversary trying to predict the outcome of the measurement is not entangled with the quantum system being measured. In Section 5.1, we explain what it means for an uncertainty relation to hold when the adversary has quantum side information. After that, in Section 5.2, we introduce metric uncertainty relations with quantum side information that we call QC-extractors. We also give several efficient constructions of QC-extractors. We finally show how using such uncertainty relations, we can relate the security of two-party computations to the quantum capacity of the quantum storage of the adversary (Section 5.4).

5.1 Introduction

Let us consider uncertainty relations in the form of a game, called the uncertainty game by Berta et al. [2010]. Bob prepares a system called $A$ and sends it to Alice. Alice chooses a projective measurement $i$ at random from a set of possible measurements to perform on system $A$. She obtains an outcome that we denote $X$. She then sends $i$ to Bob whose goal is to predict $X$. In Chapter 3, we saw several constructions of measurements for which Bob has a lot of uncertainty about $X$. But in a fully quantum world, Bob might keep a quantum system $E$ that is entangled with $A$ that could help him in predicting $X$. As an example, imagine that Bob prepares the maximally entangled state $|\Phi\rangle = \frac{1}{\sqrt{d_A}} \sum_j |j\rangle_A |j\rangle_E$. Assume the measurements that Alice performs on $A$ are obtained by first applying a unitary transformation $U_i$ on $A$ followed by a measurement in the computational basis. It is simple to see that if Bob, upon receiving the index $i$, applies $U_i^*$ on his system $E$ and performs a
measurement in the computational basis will get the exact same outcome as Alice. Thus, if Alice and Bob share a maximally entangled state, then Bob can perfectly predict the outcome that Alice obtains: there is no uncertainty at all. This makes it clear that the amount of uncertainty depends on the information available for the adversary Bob. In the previous two chapters, we considered the case where Bob prepares a quantum state and sends it completely to Alice, i.e., the $E$ system is not present. In this chapter, we will construct uncertainty relations that hold even when Bob holds a quantum system.

As was discussed above, when the systems $A$ and $E$ are maximally entangled, all the measurement outcomes can be predicted perfectly. Thus, any uncertainty relation should take into account the amount of entanglement between $A$ and the adversary $E$. After being conjectured by Renes and Boileau [2009], it was proven by Berta et al. [2010] that for any state $\rho$ on $AB$, the following inequality holds:

$$\frac{1}{2} (\mathbb{H}(X|E)_{\rho_1} + \mathbb{H}(X|E)_{\rho_2}) \geq \log(1/c) + \frac{1}{2} \mathbb{H}(A|E)_{\rho}$$

(5.1)

where $\rho_{XE} = \mathcal{M}_{A\rightarrow X}(U_i \rho_{AE} U_i^\dagger)$ for $i \in \{1, 2\}$ is the state obtained when measuring the system $A$ of the state $U_i \rho_{AE} U_i^\dagger$ in the computational basis and $c$ is the maximum overlap between the vectors defined by $U_1$ and $U_2$, $c = \max_{a,a'} |\langle a|U_1 U_2^\dagger|a'\rangle|$. $\mathcal{M}_{A\rightarrow X}$ refers to the measurement in the computational basis map: $\mathcal{M}_{A\rightarrow X}(\sigma) = \sum_a \langle a|\sigma|a\rangle |a\rangle \langle a|$. Note that the reason we renamed the system $X$ after the measurement is simply to emphasize that it is a classical system. If the state $\rho_{AE}$ is a pure state on $A$, we have $\mathbb{H}(A|E) = 0$ and recover the uncertainty relation of Maassen and Uffink [1988] in (3.1). In the case where $\rho_{AE}$ is maximally entangled, then $\mathbb{H}(A|E) = -\log d_A$, and $c$ cannot be smaller than $1/\sqrt{d_A}$. This implies that the lower bound in (5.1) is nonpositive, which as discussed earlier is unavoidable. For cryptographic applications, the most interesting case is usually when $\rho_{AE}$ is entangled but not maximally so, i.e., $-\log d_A < \mathbb{H}(A|E) < 0$.

We should mention that quantum side information is usually much harder to handle than classical side information. This is due to the fact that it is not clear how to describe a conditional state. Consider the example of the study of randomness extractors. It is not hard to prove that an extractor can handle any classical adversary as long as it can handle a classical adversary holding a trivial system\textsuperscript{1} see e.g., [König and Terhal, 2008, Proposition 1]. The situation is quite different for quantum adversaries. In fact, Gavinsky et al. [2007] gave an example of an extractor that completely fails when quantum side

\textsuperscript{1}To see this, just write one of the vector of basis 1 in basis 2: one of the squared coefficients has to be at least as large as the average of $1/d_A$.

\textsuperscript{2}Provided the conditional entropy is the same of course.
information is available. This is not to say that quantum adversaries can break any extractor, but that quantum side information can behave in unexpected ways. We now know of many constructions of extractors that do work even when the adversary holds a quantum memory [De et al., 2009, König and Terhal, 2008, Renner and König, 2005, Ta-Shma, 2009, Tomamichel et al., 2011].

As was mentioned in Chapter 3, if we want a larger average measurement entropy, we need to consider more measurements. Unfortunately, up to this day, we only know of uncertainty relations that hold in the presence of quantum memory for two measurements [Berta et al., 2010, Christandl and Winter, 2005, Coles et al., 2011a,b, 2012, Renes and Boileau, 2009, Tomamichel and Renner, 2011]. For two measurements, the incompatibility is directly related to a simple function of the pairwise inner products between vectors in the two bases. For more measurements, controlling the pairwise inner products between the different bases elements is not sufficient to guarantee a good lower bound on the uncertainty [Ballester and Wehner, 2007]. In this chapter, we will give several constructions of strong uncertainty relations for many measurements.

Our strategy will be to follow the idea introduced in Chapter 3 of quantifying the uncertainty in a set of measurement outcomes by the distance to the uniform distribution. In order to account for the possible side information that the adversary has, we also require the output to be independent of the adversary. More precisely, the condition for a set of unitaries $U_1, \ldots, U_t$ will be of the form

$$\frac{1}{t} \sum_{i=1}^{t} \left\| T_{A_1 \rightarrow A_1} (U_i \rho_{AE} U_i^\dagger) - \frac{\text{id}_{A_1}}{d_{A_1}} \otimes \rho_E \right\|_1 \leq \epsilon ,$$

(5.2)

where the map $T$ performs a measurement in the computational basis and then discards the subsystem $A_2$ while keeping $A_1$:

$$T(\cdot)_{A_1 \rightarrow A_1} = \sum_{a_1 a_2} \langle a_1 a_2 | (\cdot) | a_1 a_2 \rangle |a_1 \rangle \langle a_1| ,$$

(5.3)

where $\{|a_1\rangle\}, \{|a_2\rangle\}$ are the computational bases of $A_1, A_2$ respectively. Here, $A_1$ plays the role of the hard to predict outcome called $X$ in the earlier discussion. Equation (5.2) is analogous to Definition 3.2.1 except that we also require that the outcome $A_1$ be decoupled from the adversary $E$. Motivated by the similarity between equation (5.2) and randomness extractors (already introduced in Definition 3.4.4), we call such a set of unitaries a QC-extractor. More details on randomness extractors and related constructions are given in Section 5.2, where we also give constructions of QC-extractors. We will show in Section 5.3 that if $U_1, \ldots, U_t$ satisfy (5.2), they also satisfy an entropic uncertainty relation, as was
done in Chapter 3. Section 5.4 is devoted to cryptographic applications of these uncertainty relations.

5.2 Quantum to Classical randomness extractors (QC-extractors)

Randomness extractors were introduced by Nisan and Zuckerman [1996] in the context of derandomization. An extractor is a function that transforms a weak source of randomness into almost uniform random bits. The initial motivating applications were related to complexity theory, e.g., derandomization of space-bounded computations [Nisan and Zuckerman, 1996], simulating randomized algorithms with a weak random source [Zuckerman, 1996b] or also as a tool for proving hardness of approximation [Zuckerman, 1996a]. The definition of randomness extractors was actually predated by the similar idea of privacy amplification introduced in a cryptographic context, more precisely for quantum key distribution [Bennett et al., 1988, 1995]. There, the setting is as follows. Suppose Alice and Bob share a bitstring $X$ about which Eve might have some information $E$. They want to extract a secret key about which Eve has almost no information. Here $X$ viewed from the point of view of $E$ is a weak source of randomness from which we want to distill almost perfect random bits. It is particularly clear in this picture that an extractor should work subject only to the assumption that the source contains some randomness, and not make any assumption on where this randomness is. The reason is that depending on her attack, Eve can obtain information about different parts of $X$. In Chapters 3 and 4 we saw yet other applications of randomness extractors to uncertainty relations and low-distortion norm embeddings. For more background on extractors, their constructions and applications, see the surveys [Shaltiel, 2002, Vadhan, 2007].

Classical sources of randomness are described by probability distributions and the randomness extractors are families of (deterministic) functions taking each possible value of the source to a binary string. To understand the definition of quantum extractors, it is convenient to see a classical extractor as a family of permutations acting on the possible values of the source. This family of permutations should satisfy the following property: for any probability distribution on input bit strings with high min-entropy, applying a typical permutation from the family to the input induces an almost uniform probability distribution on a prefix of the output; see 3.4.4 for a definition. We define a quantum to quantum

\footnote{Throughout this thesis, we will only deal with what are known as seeded extractors. For an overview of the different kinds of extractors for different kinds of sources, see [Shaltiel, 2002].}
extractor in a similar way by allowing the operations performed to be general unitary transformations and the input to the extractor to be quantum.

**Definition 5.2.1 (QQ-Extractors).** Let \( A = A_1 A_2 \) with \( n = \log d_A \).

For \( k \in [-n, n] \) and \( \epsilon \in [0, 1] \), a \((k, \epsilon)\)-QQ-extractor is a set \( \{U_1, \ldots, U_t\} \) of unitary transformations on \( A \) such that for all states \( \rho_{AE} \in S(AE) \) satisfying \( H_{\min}(A|E)_\rho \geq k \), we have

\[
\frac{1}{t} \sum_{i=1}^{t} \left\| \text{tr}_{A_2} \left[ U_i \rho_{AE} U_i^\dagger \right] - \frac{\text{id}_{A_1}}{d_{A_1}} \otimes \rho_E \right\|_1 \leq \epsilon .
\] (5.4)

\( \log t \) is called the seed size of the QQ-extractor.

Before making a few remarks on the definition, we recall the definition of CC-extractors which are simply randomness extractors that work in the presence of a quantum adversary.

**Definition 5.2.2 (CC-Extractors).** For \( k \in [0, n] \) and \( \epsilon \in [0, 1] \), a \((k, \epsilon)\)-CC-extractor is a set \( \{f_1, \ldots, f_t\} \) of functions from \( \{0, 1\}^n \) to \( \{0, 1\}^m \) such that for all states \( \rho_{XE} \in S(XE) \) satisfying \( H_{\min}(X|E)_\rho \geq k \), we have

\[
\frac{1}{t} \sum_{i=1}^{t} \left\| \rho_{f_i(X)E} - \frac{\text{id}_Z}{d_Z} \otimes \rho_E \right\|_1 \leq \epsilon ,
\] (5.5)

where the system \( f_i(X) \) is obtained by applying the function \( f_i \) to the system \( X \).

First, we should stress that the same set of unitaries should satisfy (5.4) for all states \( \rho_{AE} \) that meet the conditional min-entropy criterion \( H_{\min}(A|E)_\rho \geq k \). In particular, the system \( E \) can have arbitrarily large dimension. The quantity \( H_{\min}(A|E)_\rho \) measures the uncertainty that an adversary has about the system \( A \). As it is usually impossible to model the knowledge of an adversary, a bound on the conditional min-entropy is often all one can get. A notable difference with the classical setting is that the conditional min-entropy \( k \) can be negative when the systems \( A \) and \( E \) are entangled.

A statement of the form of equation (5.4) is more commonly known as a decoupling result \cite{Abeyesinghe2009, Dupuis2010, Dupuis2010a, Hayden2008, Horodecki2005b, Horodecki2006}. Such statements play an important role in quantum information theory and many coding theorems amount to proving a decoupling theorem.

In fact, it was shown that a set of unitaries forming a unitary 2-design (see Definition 5.2.4) define a \((k, \epsilon)\)-QQ-extractor as long as the output size \( \log d_{A_1} \leq (n + k)/2 - \log(1/\epsilon) \). The decoupling theorem of \cite{Dupuis2010, Dupuis2010a} is actually more general than this: it holds even if we replace \( \text{tr}_{A_2} \) by any completely positive trace preserving map \( T \). Of
course, then the value of $\epsilon$ depends on an entropic quantity that is a function of the map $T$ in addition to the term $H_{\text{min}}(A|E)_\rho$, which was already present for QQ-extractors.

A definition of quantum extractors was also proposed in [Ben-Aroya et al., 2010, Definition 5.1]. A set of unitaries $\{U_1, \ldots, U_t\}$ acting on $A$ is a $(k, \epsilon)$-quantum extractor if for all $\rho \in S(A)$ with $H_{\text{min}}(A)_\rho \geq k$, we have

$$\left\| \frac{1}{t} \sum_{i=1}^{t} U_i \rho U_i^\dagger - \frac{\text{id}_A}{d_A} \right\|_1 \leq \epsilon.$$  

(5.6)

First we note that in their definition, the extractor outputs the whole system (nothing is discarded). This is only possible because they measure the distance between the randomized state $\frac{1}{t} \sum_{i} U_i \rho U_i^\dagger$ and the maximally mixed state (this condition refers to weak extractors), whereas in our definition we ask for the average of the trace distances to be small (strong extractors). It is easy to see that by the triangle inequality, a strong extractor is also a weak extractor. For strong extractors, the seed can be made public, i.e., even conditioned on the value of the random seed, the outcome is close to random. This is not the case for weak extractors. In cryptography, weak extractors are usually not good enough because the seed $i$ is made public during the protocol (see for example Section 5.4). Our definition is also stronger in another respect, we require the extractor to decouple the $A$ system from any quantum side information held in the system $E$.

Ben-Aroya et al. [2010] introduced their definition in the context of studying quantum expanders. In fact, they obtain extractors for high min-entropy sources using their construction of quantum expanders \(^4\) as well as the construction of Ambainis and Smith [2004]; see also [Desrosiers and Dupuis, 2010, Theorem 3] where the construction based on Ambainis and Smith [2004] is studied the language of approximately randomizing maps [Hayden et al., 2004]. Ben-Aroya et al. [2010] applied their extractor construction to prove that the quantum entropy difference problem is in the complexity class QSZK; see the paper for more details.

In the context of cryptography, a QQ-extractor is often more than one needs. In fact, it is usually sufficient to extract random classical bits, which is in general easier to obtain than random qubits. This motivates the following definition, which differs from a QQ-extractor in that the output system $A_1$ is measured in the computational basis. In particular, any $(k, \epsilon)$-QQ-extractor is also a $(k, \epsilon)$-QC-extractor.

**Definition 5.2.3 (QC-Extractors).** Let $A = A_1A_2$ with $n = \log d_A$, and let $T_{A \rightarrow A_1}$ be the measurement map defined in equation (5.3).
For $k \in [-n, n]$ and $\epsilon \in [0, 1]$, a $(k, \epsilon)$-QC-extractor is a set $\{U_1, \ldots, U_t\}$ of unitary transformations on $A$ such that for all states $\rho_{AE} \in S(AE)$ satisfying $H_{\min}(A|E) \geq k$, we have

$$\frac{1}{t} \sum_{i=1}^{t} \left\| T_{A \rightarrow A_1}(U_i \rho_{AE} U_i^\dagger) - \frac{\text{id}_{A_1}}{d_{A_1}} \otimes \rho_E \right\|_1 \leq \epsilon .$$ (5.7)

\(\log t\) is called the seed size of the QC-extractor.

Observe that Definition 5.2.3 only allows a specific form of measurements obtained by applying a unitary transformation followed by a measurement in the computational basis of $A_1$. The reason we restrict the measurements to be of this form is that we want the output of the extractor to be perfectly determined by the source and the choice of the seed. In the classical setting, an extractor is a family of deterministic functions of the source and the seed. In the quantum setting, a natural way of translating this requirement is by imposing that an adversary holding a system that is maximally entangled with the source can perfectly predict the output. This condition is satisfied by the form of measurements dictated by Definition 5.2.3. Allowing generalized measurements (POVMs) already (implicitly) allows the use of randomness for free. Note also, that in the case where the system $E$ is trivial, a $(0, \epsilon)$-QC-extractor is the same as an $\epsilon/2$-metric uncertainty relation (Definition 3.2.1).

Our definition of QC-extractors has some connections with some recent work on device independent randomness generation [Acín et al., 2012, Colbeck, 2006, Colbeck and Kent, 2011, Fehr et al., 2011, Pironio and Massar, 2011, Pironio et al., 2010, Vazirani and Vidick, 2011]. The objective of this line of work is to build protocols to generate bits that are certified to be random. The setting is as follows. The system we consider has a special structure: it is composed of two parts $A$ and $B$ that are spatially separated. Using a small random seed, a pair of measurements is chosen to be performed on $A$ and $B$ obtaining outcomes $X$ and $Y$. Then, statistical tests are performed on $X$ and $Y$ to record a Bell inequality violation. Such a violation is then evidence that the systems $A$ and $B$ are entangled, which implies in particular that say $H_{\max}(A|B)$ is significantly smaller than $\log d_A$. But this means that if the adversary holds a purification of the system $AB$, we have $H_{\min}(A|E) = -H_{\max}(A|B) \gg -\log d_A$. Thus, by applying a QC-extractor, one can generate almost perfect random bits. The challenge in that context is to detect Bell inequality violations with a small number of measurements.
5.2. Quantum to Classical randomness extractors (QC-extractors)

5.2.1 Examples and limitations of QC-extractors

Universal (or two-independent) hashing is probably one of the most important extractor constructions, which even predates the general definition of extractors [Impagliazzo et al., 1989]. Unitary 2-designs can be seen as a quantum generalization of two-independent hash functions.

**Definition 5.2.4.** A set of unitaries \( \{U_1, \ldots, U_t\} \) acting on \( A \) is said to be a 2-design if for all \( M \in L(A) \), we have

\[
\frac{1}{t} \sum_{i=1}^{t} U_i \otimes^2 M (U_i^\dagger) \otimes^2 = \int U \otimes^2 M (U^\dagger) \otimes^2 dU \tag{5.8}
\]

where the integration is with respect to the Haar measure on the unitary group.

Many efficient constructions of unitary 2-designs are known [Dankert et al., 2009, Gross et al., 2007], and in an \( n \)-qubit space, such unitaries can typically be computed by circuits of size \( O(n^2) \). However, observe that the number of unitaries of a 2-design is at least \( t \geq d_A^4 - 2d_A^2 + 2 \) [Gross et al., 2007]. The following is immediate using a general decoupling result of Dupuis [2010], Dupuis et al. [2010a] (see Lemma A.3.2).

**Corollary 5.2.5.** Let \( A = A_1 A_2 \) with \( n = \log d_A \). For all \( k \in [-n, n] \) and all \( \epsilon > 0 \), a unitary 2-design \( \{U_1, \ldots, U_t\} \) on \( A \) is a \((k, \epsilon)\)-QC-extractor with output size

\[
\log d_{A_1} = \min(n, n + k - 2 \log(1/\epsilon)). \tag{5.9}
\]

Similar results also hold for almost unitary 2-designs; see [Szehr et al., 2011]. Using the results of Harrow and Low [2009], this shows for instance that random quantum circuits of size \( O(n^2) \) are QC-extractors with basically the same parameters as in Corollary 5.2.5.

Proposition 5.2.7 below shows that the output size of these QC-extractors is basically optimal. In fact, even if we are looking for a QC-extractor that works for a particular state \( \rho_{AE} \), the output size is at most \( n + H_{\min}^c(A|E)_{\rho} \), where \( n \) denotes the size of the input. In order to do that, we start by proving that when we measure a quantum system \( A \), the min-entropy increases by at most the logarithm of the dimension of the system being measured.

**Lemma 5.2.6.** Let \( \rho_{AB} \in \mathcal{S}(AB) \), \( \epsilon \geq 0 \), and \( \{P_x\}_{x=1}^{d_X} \) be a projective rank-one measurement on \( A \). Then

\[
H_{\min}^c(X|B)_{\rho} \leq H_{\min}^c(A|B)_{\rho} + \log d_{X}. 
\]
Proof Let $V_{A \rightarrow X'}$ be an isometric purification of $\{P_x\}$ and $\rho_{XX'BB'}$ a purification of $\rho_{XX'B} = V \rho_{AB} V^\dagger$. By the invariance of the min-entropy under local isometries [Tomamichel et al. 2010, Lemma 13] and the duality between the min- and max-entropy (Lemma 2.2.3), the proposition becomes equivalent to

$$H_{\max}^\epsilon(X|B')_\rho \leq H_{\max}^\epsilon(X'B')_\rho + \log d_X,$$

By the definition of the smooth max-entropy (equations (2.15)-(2.16)), there exists $\hat{\rho}_{XX'B'} \in B'(\rho_{XX'B'})$ and $\hat{\sigma}_{X'B'} \in S(X'B')$ such that

$$H_{\max}^\epsilon(X|B')_\rho = \log F(\hat{\rho}_{XX'B'}, \text{id}_X \otimes \hat{\sigma}_{X'B'})^2,$$

as well as $\bar{\rho}_{XX'B'} \in B'(\rho_{XX'B'})$ and $\bar{\sigma}_B \in S(B)$ such that

$$H_{\max}^\epsilon(X'B')_\rho = \log F(\bar{\rho}_{XX'B'}, \text{id}_X \otimes \bar{\sigma}_B)^2.$$

Now observe that

$$H_{\max}^\epsilon(X|B')_\rho \leq \log \left( d_X \cdot F(\hat{\rho}_{XX'B'}, \text{id}_X \otimes \frac{\text{id}_{X'}}{d_X} \otimes \hat{\sigma}_{X'B'})^2 \right)$$

$$\leq \max_{\sigma_{X'B'}} \log F(\hat{\rho}_{XX'B'}, \text{id}_X \otimes \sigma_{X'B'})^2 + \log d_X$$

$$= H_{\max}^\epsilon(X|B')_\rho + \log d_X.$$

$\square$

**Proposition 5.2.7** (Upper bound on the output size). Let $A = A_1 A_2$, $\rho_{AE} \in S(AE)$, $\{U_1, \ldots, U_t\}$ a set of unitaries on $A$, and $T_{A \rightarrow A_1}$ defined as in equation (5.3), such that

$$\frac{1}{t} \sum_{i=1}^t \left\| T_{A \rightarrow A_1} \left( U_i \rho_{AE} U_i^\dagger \right) - \frac{\text{id}_{A_1}}{d_{A_1}} \otimes \rho_E \right\|_1 \leq \epsilon. \quad (5.10)$$

Then,

$$\log d_{A_1} \leq \log d_A + H_{\min}^{\sqrt{\epsilon}}(A|E)_\rho.$$

**Proof** Consider the projective rank-one measurements $\{P_x\}$ obtained by performing $U_i$ followed by a measurement in the computational basis of $A$. As a result, we can apply Lemma 5.2.6 and obtain for all $i \in \{1, \ldots, t\}$

$$H_{\min}^{\sqrt{\epsilon}}(A|E)_\rho + \log d_A \geq H_{\min}^{\sqrt{\epsilon}}(X_i|E)_\rho,$$
where $X_i$ denotes the outcome of the measurement $\{P^i\}$. But condition (5.10) implies that there exists $i \in \{1, \ldots, t\}$ such that

$$\left\| \mathcal{T}_{A \to A_1} \left( U_i \rho_{AE} U_i^\dagger \right) - \frac{id_{A_1}}{d_{A_1}} \otimes \rho_E \right\|_1 \leq \epsilon.$$ 

In other words, $\frac{id_{A_1}}{d_{A_1}} \otimes \rho_E \in B^{\sqrt{\tau}} \left( \mathcal{T}_{A \to A_1} \left( U_i \rho_{AE} U_i^\dagger \right) \right)$. By monotonicity of the min-entropy for classical registers [Berta et al., 2011c, Lemma C.5], we have that

$$H_{\min}^{\sqrt{\tau}}(X_i|E)_\rho \geq H_{\min}^{\sqrt{\tau}}(A_1|E)_{\mathcal{T}_{A \to A_1}(U_i \rho_{AE} U_i^\dagger)} \geq \log d_{A_1},$$

which proves the desired result. 

We now study the seed size property. We prove that choosing a reasonably small set of unitaries at random defines a QC-extractor with high probability. The seed size in this case is of the same order as the output size of the extractor. We expect that a much smaller seed size would be sufficient. However, as will be proved in Proposition 5.2.10 below, different methods would have to be used in order to prove that.

**Theorem 5.2.8.** Let $A = A_1 A_2$ with $n = \log d_A$ and $\mathcal{T}_{A \to A_1}$ be the measurement map defined in equation (5.3). Let $\epsilon > 0$, $c$ be a sufficiently large constant,

$$\log d_{A_1} \leq n + k - 4 \log(1/\epsilon) - c \quad \text{and} \quad \log t \geq \log d_{A_1} + \log n + 4 \log(1/\epsilon) + c.$$ 

Then, choosing $t$ unitaries $\{U_1, \ldots, U_t\}$ independently according to the Haar measure defines a $(k, \epsilon)$-QC-extractor with high probability. See (5.15) for a probability bound.

**Proof** The proof uses one-shot decoupling techniques [Dupuis, 2010, Dupuis et al., 2010a, Szehr et al., 2011] combined with an operator Chernoff bound [Ahlswede and Winter, 2002] (see Lemma A.3.5).

Let $U$ be a unitary on $A$. We use the Hölder-type inequality (see e.g., [Bhatia, 1997, Corollary IV.2.6])

$$\|\alpha \beta \gamma\|_1 \leq \|\alpha\|_r^{1/r} \|\beta\|_s^{1/s} \|\gamma\|_{r'}^{1/r'}$$

where $1/r + 1/s + 1/r' = 1$. We use it with $r = r' = 4$, $s = 2$, and $\alpha = \gamma = (id_{A_1} \otimes \rho_E)^{1/4}$,
\[ \beta = (\text{id}_{A_1} \otimes \rho_E)^{-1/4} \left( \mathcal{T}(U \rho_{AE} U^\dagger) - \frac{\text{id}_{A_1}}{d_{A_1}} \otimes \rho_E \right) (\text{id}_{A_1} \otimes \rho_E)^{-1/4} \] to get that\footnote{The inverses are generalized inverses.}

\[ \left\| \mathcal{T}(U \rho_{AE} U^\dagger) - \frac{\text{id}_{A_1}}{d_{A_1}} \otimes \rho_E \right\|_1 \leq d_{A_1}^{1/4} \sqrt{\text{tr} \left[ (\text{id}_{A_1} \otimes \rho_E)^{-1/4} \left( \mathcal{T}(U \rho_{AE} U^\dagger) - \frac{\text{id}_{A_1}}{d_{A_1}} \otimes \rho_E \right) (\text{id}_{A_1} \otimes \rho_E)^{-1/4} \right]^2 d_{A_1}^{1/4}} \]

\[ = d_{A_1}^{1/2} \left\| (\text{id}_{A_1} \otimes \rho_E)^{-1/4} \left( \mathcal{T}(U \rho_{AE} U^\dagger) - \frac{\text{id}_{A_1}}{d_{A_1}} \otimes \rho_E \right) (\text{id}_{A_1} \otimes \rho_E)^{-1/4} \right\|_2 \]

\[ = d_{A_1}^{1/2} \left\| \mathcal{T}(U \tilde{\rho}_{AE} U^\dagger) - \frac{\text{id}_{A_1}}{d_{A_1}} \otimes \tilde{\rho}_E \right\|_2, \]

where \( \tilde{\rho}_{AE} = (\text{id}_A \otimes \rho_E)^{-1/4} \rho_{AE} (\text{id}_A \otimes \rho_E)^{-1/4} \). Together with the concavity of the square root function, this implies

\[ \frac{1}{t} \sum_{i=1}^t \left\| \mathcal{T}(U_i \rho_{AE} U_i^\dagger) - \frac{\text{id}_{A_1}}{d_{A_1}} \otimes \rho_E \right\|_1 \leq \sqrt{\frac{1}{t} \sum_{i=1}^t \left\| \mathcal{T}(U_i \rho_{AE} U_i^\dagger) - \frac{\text{id}_{A_1}}{d_{A_1}} \otimes \rho_E \right\|_1^2} \]

\[ \leq \sqrt{d_{A_1} \frac{1}{t} \sum_{i=1}^t \left\| \mathcal{T}(U_i \tilde{\rho}_{AE} U_i^\dagger) - \frac{\text{id}_{A_1}}{d_{A_1}} \otimes \tilde{\rho}_E \right\|_2^2} \]

\[ = \sqrt{d_{A_1} \frac{1}{t} \sum_{i=1}^t \text{tr} \left[ \mathcal{T}(U_i \tilde{\rho}_{AE} U_i^\dagger) - \frac{\text{id}_{A_1}}{d_{A_1}} \otimes \tilde{\rho}_E \right]^2}. \quad (5.11) \]

We continue with

\[ \frac{1}{t} \sum_{i=1}^t \text{tr} \left[ \mathcal{T}(U_i \tilde{\rho}_{AE} U_i^\dagger) - \frac{\text{id}_{A_1}}{d_{A_1}} \otimes \tilde{\rho}_E \right]^2 \]

\[ = \frac{1}{t} \sum_{i=1}^t \text{tr} \left[ \mathcal{T}(U_i \tilde{\rho}_{AE} U_i^\dagger) \right]^2 - 2 \text{tr} \left[ \mathcal{T}(U_i \tilde{\rho}_{AE} U_i^\dagger) \left( \frac{\text{id}_{A_1}}{d_{A_1}} \otimes \tilde{\rho}_E \right) \right] + \text{tr} \left[ \frac{\text{id}_{A_1}}{d_{A_1}} \otimes \tilde{\rho}_E \right] \]

\[ = 2 \sum_{i=1}^t \text{tr} \left[ \mathcal{T}(U_i \tilde{\rho}_{AE} U_i^\dagger) \right] - \text{tr} \left[ \mathcal{T}(U_i \tilde{\rho}_{AE} U_i^\dagger) \left( \frac{\text{id}_{A_1}}{d_{A_1}} \otimes \tilde{\rho}_E \right) \right] + \text{tr} \left[ \frac{\text{id}_{A_1}}{d_{A_1}} \otimes \tilde{\rho}_E \right] \]

\[ = \frac{1}{d_{A_1}} \text{tr} \left[ \tilde{\rho}_E^2 \right]. \quad (5.12) \]
Going back to equation (5.12), we obtain

$$\frac{1}{t} \sum_{i=1}^{t} \text{tr} \left[ T \left( U_i \tilde{\rho}_{AE} U_i^\dagger \right) - \frac{\text{id}_{A_1}}{d_{A_1}} \otimes \tilde{\rho}_E \right]^2 = \frac{1}{t} \sum_{i=1}^{t} \text{tr} \left[ T \left( U_i \tilde{\rho}_{AE} U_i^\dagger \right) \right]^2 - \frac{1}{d_{A_1}} \text{tr} \left[ \tilde{\rho}_E^2 \right].$$  

(5.13)

Let $F_{AA'}$ denote the swap operator $F_{AA'} = \sum_{aa'} |aa'\rangle\langle a'a|$. We now compute the first term using the “swap trick” (Lemma A.3.4)

$$\text{tr} \left[ T(U \tilde{\rho}_{AE} U^\dagger) \right]^2$$

$$= \text{tr} \left[ \sum_{a_1a_2} \langle a_1a_2|U \tilde{\rho}_{AE} U^\dagger|a_1a_2\rangle|a_1\rangle\langle a_1| \right]^2$$

$$= \text{tr} \left[ \sum_{a_1a_2a_1'a_2'} \langle a_1a_2a_1'a_2'|U^{\otimes 2}\tilde{\rho}_{AE}^{\otimes 2}(U^{\otimes 2})^\dagger|a_1a_2a_1'a_2'|a_1a_1'|\langle F_{A_1A_1'} \otimes F_{EE'} \rangle \right]$$

$$= \sum_{a_1a_2a_1'a_2'} \text{tr} \left[ \tilde{\rho}_{AE}^{\otimes 2}(U^{\otimes 2})^\dagger|a_1a_2a_1'a_2'|\langle a_1a_1'|\left( F_{A_1A_1'} \otimes F_{EE'} \right) |a_1a_1'\rangle|a_1a_2a_1'a_2'|U^{\otimes 2} \right].$$

In the last equality, we used the fact that $|a_1a_1'|$ commutes with the scalar $\langle a_1a_2a_1'a_2'|U^{\otimes 2}\tilde{\rho}_{AE}^{\otimes 2}(U^{\otimes 2})^\dagger|a_1a_2a_1'a_2'\rangle$ and the cyclicity of the trace. Taking the average over the set $\{U_1, \ldots, U_t\}$, we get

$$\frac{1}{t} \sum_{i=1}^{t} \text{tr} \left[ T \left( U_i \tilde{\rho}_{AE} U_i^\dagger \right) \right]^2$$

$$= \sum_{a_1a_2a_1'a_2'} \text{tr} \left[ \tilde{\rho}_{AE}^{\otimes 2} \frac{1}{t} \sum_{i=1}^{t} \left\{ \left( U_i^{\otimes 2} \right)^\dagger |a_1a_2a_1'a_2'|F_{A_1A_1'} |a_1a_1'| \langle a_1a_2a_1'a_2'|U_i^{\otimes 2} \right\} \otimes F_{EE'} \right]$$

$$= \text{tr} \left[ \tilde{\rho}_{AE}^{\otimes 2} \frac{1}{t} \sum_{i=1}^{t} \left\{ \left( U_i^{\otimes 2} \right)^\dagger \sum_{a_1a_2a_1'a_2'} |a_1a_2a_1'a_2'| \langle a_1a_2a_1'a_2'|U_i^{\otimes 2} \right\} \otimes F_{EE'} \right].$$  

(5.14)

Using for example [Dupuis et al., 2010a, Lemma 3.4], if $U$ is distributed according to the Haar measure on the group of unitaries acting on $A$, then

$$E_U \left\{ \left( U^{\dagger} \right)^{\otimes 2} \sum_{a_1a_2a_1'a_2'} |a_1a_2a_1'a_2'| \langle a_1a_2a_1'a_2'|U^{\otimes 2} \right\}$$

$$= \left( \frac{d_Ad_{A_2}}{d_A^2 - 1} \right) \text{id}_{AA'} + \frac{d_A - d_{A_2}}{d_A^2 - 1} F_{AA'}.$$
We use the shorthand $\Gamma_{AA'}$ for the expression above. Now we note that $\frac{\delta_{A}d_{A2} - 1}{d_{A}^2 - 1} \geq \frac{1}{2d_{A}}$, and apply the operator Chernoff bound (Lemma A.3.5) to get

$$\Pr \left\{ \frac{1}{t} \sum_{i=1}^{t} (U_{i}^\dagger)^{\otimes 2} \sum_{a_{1}a_{2}a_{1}'a_{2}'} |a_{1}a_{2}a_{1}'a_{2}'\rangle \langle a_{1}a_{2}a_{1}'a_{2}'| U_{i}^{\otimes 2} \leq (1 + \eta) \Gamma \right\}$$

$$\geq 1 - d_{A} \exp \left( -\frac{\eta^2}{\left(\frac{d_{A} - d_{A2}}{2d_{A} - 1}\right)^2} \right).$$

(5.15)

This shows that if $t \geq 2 \cdot 4 \ln 2 \cdot d_{A} \log d_{A}/\eta^2$, the unitaries $U_{1}, \ldots, U_{t}$ satisfy the above operator inequality with high probability. In the rest of the proof, we show that such unitaries define QC-extractors. Putting these unitaries in equation (5.14), we get

$$\frac{1}{t} \sum_{i=1}^{t} \tr \left[ T \left( U_{i} \tilde{\rho}_{AE} \left( U_{i}^\dagger \right)^{\dagger} \right) \right] \leq (1 + \eta) \left( \frac{d_{A}d_{A2} - 1}{d_{A}^2 - 1} \tr \left[ \rho_{E}^2 \right] + \frac{d_{A} - d_{A2}}{d_{A}^2 - 1} \tr \left[ \rho_{AE}^2 \right] \right).$$

Plugging this expression in equation (5.13) and then in equation (5.11), we get

$$\frac{1}{t} \sum_{i=1}^{t} \left\| T \left( U_{i} \rho_{AE} \left( U_{i}^\dagger \right)^{\dagger} \right) \rho_{E} \right\| _{1} - \frac{\text{id}_{A_{1}}}{d_{A_{1}}} \otimes \rho_{E} \right\| _{1} \leq \sqrt{(1 + \eta) \left( \frac{d_{A}^2 - d_{A1}}{d_{A}^2 - 1} \tr \left[ \rho_{E}^2 \right] + (1 + \eta) \left( \frac{d_{A1}d_{A} - d_{A2}}{d_{A}^2 - 1} \tr \left[ \rho_{AE}^2 \right] \right) \right)}$$

$$\leq \sqrt{\eta + (1 + \eta) \frac{d_{A1}}{d_{A} + 1} \tr \left[ \rho_{AE}^2 \right]},$$

since $\tr \left[ \rho_{E}^2 \right] = \tr \left[ \left( \text{id}_{A} \otimes \rho_{E}^{-1/4} \right) \rho_{AE} \left( \text{id}_{A} \otimes \rho_{E}^{-1/4} \right) \right]^2 = \tr \left[ \rho_{E} \right] = 1$. By the definition of the conditional collision entropy (equation (2.17)) and Lemma 2.2.4, it follows that,

$$\frac{1}{t} \sum_{i=1}^{t} \left\| T \left( U_{i} \rho_{AE} \left( U_{i}^\dagger \right)^{\dagger} \right) \rho_{E} \right\| _{1} - \frac{\text{id}_{A_{1}}}{d_{A_{1}}} \otimes \rho_{E} \right\| _{1} \leq \sqrt{(1 + \eta) \frac{d_{A1}}{d_{A} + 1} 2^{-H_{2}(A|E)_{\rho_{E}}} \rho_{E}}$$

$$\leq \sqrt{\eta + (1 + \eta) \frac{d_{A1}}{d_{A} + 1} 2^{-H_{\min}(A|E)_{\rho_{E}}} \rho_{E}}.$$

(5.16)

Now let $\rho_{AE}' \in B^{\delta + \delta'}(\rho_{AE})$ be such that $H_{\min}(A|E)_{\rho_{E}} = H_{\min}(A|E)_{\rho_{E}'}$. Since we have $\| \rho_{AE}' - \rho_{AE} \| _{1} \leq 2(\delta + \delta')$ (by equation (2.7)), we know that by the triangle inequality and the monotonicity of the trace distance,

$$\left\| T \left( U_{\rho_{AE}U^{\dagger}} \right) - \frac{\text{id}_{A_{1}}}{d_{A_{1}}} \otimes \rho_{E} \right\| _{1} - \left\| T \left( U_{\rho_{AE}'U^{\dagger}} \right) - \frac{\text{id}_{A_{1}}}{d_{A_{1}}} \otimes \rho_{E} \right\| _{1} \leq \left\| T \left( U_{\rho_{AE}U^{\dagger}} \right) - T \left( U_{\rho_{AE}'U^{\dagger}} \right) \right\| _{1}$$

$$\leq \| \rho_{AE}' - \rho_{AE} \| _{1} \leq 2(\delta + \delta'),$$
and hence applying (5.16) to \( \rho'_{AB} \), we get
\[
\frac{1}{t} \sum_{i=1}^{t} \left\| \mathcal{T} \left( U_i \rho_{AE} (U_i)^\dagger \right) - \frac{id_{A_1}}{d_{A_1}} \otimes \rho_E \right\|_1 \\
\leq \sqrt{\eta + (1 + \eta) \frac{d_{A_1}}{d_A + 1} 2^{-H_{\min}(A|E)_\rho} + 2(\delta + \delta')}.
\]
We then use Lemma 2.2.2 about the equivalence of the different conditional min-entropies to get
\[
\frac{1}{t} \sum_{i=1}^{t} \left\| \mathcal{T} \left( U_i \rho_{AE} (U_i)^\dagger \right) - \frac{id_{A_1}}{d_{A_1}} \otimes \rho_E \right\|_1 \\
\leq \sqrt{\eta + (1 + \eta) \frac{d_{A_1}}{d_A + 1} 2^{-H_{\min}(A|E)_\rho + z} + 2(\delta + \delta')},
\]
(5.17)
with \( z = \log(2/\delta^2 + 1/(1 - \delta)) \). Setting \( \eta = \epsilon^2/4, \delta = 0, \delta' = \epsilon/4 \), and assuming \( \log d_{A_1} \leq n + k - 4 \log(1/\epsilon) - c \) with \( k = H_{\min}(A|E)_\rho \), we can upper bound equation (5.17) for large enough \( c \) by
\[
\frac{\epsilon}{2} + \sqrt{\epsilon^2/4 + 2 \cdot 2^{k-4 \log(1/\epsilon) - c - k + \log(8/\epsilon^2+1)}} \\
\leq \frac{\epsilon}{2} + \sqrt{\epsilon^2/4 + \epsilon^2 \cdot 2^{1-c+4}} \leq \epsilon.
\]
\( \square \)

The following simple argument shows that the number of unitaries of a QC-extractor has to be at least \( 1/\epsilon \).

**Proposition 5.2.9** (Lower bound on seed size). Let \( A = A_1 A_2 \). Any \((k, \epsilon)\)-QC-extractor with \( k \leq \log d_A - 1 \) is composed of a set of unitaries on \( A \) of size at least \( t \geq 1/\epsilon \).

**Proof** Let \( S \subseteq [d_{A_1}] \) be an arbitrary subset of \( d_{A_1}/2 \) basis elements of \( A_1 \). Then consider the state
\[
\rho_A = \frac{2}{d_{A_1}} \sum_{a_1 \in S, a_2 \in [d_{A_2}]} U_1^\dagger |a_1 a_2 \rangle \langle a_1 a_2 | U_1.
\]
Note that \( \mathcal{T}(U_1 \rho_A U_1^\dagger) = \frac{2}{d_{A_1}} \sum_{a_1 \in S} |a_1 \rangle \langle a_1 | \) and thus \( \| \mathcal{T}(U_1 \rho_A U_1^\dagger) - \frac{id_{A_1}}{d_{A_1}} \|_1 = 1 \). This implies the claim.
\( \square \)

Observe that in the case where the system \( E \) is trivial (or classical), we showed in Theorem 3.3.2 that there exists QC-extractors composed of \( t = O(\log(1/\epsilon) \epsilon^{-2}) \) unitaries.
This is a difference with classical extractors, for which the number of possible values of the seed has to be at least $\Omega\left( (n-k)\epsilon^{-2} \right)$ [Radhakrishnan and Ta-Shma, 2000] and a probabilistic construction shows that this is tight. It is not clear whether this is an important difference or whether it simply comes from the fact that the analogue for $(0,\epsilon)$-QC-extractors should be $(n,\epsilon)$-CC-extractors. An interesting question in this regard is to see whether one can prove an analogous lower bound on the seed size for $(k,\epsilon)$-QC-extractors with negative $k$.

Observe that in the analysis of Theorem 5.2.8, we actually proved something stronger than condition (5.7). There and actually in all the constructions given in this chapter, we prove that the stronger condition (5.18) below holds. The following proposition shows that with such a strong definition, the seed has to be quite large. In particular, to show the existence of QC-extractors (or even QQ or CC-extractors) with a short seed, one should use different techniques to bound the trace distance directly.

**Proposition 5.2.10** (Extractors for the 2-norm). Let $A = A_1A_2$. Let $\{U_1, \ldots, U_t\}$ be unitaries such that

$$
\frac{1}{t} \sum_{i=1}^{t} \left\| T_{A_1 	o A_1} (U_i \rho_{AE} U_i^\dagger) - \frac{id_{A_1}}{d_{A_1}} \otimes \rho_E \right\|_2^2 \leq \frac{\epsilon^2}{d_{A_1}}.
$$

Then, $t \geq 1/\epsilon^2 \cdot \min \left( \frac{d_{A_2}}{2^{2k}}, d_{A_1}/4 \right)$.

**Proof** Let $S \subseteq [d_{A_1}]$ be an arbitrary subset of $\max(1, \lceil 2^k/d_{A_2} \rceil)$ basis elements of $A_1$. Then consider the state

$$
\rho_A = \frac{1}{|S| \cdot d_{A_2}} \sum_{a_1 \in S, a_2 \in [d_{A_2}]} U_i^\dagger |a_1 a_2 \rangle \langle a_1 a_2| U_i.
$$

We have $H_{\text{min}}(A)_\rho \geq k$ and

$$
T_2(U_1 \rho_A U_1^\dagger) = \frac{1}{|S|} \sum_{a_1 \in S} |a_1 \rangle \langle a_1|.
$$

We can then compute

$$
\left\| T(U_1 \rho_A U_1^\dagger) - \frac{id_{A_1}}{d_{A_1}} \right\|_2^2 \geq \sum_{a_1 \in S} \left( \frac{1}{|S|} - \frac{1}{d_{A_1}} \right)^2,
$$

$$
= |S| \left( \frac{1}{|S|} - \frac{1}{d_{A_1}} \right)^2.
$$

As a result, we have $\frac{1}{t} \cdot |S| \left( \frac{1}{|S|} - \frac{1}{d_{A_1}} \right)^2 \leq \frac{\epsilon^2}{d_{A_1}}$. In the case $2^k/d_{A_2} \leq 1$, we obtain a lower bound of $t \geq 1/\epsilon^2 \cdot d_{A_1}/(1 - 1/d_{A_1})^2$. In the other case, we get

$$
t \geq 1/\epsilon^2 \cdot d_{A_1} / |S| \geq 1/\epsilon^2 \cdot \frac{d_{A_1}}{2 \cdot 2^k/d_{A_2}} = 1/\epsilon^2 \cdot \frac{d_{A}}{2 \cdot 2^k}.
$$
Our results about QC-extractors are summarized in Table 5.1.

| Seed | CC-extractors | QC-extractors |
|------|---------------|---------------|
| LB   | \(\log(n-k) + 2\log(1/\epsilon)\) [2000] | \(\log(1/\epsilon)\) |
| UB   | \(\log(n-k) + 2\log(1/\epsilon)\) \(c \cdot \log(n/\epsilon)\) [2009] | \(m + \log n + 4\log(1/\epsilon)\) [5.2.8] |

| Output |
|--------|
| UB     | \(k - 2\log(1/\epsilon)\) [2000] |
| LB     | \(k - 2\log(1/\epsilon)\) [1989, 2005] |

Table 5.1: Known lower bounds (LB) and upper bounds (UB) on the seed size and output size in terms of (qu)bits for different kinds of \((k, \epsilon)\)-randomness extractors. \(n\) refers to the number of input (qu)bits, \(m\) the number of output (qu)bits and \(k\) the min-entropy of the input \(H_{\text{min}}(A|E)\). Note that for QC-extractors, \(k\) can be as small as \(-n\). Additive absolute constants are omitted. We note that the constructions corresponding to the second line are non-explicit.

5.2.2 Full set of mutually unbiased bases (MUBs)

We saw that unitary 2-designs define QC-extractors. As unitary 2-designs also define QQ-extractors, it is natural to expect that we can build smaller and simpler sets of unitaries if we are only interested in extracting random classical bits. To that end, in this section, we construct simpler sets of unitaries that define QC-extractors. Two ingredients are used: a full set of mutually unbiased bases and a family of pairwise independent permutations.

A set of unitaries \(\{U_1, \ldots, U_t\}\) acting on \(A\) is said to define mutually unbiased bases if for all elements \(|a\rangle, |a'\rangle\) of the computational basis of \(A\), we have \(|\langle a' | U_j U_i^\dagger | a \rangle|^2 \leq d_A^{-1}\) for all \(i \neq j\). In other words, a state described by a vector \(U_i^\dagger |a\rangle\) of the basis \(i\) gives a uniformly distributed outcome when measured in basis \(j\) for \(i \neq j\). For example the two bases, sometimes called computational and Hadamard bases (used in most quantum cryptographic protocols), are mutually unbiased. There can be at most \(d_A + 1\) mutually unbiased bases for \(A\). Constructions of full sets of \(d_A + 1\) MUBs are known in prime power dimensions [Bandyopadhyay et al., 2002, Wootters and Fields, 1989]. Such unitaries can be implemented by quantum circuits of almost linear size; see Lemma 3.4.1. Mutually unbiased bases also have applications in quantum state determination [Ivonovic, 1981, Wootters and Fields, 1989].
To state our result, we will need one more notion. A family \( P \) of permutations of a set \( X \) is **pairwise independent** if for all \( x_1 \neq x_2 \) and \( y_1 \neq y_2 \), and if \( \pi \) is uniformly distributed over \( P \), \( \Pr \{ \pi(x_1) = y_1, \pi(x_2) = y_2 \} = \frac{1}{d_X(d_X - 1)} \). If \( X \) has a field structure, i.e., if \( d_X \) is a prime power, it is simple to see that the family \( P = \{ x \mapsto ax + b : a \in X^*, b \in X \} \) is pairwise independent. In the following, a permutation of basis elements of a Hilbert space \( A \) should be seen as a unitary transformation on \( A \).

**Theorem 5.2.11.** Let \( A = A_1A_2 \) with \( n = \log d_A \), \( d_A \) a prime power, and consider the map \( T_{A \rightarrow A_1} \) as defined in equation (5.3). Then, if \( \{ U_1, \ldots, U_{d_A+1} \} \) defines a full set of mutually unbiased bases, we have for \( \delta \geq 0 \),

\[
\left\| \frac{1}{|P|} \sum_{P \in P} \sum_{i=1}^{d_A+1} T_{A \rightarrow A_1} \left( PU_i \rho_{AE} (PU_i)^\dagger \right) - \frac{id_{A_1} \otimes \rho_E}{d_{A_1}} \right\|_1 \\
\leq \sqrt{\frac{d_{A_1}}{d_A + 1} 2^{-H_{\min}(A|E)_\nu}} + 2\delta, \tag{5.19}
\]

where \( P \) is a set of pairwise independent permutation matrices. In particular, the set \( \{ PU_i : P \in P, i \in [d_A + 1] \} \) defines a \((k, \epsilon)-QC\)-extractor provided

\[
\log d_{A_1} \leq n + k - 2 \log(1/\epsilon),
\]

and the number of unitaries is

\[
t = (d_A + 1)|P|,
\]

which for the pairwise independent permutations described above gives \( t = (d_A + 1)d_A(d_A - 1) \).

**Proof** The idea is to bound the trace norm in equation (5.19) by the Hilbert-Schmidt norm of some well-chosen operator. This term is then computed exactly using the fact that the set of all the MUB vectors form a complex projective 2-design (Lemma A.3.3), and the fact that the set of permutations is pairwise independent.

Similar to the proof of Theorem 5.2.8, but with the difference that now \( \tilde{\rho}_{AE} = (id_A \otimes \sigma_E)^{-1/4} \rho_{AB} (id_A \otimes \sigma_E)^{-1/4} \) for some \( \sigma_E \in S(E) \) to be chosen later, we get

\[
\left\| \frac{1}{|P|} \sum_{P \in P} \sum_{i=1}^{d_A+1} T \left( PU_i \rho_{AE} (PU_i)^\dagger \right) - \frac{id_{A_1} \otimes \rho_E}{A_1} \right\|_1 \\
\leq \sqrt{d_{A_1} \sum_{a_1a_2a_2'} \operatorname{tr} \left[ \tilde{\rho}_{AE}^2 \Gamma_{a_1a_2a_2'} \otimes F_{EE'} \right] - \operatorname{tr} \left[ \tilde{\rho}_{AE}^2 \right]} , \tag{5.20}
\]
where \( \Gamma_{a_1,a_2,a_2'} = \mathbb{E}_{P,i} \left\{ (U_i^\dagger P i)^{\otimes 2} |a_1a_2a_1a_2'|a_1a_2a_1a_2'| (PU_i)^{\otimes 2} \right\} \) and \( F_{EE'} \) is the swap operator. We now compute \( \Gamma_{a_1,a_2,a_2'} \) handling the case \( a_2 = a_2' \) and the case \( a_2 \neq a_2' \) differently. When \( a_2 = a_2' \), we have \( (U_i^{\dagger P i})^{\otimes 2} |aa\rangle\langle aa|U_i^{\otimes 2} = (U_i^{\dagger i})|a|\langle a|U_i^{\otimes 2} \), where \( a = P^{-1}(a_1a_2) \). As \( \{U_1, \ldots, U_{d_\Lambda + 1}\} \) form a full set of mutually unbiased bases, the vectors \( \{U_i|a\}\}_{i,a} \) define a complex projective 2-design (Lemma A.3.3), and we get

\[
\sum_{a_1,a_2,a_2'=a_2} \mathbb{E}_{P,i} \left\{ (U_i^{\dagger P i})^{\otimes 2} |a_1a_2a_1a_2'|a_1a_2a_1a_2'| (PU_i)^{\otimes 2} \right\} = \sum_a \mathbb{E}_i \left\{ (U_i^{\dagger i})^{\otimes 2} |aa\rangle\langle aa|U_i^{\otimes 2} \right\} = d_A \frac{2\Pi^{\text{sym}}_{AA'}}{(d_A + 1)d_A} = \frac{\text{id}_{AA'} + F_{AA'}}{d_A + 1},
\]

(5.21)

where \( \Pi^{\text{sym}}_{AA'} \) is the projector onto the symmetric subspace of \( AA' \), i.e., the subspace spanned by vectors \( |a'a\rangle + |aa'\rangle \). We now consider \( a_2 \neq a_2' \). We have

\[
\mathbb{E}_P \left\{ (P^i)^{\otimes 2} |a_1a_2a_1a_2'|a_1a_2a_1a_2'| P^{\otimes 2} \right\} = \mathbb{E}_P \left\{ |P^{-1}(a_1a_2)\rangle\langle P^{-1}(a_1a_2)| \otimes |P^{-1}(a_1a_2')\rangle\langle P^{-1}(a_1a_2')| \right\} = \sum_{a\neq a'} \text{Pr}_P \left\{ P^{-1}(a_1a_2) = a, P^{-1}(a_1a_2') = a' \right\} |a\rangle\langle a| \otimes |a'\rangle\langle a'|
\]

\[
= \frac{1}{d_A(d_A - 1)} \sum_{a\neq a'} |a\rangle\langle a| \otimes |a'\rangle\langle a'|
\]

\[
= \frac{\text{id}_{AA'}}{d_A(d_A - 1)} - \frac{1}{d_A(d_A - 1)} \sum_a |aa\rangle\langle aa|
\]

(5.22)

Going back to equation (5.21), we get together with equation (5.22) that for any \( a_2 \neq a_2' \),

\[
\mathbb{E}_{P,i} \left\{ (U_i^{\dagger i})^{\otimes 2} (P^i)^{\otimes 2} |a_1a_2a_1a_2'|a_1a_2a_1a_2'| P^{\otimes 2} U_i^{\otimes 2} \right\} = \frac{\text{id}_{AA'}}{d_A(d_A - 1)} - \frac{1}{d_A(d_A - 1)} \sum_a \mathbb{E}_i \left\{ (U_i^{\dagger i})^{\otimes 2} |aa\rangle\langle aa|U_i^{\otimes 2} \right\}
\]

\[
= \frac{\text{id}_{AA'}}{d_A(d_A - 1)} + \frac{\text{id}_{AA'} + F_{AA'}}{d_A(d_A - 1)(d_A + 1)}
\]

\[
= \frac{d_A \text{id}_{AA'} - F_{AA'}}{d_A(d_A^2 - 1)}.
\]
This being true for all $a_1, a_2, a'_2$, it follows with equation (5.20) that,

$$E_{P,i} \left\{ \| T \left( PU_i \rho_{AE} (PU_i)^\dagger \right) - \frac{id_{A_1}}{d_{A_1}} \otimes \rho_E \|_1 \right\}$$

$$= \sqrt{d_{A_1}} \left[ \bar{\rho}_{AE}^2 \left( \frac{id_{A_1} + F_{AA'}}{d_{A_1} + 1} + d_A(d_A - 1) \frac{d_{A_1}id_{A_1} - F_{AA'}}{d_{A_1}^2 - 1} \right) \otimes F_{EE'} \right] - \text{tr} \left[ \bar{\rho}_{E}^2 \right].$$

(5.23)

Expanding the expression inside the square root, we obtain

$$d_{A_1} \left( \frac{1}{d_{A_1} + 1} + \frac{d_A(d_A - 1)}{d_{A_1}^2 - 1} \right) \text{tr} \left[ \bar{\rho}_{AE}^2 \left( id_{A_1} \otimes F_{EE'} \right) \right]$$

$$+ d_{A_1} \left( \frac{1}{d_{A_1} + 1} - \frac{d_A - 1}{d_{A_1}^2 - 1} \right) \text{tr} \left[ \bar{\rho}_{AE}^2 \left( F_{AA'} \otimes F_{EE'} \right) \right] - \text{tr} \left[ \bar{\rho}_{E}^2 \right]$$

$$= d_{A_1} \frac{d_A d_A - 1}{d_{A_1}^2 - 1} \text{tr} \left[ \bar{\rho}_{AE}^2 \left( id_{A_1} \otimes F_{EE'} \right) \right]$$

$$+ d_{A_1} \frac{d_A - d_A}{d_{A_1}^2 - 1} \text{tr} \left[ \bar{\rho}_{AE}^2 \right] - \text{tr} \left[ \bar{\rho}_{E}^2 \right].$$

Continuing from (5.23), we get

$$= \sqrt{\frac{d_{A_1}^2 - d_{A_1}}{d_{A_1}^2 - 1}} \text{tr} \left[ \bar{\rho}_{E}^2 \right] + \left( \frac{d_A d_A - d_A}{d_{A_1}^2 - 1} \right) \text{tr} \left[ \bar{\rho}_{AE}^2 \right]$$

$$\leq \sqrt{\frac{d_{A_1}}{d_{A_1} + 1}} \text{tr} \left[ \bar{\rho}_{AE}^2 \right] = \sqrt{\frac{d_{A_1}}{d_{A_1} + 1} 2^{-H_2(A|E)_{\rho|\sigma}}} ,$$

where we used the definition of the conditional collision entropy (equation (2.17)) in the last step. Now, by choosing $\sigma_E$ appropriately, and an argument analogous to the very end of the proof of Theorem 5.2.8, we conclude that,

$$E_{P,i} \left\{ \| T \left( PU_i \rho_{AE} (PU_i)^\dagger \right) - \frac{id_{A_1}}{d_{A_1}} \otimes \rho_E \|_1 \right\} \leq \sqrt{\frac{d_{A_1}}{d_{A_1} + 1} 2^{-H_2(A|E)_{\rho|\sigma}}} + 2\delta .$$

In terms of output size, this construction is almost optimal, but the number of unitaries is again much larger than we expect should be possible.

### 5.2.3 Bitwise QC-extractor

The unitaries we construct in this section are even simpler. They are composed of unitaries $V$ acting on single qudits followed by permutations $P$ of the computational basis elements.
5.2. Quantum to Classical randomness extractors (QC-extractors)

Note that this means that the measurements defined by these unitaries can be implemented with current technology. As the measurement $\mathcal{T}$ commutes with the permutations $P$, we can first apply $V$, then measure in the computational basis and finally apply the permutation to the (classical) outcome of the measurement. In addition to the computational efficiency, the fact that the unitaries act on single qudits, is often a desirable property for the design of cryptographic protocols. In particular, the application to the noisy storage model that we present in Section 5.4 does make use of this fact. But the price we pay is that the parameters (both output and seed size) are worse than the previous construction.

Let $d \geq 2$ be a prime power so that there exists a complete set of mutually unbiased bases in dimension $d$. We represent such a set of bases by a set of unitary transformations $\{V_0, V_1, \ldots, V_d\}$ mapping these bases to the standard basis. For example, for the qubit space ($d = 2$), we can choose

$$V_0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad V_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}, \quad V_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & i \\ i & -1 \end{pmatrix}. $$

We define the set $V_{d,n}$ of unitary transformations on $n$ qudits by

$$V_{d,n} \overset{\text{def}}{=} \{V = V_{u_1} \otimes \cdots \otimes V_{u_n} | u_i \in \{0, \ldots, d\}\}.$$

As in the previous section, $\mathcal{P}$ denotes a family of pairwise independent permutations.

**Theorem 5.2.12.** Let $A = A_1 A_2$ with $d_A = d^n$, $d_{A_1} = d^{\xi n}$, $d_{A_2} = d^{(1-\xi)n}$, and $d$ a prime power. Consider the map $\mathcal{T}_{A \rightarrow A_1}$ as defined in equation (5.3). Then for $\delta \geq 0$ and $\delta' > 0$,

$$\frac{1}{|\mathcal{P}|} \left( \frac{1}{d+1} \right)^n \sum_{P \in \mathcal{P}} \sum_{V \in V_{d,n}} \left\| \mathcal{T}_{A \rightarrow A_1} \left( PV \rho_{A_E} (PV)^\dagger \right) - \frac{id_{A_1}}{d_{A_1}} \otimes \rho_E \right\|_1 \leq \sqrt{2(1-\log(d+1)+\xi \log d)n \left( 1 + 2^{-H_{\min}(A|E)_{\nu+z}} \right) + 2(\delta + \delta')},$$

(5.24)

where $V_{d,n}$ is defined as above, $\mathcal{P}$ is a set of pairwise independent permutation matrices, and $z = \log \left( \frac{2}{\delta^2} + \frac{1}{1-\delta} \right)$. In particular, the set $\{PV : P \in \mathcal{P}, V \in V_{d,n}\}$ is a $(k, \epsilon)$-extractor provided

$$\log d_{A_1} \leq (\log(d+1)-1)n + \min \{0, k\} - 4 \log(1/\epsilon) - 7$$

and if we choose the pairwise independent permutations described in Theorem 5.2.11, the number of unitaries is

$$t = (d+1)^n d^n (d^n - 1).$$
The analysis uses the same technique as in the proof of Theorem 5.2.11. The main difference is that we were not able to express the Hilbert-Schmidt norm exactly in terms of the conditional min-entropy $H_{\min}(A|E)_{\rho}$. Instead, we use some additional inequalities, which account for the slightly more complicated expression we obtain.

**Proof** We use the same strategy as in the proofs of Theorem 5.2.8 and Theorem 5.2.11; here again with $\tilde{\rho}_{AE} = (id_A \otimes \rho_E)^{-1/4} \rho_{AE} (id_A \otimes \rho_E)^{-1/4}$. As in (5.20) and (5.22), we get

$$E_{PV} \left\{ \left\| T \left( PV \rho_{AE} (PV)^\dagger \right) - \frac{id_{A_1}}{d_{A_1}} \otimes \rho_E \right\|_1 \right\} \leq \sqrt{d_{A_1} \text{tr} \left[ \tilde{\rho}_{AE}^{\otimes 2} \left( \sum_a \Gamma_a + d_A (d_{A_2} - 1) \frac{id_{AA'}}{d_A (d_A - 1)} \right) \otimes F_{EE'} \right] - \text{tr} [\tilde{\rho}_E^2]} , \quad (5.25)$$

where $\Gamma_a = E_{V} \left\{ (V^\dagger |a\rangle \langle a| V)^{\otimes 2} \right\}$. We calculate

$$\sum_a E_{V} \left\{ (V^\dagger |a\rangle \langle a| V)^{\otimes 2} \right\} = \frac{1}{(d + 1)^n} \sum_{a_1, a_2, \ldots, a_n} \sum_{V_1, \ldots, V_n} \bigotimes_i \left( (V_i^\dagger |a_i\rangle \langle a_i| V_i)^{\otimes 2} \right)$$

$$= \frac{1}{(d + 1)^n} \bigotimes_i \left( \sum_{a_i, V_i} (V_i^\dagger |a_i\rangle \langle a_i| V_i)^{\otimes 2} \right) . \quad (5.26)$$

As $\{V_0, \ldots, V_d\}$ form a maximal set of mutually unbiased bases in dimension $d$, and with this form a complex projective 2-design (Lemma A.3.3), we have

$$\sum_{a \in \{0, \ldots, d\}, V \in V_{d,1}} (V^\dagger |a\rangle \langle a| V)^{\otimes 2} = 2\Pi^{\text{sym}} ,$$

where $\Pi^{\text{sym}}$ is the projector onto the symmetric subspace, i.e., the subspace spanned by vectors $|a'a\rangle + |aa'\rangle$. Furthermore $(\Pi_B^{\text{sym}})^{\otimes n} \leq \Pi_B^{\text{sym}}$ for any inner product space $B$, and hence we obtain

$$\frac{1}{(d + 1)^n} \bigotimes_i \left( \sum_{a_i, V_i} (V_i^\dagger |a_i\rangle \langle a_i| V_i)^{\otimes 2} \right) \leq \left( \frac{2}{d + 1} \right)^n \Pi^{\text{sym}}_{AA'}$$

$$= \left( \frac{2}{d + 1} \right)^n \frac{id_{AA'} + F_{AA'}}{2} . \quad (5.27)$$

Plugging equation (5.27) into the expression inside the square root in equation (5.25), we
5.2. Quantum to Classical randomness extractors (QC-extractors)

can bound it by

\[ \leq d_{A_1} \text{tr} \left[ \tilde{\rho}_{AE}^{\otimes 2} \left( \left( \frac{2}{d + 1} \right)^n \text{id}_{A A'} + F_{A A'} \frac{d_{A_2} - 1}{d_A(d_A - 1)} \right) \otimes F_{E E'} \right] \]

\[ - \text{tr} \left[ \tilde{\rho}_{E}^{2} \right] \]

\[ = \left( \frac{d_A - d_{A_1}}{d_A - 1} + \frac{d_{A_1}}{d_A} \right) \left( \frac{2}{d + 1} \right)^n \text{tr} \left[ \tilde{\rho}_{AE}^{\otimes 2} (\text{id}_{A A'} \otimes F_{E E'}) \right] \]

\[ + \frac{d_{A_1}}{2} \left( \frac{2}{d + 1} \right)^n \text{tr} \left[ \tilde{\rho}_{AE}^{\otimes 2} (F_{A A'} \otimes F_{E E'}) \right] - \text{tr} \left[ \tilde{\rho}_{E}^{2} \right] \]

\[ \leq (1 + 2^{(1-\log(d+1)+\xi \log d)n}) \text{tr} \left[ \theta_{A A'}^{\otimes 2} (\text{id}_{A A'} \otimes F_{E E'}) \right] \]

\[ + 2^{(1-\log(d+1)+\xi \log d)n} \text{tr} \left[ \theta_{AE}^{2} - \text{tr} \left[ \tilde{\rho}_{E}^{2} \right] \right]. \]

Continuing from (5.25), we get

\[ \leq \sqrt{2^{(1-\log(d+1)+\xi \log d)n} \text{tr} \left[ \tilde{\rho}_{E}^{2} \right] + 2^{(1-\log(d+1)+\xi \log d)n} \text{tr} \left[ \tilde{\theta}_{AE}^{2} \right]} \]

\[ = \sqrt{2^{(1-\log(d+1)+\xi \log d)n} \left( 1 + 2^{-H_{\text{min}}(A|E)_{\rho_{\rho}}} \right)} , \]

where we used the definition of the conditional collision entropy (equation (2.17)) in the last step. Again by an argument analogous to the very end of the proof of Theorem 5.2.8, we conclude that,

\[ \mathbf{E}_{P, V} \left\{ \left\| T \left( P V \rho_{A E} (P V)^{\dagger} \right) - \frac{\text{id}_{A_1}}{d_{A_1}} \otimes \rho_{E} \right\|_1 \right\} \]

\[ \leq \sqrt{2^{(1-\log(d+1)+\xi \log d)n} \left( 1 + 2^{-H_{\text{min}}(A|E)_{\rho_{\rho}}} \right) + 2(\delta + \delta')}, \]

where \( z = \log \left( \frac{2}{\delta^2} + \frac{1}{1-\delta} \right) \). Setting \( \delta = 0 \) and \( \delta' = \epsilon / 4 \), we conclude that the set \( \{ P V : P \in \mathcal{P}, V \in \mathcal{V}_{d,n} \} \) is a \((k, \epsilon)\)-QC-extractor provided

\[ \log d_{A_1} = n \cdot \xi \log d \]

\[ \leq \left( \log (d + 1) - 1 \right) n - \log (1 + 2^{-k+\log(8/\epsilon^2+1)}) + \log((\epsilon/2)^2) \]

\[ \leq \left( \log (d + 1) - 1 \right) n + \min \{ 0, k - \log(8/\epsilon^2 + 1) \} - 2 \log(1/\epsilon) - 3 \]

\[ \leq \left( \log (d + 1) - 1 \right) n + \min \{ 0, k \} - 4 \log(1/\epsilon) - 7. \]

\[ \square \]

It seems that the parameters proved for this QC-extractor construction are not optimal. In fact, equation (5.24) does not give anything non-trivial when \( H_{\text{min}}(A|E)_{\rho} < -\left( \log (d + 1) - 1 \right) n \). We believe however that it should be possible to obtain a non-trivial statement for
any min-entropy as long as it is larger than $-cn \log d$ for some $c < 1$. Such an improvement would be quite interesting for the application we provide to two-party secure computation in Section 5.4 (see the discussion following Theorem 5.4.2). We think that the place where the analysis should be improved is the inequality (5.27). If we do not use this inequality, we end up with having to handle an expression of the form

$$\sum_{S \subseteq [n]} \text{tr}\left[\tilde{\rho}_{A_S}^2\right],$$

(5.28)

where $A_S$ refers to the qudits of $A$ indexed by elements of $S$. Because we have $\text{tr}[\tilde{\rho}_{A_S}^2] \leq 2^{|S| \log d}$ for any $S$, the sum in equation (5.28) is always bounded by $(d + 1)^n$. It would be interesting for example to show that whenever $H_{\min}(A|E) \geq -cn \log d$ for some $c < 1$, then there exists some $\beta < (d + 1)^n$ such that the sum in (5.28) is bounded by $\beta^n$. This kind of statement is related to min-entropy sampling [Konig and Renner, 2011]. The problem there is to prove that for most subsets $S$ of $[n]$ of size $r$, we have $H_{\min}(A_S|E) \gtrsim \frac{r}{n} H_{\min}(A|E)$. Such a statement was proved by [Konig and Renner, 2011] in the case where $A$ is classical. It would be interesting to see if such a result holds when $A$ is a general quantum system.

### 5.3 Entropic uncertainty relations with quantum side information

In this section, we show how to obtain entropic uncertainty relations from general QC-extractors.

#### 5.3.1 Min-entropy uncertainty relations

We start by proving uncertainty relations for the smooth min-entropy, which is usually the relevant measure in the context of cryptography. Consider the state $\rho_{X_1 ... X_t} = \frac{1}{t} \sum_{j=1}^t \mathcal{M}_{A \rightarrow X}(U_j \rho_{AE} U_j^\dagger) \otimes |j\rangle \langle j|$. We note that unlike for the von Neumann entropy, the conditional entropy $H_{\min}(X|E)_{\rho}$ is not the same as the average $\frac{1}{t} \sum_{j=1}^t H_{\min}(X|E)_{\rho^j}$, where $\rho^j = \mathcal{M}_{A \rightarrow X}(U_j \rho_{AE} U_j^\dagger)$. However, by concavity of the logarithm, we have

$$\frac{1}{t} \sum_{j=1}^t H_{\min}(X|E)_{\rho^j} \geq - \log \left[ \frac{1}{t} \sum_{j=1}^t 2^{-H_{\min}(X|E)_{\rho^j}} \right] = H_{\min}(X|E)_{\rho}.$$ 

Here, we used the expression for the conditional min-entropy in (2.11). It follows that proving lower bounds for $H_{\min}(X|E)$ is stronger and directly gives lower bounds for
the average measurement entropy. For this reason, we use the conditional min-entropy $H_{\min}(X|EJ)$ in place of the average entropy.

The following lemma shows that a QC-extractor directly satisfies an uncertainty relation for the smooth min-entropy. The idea is simple: if the outcome of the QC-extractor is $\epsilon$-close to $\frac{id_{A_1}}{d_{A_1}} \otimes \rho_E$, then by the definition of the smoothed entropies, the smooth min-entropy of the outcome has to be at least $\log d_{A_1}$.

**Lemma 5.3.1.** Let $\rho_{AE} \in S(AE)$, and $\{U_1, \ldots, U_t\}$ be a set of unitaries on $A$ such that,

$$\frac{1}{t} \sum_{j=1}^{t} \left\| T_{A \rightarrow A_1}(U_j \rho_{AE} U_j^\dagger) - \frac{id_{A_1}}{d_{A_1}} \otimes \rho_E \right\|_1 \leq \epsilon(\rho), \quad (5.29)$$

for some $\epsilon(\rho)$ depending on the input state $\rho_{AE}$. Then

$$H_{\min}^{\sqrt{2\epsilon(\rho)}}(X|EJ)_\rho \geq \log d_{A_1},$$

where $\rho_{XE,J} = \frac{1}{t} \sum_{j=1}^{t} M_{A \rightarrow X}(U_j \rho_{AE} U_j^\dagger) \otimes |j\rangle \langle j|_J$ and $M_{A \rightarrow X}$ is the measurement in the computational basis.

**Proof** By the definition of the smooth min-entropy and the inequality (2.7) between the purified and trace distance, condition (5.29) directly translates into

$$H_{\min}^{\sqrt{2\epsilon(\rho)}}(A_1|EJ)_\rho \geq H_{\min}(A_1|EJ)_{\frac{\rho_{A_1} \otimes \rho_E \otimes \rho_{T}}{\pi_{A_1}} = \log d_{A_1}},$$

where $\rho = \frac{1}{t} \sum_{j=1}^{t} |j\rangle \langle j|_J \otimes T_{A \rightarrow A_1}(U_j \rho_{AE} U_j^\dagger)$. Recall that $T$ performs a measurement in the computational basis and then discards a (classical) system called $A_2$. Because we are only discarding a classical system, the min-entropy of the whole measurement outcome is at least the min-entropy in the register $A_1$ (see Lemma 2.2.6). As a result,

$$H_{\min}^{\sqrt{2\epsilon(\rho)}}(X|EJ)_\rho \geq H_{\min}(A_1|EJ)_{\frac{\rho_{A_1} \otimes \rho_E \otimes \rho_{T}}{\pi_{A_1}} = \log d_{A_1}}.$$

This allows us to translate all our constructions from Section 5.2 into a min-entropy uncertainty relation form. Note that conversely, we can convert a min-entropy uncertainty relation into a QC-extractor simply by applying a CC-extractor. We state below uncertainty relations for mutually unbiased bases and for “single-qudit” bases.
Corollary 5.3.2. Let \( A \) be a Hilbert space such that \( d_A \) is a prime power. Let \( \{U_1, \ldots, U_{d_A+1}\} \) be a full set of mutually unbiased bases. For any state \( \rho_{AE} \), we have for all \( \epsilon > 0 \) and \( 0 \leq \delta < \epsilon^2/4 \),

\[
H_{\min}^\epsilon(X|EJ)_\rho \geq \log (d_A + 1) + H_{\min}^\delta(A|E)_\rho - \log \left( \frac{1}{(\epsilon^2/2 - 2\delta)^2} \right) - 1,
\]

where \( \rho_{XEJ} = \frac{1}{d_A+1} \sum_j M_{A \rightarrow X}(U_j \rho U_j^\dagger) \otimes |j\rangle\langle j| \).

Proof. Recall that the unitaries of the QC-extractor of Theorem 5.2.11 are composed of mutually unbiased bases \( U_1, \ldots, U_{d_A+1} \) but also some permutations \( P \in \mathcal{P} \). Letting \( \rho_{XEJ} = \frac{1}{d_A+1} \sum_j M_{A \rightarrow X}(U_j \rho U_j^\dagger) \otimes |j\rangle\langle j| \), Lemma 5.3.1 gives \( H_{\min}^\epsilon(P(X|EJP)_\rho \geq \log d_A \). But \( P \) is a permutation that simply relabels the measurement outcomes, and thus does not change the entropy. It follows that

\[
H_{\min}^\epsilon(X|EJ)_\rho \geq \log d_A.
\]

Now it only remains to choose the dimension \( d_A \). We pick

\[
d_A = \left\lfloor (\epsilon' - 2\delta)^2 \frac{d_A + 1}{2 - H_{\min}^\delta(A|E)} \right\rfloor.
\]

Plugging this value of \( d_A \) in (5.19), we get

\[
\frac{1}{|\mathcal{P}|} \frac{1}{d_A + 1} \sum_{P \in \mathcal{P}} \sum_{i=1}^{d_A+1} \left\| T_{A \rightarrow A_1} \left( PU_i \rho_{AE} (PU_i)^\dagger \right) - \frac{\text{id}_{A_1}}{d_{A_1}} \otimes \rho_E \right\|_1 \leq \epsilon'.
\]

As a result, condition (5.29) is satisfied with \( \epsilon(\rho) = \epsilon' \). The desired result follows from the fact that \( \log d_A \geq \log(1/2) + \log \left( (\epsilon' - 2\delta)^2 \frac{d_A + 1}{2 - H_{\min}^\delta(A|E)} \right) \).

From the point of view of applications, the following entropic uncertainty relation for single-qudit measurements is probably the most interesting. It can be seen as a generalization to allow for quantum side information of uncertainty relations obtained by Damgård et al. [2007]. The proof is very similar to the proof of the previous corollary.

Corollary 5.3.3. Let \( d \geq 2 \) be a prime power. For any state \( \rho_{AE} \), we have

\[
H_{\min}^\epsilon(X|EJ)_\rho \geq n \cdot (\log(d + 1) - 1) + \min \left\{ 0, H_{\min}^\delta(A|E)_\rho - \log \left( \frac{2}{\delta^2} + \frac{1}{1 - 2\delta} \right) \right\}
\]

\[
- \log \left( \frac{1}{(\epsilon^2/2 - 2(\delta + \delta'))^2} \right) - 2,
\]

where \( \rho_{XEJ} = \frac{1}{(d+1)^n} \sum_j M_{A \rightarrow X}(V_j \rho V_j^\dagger) \otimes |j\rangle\langle j| \) and \( \{V_j\}_j = \mathcal{V}_{d,n} \) as defined in Theorem 5.2.12.
5.3. Entropic uncertainty relations with quantum side information

Proof We choose the dimension $d_{A_1}$ of the $A_1$ system of Theorem 5.2.2 to be

$$d_{A_1} = \left\lfloor (\epsilon' - 2(\delta + \delta'))^2 \frac{2(\log(d+1) - 1)n}{1 + 2^{-H_{\min}(A|E)_{\rho} + z}} \right\rfloor.$$ 

We also compute

$$\log d_{A_1} \geq -1 - \log \left( \frac{1}{(\epsilon' - 2(\delta + \delta'))^2} \right) + n \cdot (\log(d + 1) - 1) - \log \left( 1 + 2^{-H_{\min}(A|E)_{\rho} + z} \right)$$

$$= n \cdot (\log(d + 1) - 1) - \log \left( \frac{1}{(\epsilon' - 2(\delta + \delta'))^2} \right) - 2.$$

Setting $\sqrt{2\epsilon'} = \epsilon$, we achieve the desired result. \qed

5.3.2 Uncertainty relations for the von Neumann entropy

Uncertainty relations for the conditional von Neumann entropy can also be obtained as in Proposition 3.2.2.

Lemma 5.3.4. Let $\rho_{AE} \in S(AE)$, and $\{U_1, \ldots, U_t\}$ be a set of unitaries on $A$ such that

$$\frac{1}{t} \sum_{j=1}^{t} \left\| T_{A \rightarrow A_1}(U_j \rho_{AE} U_j^\dagger) - \frac{id_{A_1}}{d_{A_1}} \otimes \rho_E \right\|_1 \leq \epsilon(\rho),$$

for some $\epsilon(\rho)$ depending on the input state $\rho_{AE}$. Then

$$\frac{1}{t} \sum_{j=1}^{t} H(X|E)_{\rho^j} = H(X|E)_{\rho} \geq (1 - 4\epsilon(\rho)) \log d_{A_1} - 2h_2(\epsilon(\rho)),$$

where $\rho^j = \mathcal{M}_{A \rightarrow X}(U_j \rho_{AE} U_j^\dagger)$ and $\rho_{XEJ} = \frac{1}{t} \sum_{j=1}^{t} \rho^j \otimes |j\rangle \langle j|.$

Proof The argument is the same as the proof of Lemma 5.3.1, except that instead of just obtaining a bound on the smooth entropy, we use the Alicki-Fannes inequality (Lemma 2.2.5). \qed

This lemma can naturally be applied directly to all the constructions of QC-extractors. For example, for a full set of mutually unbiased bases, by choosing $d_{A_1} = \left\lfloor (\epsilon - 2\delta)^2(d_A + 1)2^{H_{\min}(A|E)} \right\rfloor$, we can get

$$\frac{1}{t} \sum_j H(X|E)_{\rho^j} \geq (1 - 4\epsilon) \left( \log (d_A + 1) + H_{\min}(A|E)_{\rho} - \log \left( \frac{1}{(\epsilon - 2\delta)^2} \right) - 1 \right) - 2h_2(\epsilon).$$
Using the asymptotic equipartition property for the smooth min-entropy, we can obtain an uncertainty relation only in terms of von Neumann entropies.

**Proposition 5.3.5.** Let \( d \geq 2 \) be a prime power, and \( \{V_0, V_1, \ldots, V_d\} \) define a complete set of MUBs of \( \mathbb{C}^d \). Consider the set of measurements on the \( n \) qudit space \( A \) defined by the unitary transformations \( V = V_{u_1} \otimes \cdots \otimes V_{u_n} |u_i \in \{0, \ldots, d\} \) that we index by numbers from 1 to \((d+1)^n\) as \( \{V_j\}_{j \in \{1,\ldots,(d+1)^n\}} \). Then for all \( \rho_{AE} \in S(AE) \), we have

\[
\frac{1}{(d+1)^n} \sum_{j=1}^{(d+1)^n} H(X|E)_{\rho_j} \geq n \cdot (\log(d+1) - 1) + \min\{0, H(A|E)_{\rho}\},
\]

where \( \rho_j = \mathcal{M}_{A \to X}(V_j \rho V_j^\dagger) \).

**Proof** Using the QC-extractor for the single-qudit MUB of Theorem 5.2.12 with \( d_{A_1} = \left\lfloor (\epsilon - 2\delta)^2 \frac{2^{(\log(d+1)-1)n}}{1 + 2^{-H_{\min}(A|E)_{\rho|\rho}}} \right\rfloor \), we get

\[
\frac{1}{(d+1)^n} \sum_{j=1}^{(d+1)^n} H(X|E)_{\rho_j}
\geq (1 - 4\epsilon) \left( n \left( \log(d+1) - 1 \right) - \log \left( 1 + 2^{-H_{\min}(A|E)_{\rho|\rho}} \right) - \log \left( \frac{1}{(\epsilon - 2\delta)^2} \right) - 1 \right)
\]

\[
- 2h_2(\epsilon)
\geq (1 - 4\epsilon) \left( n \left( \log(d+1) - 1 \right) + \min\{0, H_{\min}(A|E)_{\rho|\rho}\} - 2 - \log \left( \frac{1}{(\epsilon - 2\delta)^2} \right) \right)
\]

\[
- 2h_2(\epsilon).
\]

(5.30)

Here, we use a version with \( H_{\min}(A|E)_{\rho|\rho} \) instead of \( H_{\min}(A|E)_{\rho} \) that is in the statement of Theorem 5.2.12. The expression with \( H_{\min}(A|E)_{\rho|\rho} \) is however easily obtained by looking at the proof. Evaluating equation (5.30) on the \( m \)-fold tensor product of the original input...
system $d^n$, and multiplying both sides with $1/m$, we obtain
\[
\frac{1}{(d+1)^n} \sum_{j=1}^{d+1} H(X|E)_{\rho^j} \geq (1 - 4\epsilon) \left( n \log(d+1) - 1 \right) + \min \left\{ 0, \frac{1}{m} H_{\min}(A|E)_{\rho^m|\rho^m} \right\}
\]
\[
- \frac{1 - 4\epsilon}{m} \left( 2 + \log \left( \frac{1}{(\epsilon - 2\delta)^2} \right) \right) - \frac{2h_2(\epsilon)}{m}
\]
\[
\geq (1 - 4\epsilon) \left( n \log(d+1) - 1 \right) + \min \left\{ 0, H(A|E)_{\rho} - \frac{4\sqrt{1 - 2\log \delta} (2 + \frac{n}{2})}{\sqrt{m}} \right\}
\]
\[
- \frac{1 - 4\epsilon}{m} \left( 2 + \log \left( \frac{1}{(\epsilon - 2\delta)^2} \right) \right) - \frac{2h_2(\epsilon)}{m} .
\]
Here we used the fully quantum asymptotic equipartition property for the smooth conditional min-entropy (Lemma 2.2.7). By first letting $m \to \infty$ and then $\epsilon \to 0$, we obtain the desired result. \hfill $\square$

Note that for $n = 1$, this again gives an uncertainty relation for the full set of MUBs only in terms of von Neumann entropies
\[
\frac{1}{d+1} \sum_{j=1}^{d+1} H(X|E)_{\rho^j} \geq \log(d+1) - 1 + \min \left\{ 0, H(A|E)_{\rho^m} \right\} . \tag{5.31}
\]
In the special case when $E$ is trivial, we arrive at
\[
\frac{1}{d+1} \sum_{j=1}^{d+1} H(X)_{\rho^j} \geq \log(d+1) - 1 , \tag{5.32}
\]
which is the best known bound for a full set of MUBs and general $d$ [Ivanovic 1992, Larsen 1990, Sanchez 1993]. But without side information and when $d$ is even, this was improved by Sanchez-Ruiz [1995] to
\[
\frac{1}{d+1} \sum_{j=1}^{d+1} H(X)_{\rho^j} \geq \frac{1}{d+1} \left( \frac{d}{2} \log \left( \frac{d}{2} \right) + \left( \frac{d}{2} + 1 \right) \log \left( \frac{d}{2} + 1 \right) \right) . \tag{5.33}
\]
For one qubit ($d = 2$) the latter gives $2/3$ (which is best possible for three measurements), whereas our bound gives $\log 3 - 1 \approx 0.585$.

### 5.4 Applications to security in the noisy-storage model

We use the min-entropy uncertainty relation of Corollary 5.3.3 to prove the security of secure function evaluation in the noisy storage model.
5.4.1 Introduction

Consider two mutually distrustful parties Alice and Bob who want to collaborate to perform a distributed computation in a secure fashion. Typically, Alice holds $x$ and Bob holds $y$, and they both want to figure out $f(x, y)$ in such a way that each party does not learn too much about the other party’s input. Unfortunately, if we are looking for information theoretic security, it turns out that even quantum communication does not allow us to solve general two-party secure function evaluation [Lo, 1997]. For example, it is known that only weak variants of (information theoretically secure) bit commitment are possible; see Section 4.1.6 and [D’Ariano et al., 2007; Lo and Chau, 1997; Mayers, 1997].

The natural question then is under which assumptions can we obtain secure protocols for two-party computations. Classically, these assumptions typically limit the computational power of a party. One then assumes that a particular problem requires a lot of computational resources to solve in some precise complexity theoretic sense, and then one proves using this assumption that a cheating strategy needs more computational resources than what is available. It goes without saying that the computational assumptions are almost always not proven. As computation is such a complicated notion to understand, a natural question is then whether one can make simpler assumptions on the devices of the parties.

Classically, it is possible to obtain security when we are willing to assume that the adversary’s memory is limited in size [Cachin and Maurer, 1997; Maurer, 1992]. But unfortunately, Dziembowski and Maurer [2004] showed that any classical protocol in which the honest players need to store $n$ classical bits to execute can be broken by an adversary who can store $O(n^2)$ bits.

Motivated by this unsatisfactory gap, it was thus suggested to assume that the attacker’s quantum storage was bounded [Damgård et al., 2005, 2007]. The central assumption in this model is that during waiting times $\Delta t$ introduced in the protocol, the adversary can only store a limited number of qubits $N$. This is the only assumption on the adversary, who is otherwise all powerful. In particular, he can store an unlimited amount of classical information, and perform any operation instantaneously. The latter implies that he is able to perform any encoding and decoding operation before and after using his memory device. König et al. [2012] based on Damgård et al. [2005, 2007] constructed a protocol for bit commitment using BB84 encoded qubits that is secure whenever Bob is only allowed to store $N$ qubits while Alice and Bob exchange more than (roughly) $2N$ qubits during the protocol.

A natural question then is to characterize the property of Bob’s storage device that allows him and Alice to implement secure two-party function evaluation. The noisy-storage model
5.4. Applications to security in the noisy-storage model

 introduced by Schaffner et al. [2008], Wehner et al. [2008] is a generalization of the bounded storage model. As in the bounded storage model, during waiting times $\Delta t$, the adversary can only keep quantum information in his quantum storage device $\mathcal{F}$. Mathematically, such a quantum storage device is simply a quantum channel $\mathcal{F} : \mathcal{S}(\mathcal{H}_{\text{in}}) \rightarrow \mathcal{S}(\mathcal{H}_{\text{out}})$ mapping input states on the space $\mathcal{H}_{\text{in}}$ to some noisy output states on the space $\mathcal{H}_{\text{out}}$. An example of a storage device would be $N$ $d$-dimensional identical memory cells, so that $\mathcal{F}$ takes the form $\mathcal{F} = N \otimes N$. In particular, in the bounded quantum storage model, the adversary is only allowed to store $N$ qubits, which means $\mathcal{F} = \text{id}_2 \otimes N$ [Damgård et al., 2005, 2007]. The kind of statement one proves in this framework is of the following form: Provided $\mathcal{F}$ cannot be used to reliably transmit $n$ bits or qubits of information, the protocol $\mathcal{P}_n$ is secure. We describe precise versions of this statement in the following sections.

5.4.2 The noisy storage model

Weak string erasure

König et al. [2012] showed that bit commitment and oblivious transfer and hence any two-party secure computation, can be implemented securely against an all-powerful quantum adversary given access to a simple primitive called weak string erasure (WSE). Hence, it suffices to construct a protocol for WSE that is secure under the assumption that the storage devices of the parties are noisy, and we will follow that approach here.

The motivation behind the weak string erasure primitive is to create a basic quantum protocol that builds up classical correlations between Alice and Bob which are later used to implement more interesting cryptographic primitives. Informally, weak string erasure achieves the following task. WSE takes no inputs from Alice and Bob. Alice receives as output a randomly chosen string $X^n = X_1, \ldots, X_n \in \{0, 1\}^n$. Bob receives a randomly chosen subset $I \subseteq [n]$ and the substring $X_I$ of $X^n$ corresponding to the bits in positions indexed by $I$. For each $i \in [n]$, we decide independently to put $i$ in the set $I$ with probability $p$. Originally, $p = 1/2$ [König et al., 2012], but any probability $0 < p < 1$ allows for the implementation of oblivious transfer [Mandayam and Wehner, 2011]. The security requirements of weak string erasure are that Alice does not learn $I$, and Bob’s min-entropy given all of his information $B$ is bounded as $H_{\text{min}}(X|B) \geq \lambda n$ for some parameter $\lambda > 0$. To summarize all relevant parameters, we thereby speak of an $(n, \lambda, \epsilon, p)$-WSE scheme.

The precise requirement of security is stated in terms of (approximate) indistinguisha-
bility between the states obtained in an execution of the real protocol and some ideal states. We should highlight that the notion of distance we use here is the trace distance, which is more relevant than the purified distance in the context of cryptography because of its interpretation in terms of distinguishing probability [Helstrom 1967]. It will be convenient to express the distribution of the random subset $I$ by a density operator:

$$\Psi(p) = \sum_{I \subseteq [n]} p^{|I|}(1 - p)^{n - |I|}|I\rangle\langle I|.$$ (5.34)

**Definition 5.4.1 (Non-uniform WSE).** An $(n, \lambda, \varepsilon, p)$-weak string erasure scheme is a protocol between $A$ and $B$ satisfying the following properties:

**Correctness:** If both parties are honest, then there exists an ideal state $\sigma_{X^n I X^n}$ such that

1. The joint distribution of the $n$-bit string $X^n$ and subset $I$ is given by

$$\sigma_{X^n I} = \frac{\text{id}_{X^n}}{2^n} \otimes \Psi(p),$$ (5.35)

2. The joint state $\rho_{AB}$ created by the real protocol is equal to the ideal state: $\rho_{AB} = \sigma_{X^n I X^n}$, where we identify $(A, B)$ with $(X^n, I X^n)$.

**Security for Alice:** If $A$ is honest, then there exists an ideal state $\sigma_{X^n B'}$ such that

1. The random variable $I$ is independent of $A' \hat{X}^n$ and distributed over $2^n$ according to the probability distribution given by (5.34):

$$\sigma_{A' \hat{X}^n I} = \sigma_{A' \hat{X}^n} \otimes \Psi(p).$$ (5.37)

2. The joint state $\rho_{A'B}$ created by the real protocol is equal to the ideal state: $\rho_{A'B} = \sigma_{A' \hat{X}^n I}$, where we identify $(A', B)$ with $(A', \hat{X}_I)$.

Note that any positive $\lambda$ allows one to build a protocol for bit commitment and oblivious transfer but of course, larger values of $\lambda$ naturally lead to better parameters. To give an example, Mandayam and Wehner [2011] prove that using a $(n, \lambda, \varepsilon, 1/3)$-WSE, one can obtain an 1-2 oblivious transfer of strings of length about $\lambda/24 \cdot n$. 
Protocol for weak string erasure

We now construct a very simple protocol for weak string erasure, and prove its security using our bitwise QC-randomness extractor. The only difference to the protocol proposed in König et al. [2012] is that we will use three MUBs per qubit instead of only two. For sake of argument, we state the protocol in a purified form where Alice generates the EPR-pairs and later measures them. Note, however, that the protocol is entirely equivalent to Alice creating single qubits and sending them directly to Bob. That is, honest Alice and Bob do not need any quantum memory to implement the protocol below. In the purified protocol, the choice of bit she encodes is determined randomly by her measurement outcome in the chosen basis on the EPR-pair. The protocol is illustrated in Figure 5.1.

### Protocol Weak string erasure (WSE): Outputs: $x^n \in \{0, 1\}^n$ to Alice, $(\mathcal{I}, z^{[\mathcal{I}]}) \in 2^{[n]} \times \{0, 1\}^{[\mathcal{I}]}$ to Bob.

1. **Alice**: Creates $n$ EPR-pairs $\Phi_i$ and sends half of each pair to Bob.

2. **Alice**: Chooses a bases-specifying string $\theta^n \in_R \{0, 1, 2\}^n$ uniformly at random. For all $i$, she measures the $i$-th qubit in the basis $\theta_i$ to obtain outcome $x_i$.

3. **Bob**: Chooses a basis string $\tilde{\theta}^n \in_R \{0, 1, 2\}^n$ uniformly at random. When receiving the $i$-th qubit, Bob measures it in the basis given by $\tilde{\theta}_i$ to obtain outcome $\tilde{x}_i$.

Both parties wait time $\Delta t$.

4. **Alice**: Sends the basis information $\theta^n$ to Bob, and outputs $x^n$.

5. **Bob**: Computes $\mathcal{I} = \{ i \in [n] \mid \theta_i = \tilde{\theta}_i \}$, and outputs $(\mathcal{I}, z^{[\mathcal{I}]}) := (\mathcal{I}, \tilde{x}_{\mathcal{I}})$.

The proof of correctness of the protocol, and security against dishonest Alice is identical to König et al. [2012], Mandayam and Wehner [2011]. It essentially follows from the fact that Bob never sends any information to Alice. The main difficulty lies in proving security against dishonest Bob. Before embarking on a formal proof, let us first consider the general form that any attack of Bob takes (see Figure 5.2). First of all, note that the noisy-storage model only assumes that Bob has to use his storage device during waiting times $\Delta t$. Let $Q$ denote Bob’s quantum register containing all $n$ qubits that he receives. Note that since there is no communication between Alice and Bob during the transmission of these $n$ qubits, we
can without loss of generality assume that Bob first waits for all $n$ qubits to arrive before mounting any form of attack.

As any operation in quantum theory is a quantum channel, Bob’s attack can be described by a quantum channel $E : S(Q) \rightarrow S(H_{\text{in}} \otimes M)$. This map takes $Q$, to some quantum state on the input of Bob’s storage device ($H_{\text{in}}$), and some arbitrarily large amount of classical information ($M$). For example, $E$ could be an encoding into an error-correcting code. By assumption of the noisy-storage model, Bob’s quantum memory is then affected by noise $F : S(H_{\text{in}}) \rightarrow S(H_{\text{out}})$. After the waiting time, the joint state held by Alice and Bob in the purified version of the protocol, i.e., before Alice measures, is thus of the form

$$\rho_{ABM} = \text{id}_A \otimes \left[ (F \otimes \text{id}_M) \circ E \right] (\Phi^\otimes n),$$

(5.38)

where $\Phi$ is an EPR-pair. After the waiting time, Bob can perform any form of quantum operation to try and recover information about $X$ from the storage device.

### 5.4.3 Security and the quantum capacity

Recall from the definition above that our goal is to show that $H_{\text{min}}^{e}(X|BM\Theta)_{\rho} \geq \lambda \cdot n$ for some parameter $\lambda > 0$. Although it was always clear that security should be related to the channel’s ability to store quantum information, i.e., the quantum capacity of $F$, proving this fact has been a challenge for several years. Partial progress to answering this
Any attack of dishonest Bob is described by an encoding attack $E$ and a guessing attack, since for classical $X$ the min-entropy $H_{\min}(X|BM\Theta)$ is directly related to the probability that Bob guesses $X$. As we will see below, it is however sufficient to consider how well a decoding attack $D$ can preserve entanglement between Alice and Bob, where $D$ acts on $BM$ on the state $\rho_{ABM}$ from equation (5.38).

The question was made in König et al. [2012] and Berta et al. [2011a], where security was linked to the classical capacity and entanglement cost of $F$, respectively. We informally state these results. Both of them prove security of the protocol described above except that two mutually unbiased bases are chosen instead of three. König et al. [2012] prove the protocol is secure whenever for some $R < 1/2$, the channel $F$ is such that any attempt to transmit $nR$ classical bits across $F$ is bound to fail with probability exponentially close to 1. Berta et al. [2011a] consider channels of the form $F = N^{\otimes N}$ and they show that the protocol is secure provided $\frac{N-E_C(N)}{n}$ is bounded away from 1/2, where $E_C(N)$ is the entanglement cost of the channel $N$. The entanglement cost is the amount of entanglement needed to simulate the channel $N$ when classical communication is given for free. $E_C(N)$ is a measure of how good the channel is for sending quantum information but it is in general larger than the quantum capacity.

Note that our objective is to make a statement about some classical information $X$ obtained by measuring $A$ in a randomly selected basis $\Theta$. That is, we effectively ask for an uncertainty relation for these measurements. Previously, however, suitable uncertainty relations were only known for classical side information. The missing ingredient was an uncertainty relation with quantum side information, linked to the channel’s ability to preserve quantum information. Here is where our uncertainty relation of Corollary 5.3.3 comes in.

To state the result, we first define the notion of channel fidelity introduced by Barnum.
et al. [2000] which is perhaps the most widely used quantity to measure how good a channel is at sending quantum information. For a channel \( \mathcal{N} : \mathcal{S}(Q) \rightarrow \mathcal{S}(Q') \), the channel fidelity \( F_c \) quantifies how well \( \mathcal{N} \) preserves entanglement with a reference:

\[
F_c(\mathcal{N}) = F(\Phi_{QA}, [\mathcal{N} \otimes \text{id}_A](\Phi_{Q'A})),
\]

(5.39)

where \( \Phi_{QA} \) is a maximally entangled state. For example, one way of defining the (one-shot) quantum capacity with free classical forward communication of a channel \( F \) is by the maximum of \( \log d_A \) over all encodings \( E : \mathcal{S}(Q) \rightarrow S(H_{in} \otimes M) \) and decodings \( D : \mathcal{S}(B \otimes M) \rightarrow \mathcal{S}(Q') \) such that \( F_c(\mathcal{D} \circ (F \otimes \text{id}_M) \circ E) \geq 1 - \epsilon \) for small enough \( \epsilon \). Here \( \text{id}_M \) refers to a noiseless classical channel.

**Theorem 5.4.2.** Let Bob’s storage device be given by \( \mathcal{F} : \mathcal{S}(H_{in}) \rightarrow \mathcal{S}(B) \). Let \( \epsilon \in (0, 1) \), \( \kappa = 8 \log(4/\epsilon) \), \( \lambda \leq \log 3 - 1 \). Assume that we have

\[
\max_{\mathcal{D}, \mathcal{E}} F_c(\mathcal{D} \circ (\mathcal{F} \otimes \text{id}_M) \circ E) \leq 2^{-(2 - \log 3 + \lambda)n - \kappa}
\]

(5.40)

where the maximum is over all quantum channels \( \mathcal{E} : \mathcal{S}((\mathbb{C}^2)^{\otimes n}) \rightarrow \mathcal{S}(H_{in} \otimes M) \) and \( \mathcal{D} : \mathcal{S}(B \otimes M) \rightarrow \mathcal{S}((\mathbb{C}^2)^{\otimes n}) \).

Then, Protocol 1 implements a \((n, \lambda, \epsilon, 1/3)\)-WSE.

**Proof** The proof of correctness of the protocol, and security against dishonest Alice is identical to König et al. [2012], Mandayam and Wehner [2011] and does not lead to any error terms.

Using the uncertainty relation of Corollary [5.3.3] with \( E = BM\Theta \) on \( \rho_{ABM\Theta} \) we get

\[
H_{\min}(X|BM\Theta)_{\rho} \gtrsim (\log(3) - 1)n + \min\{0, H_{\min}(A|BM\Theta)_{\rho}\}.
\]

(5.41)

Note that because \( \Theta \) is independent of \( ABM \), we have \( H_{\min}(A|BM\Theta)_{\rho} = H_{\min}(A|BM)_{\rho} \). To place a bound on (5.41), we would like to obtain a lower bound on

\[
\min_{\mathcal{E}} H_{\min}(A|BM)_{\rho},
\]

where the minimization is taken over all encoding attacks as described above. We will use condition (5.40) to obtain such a lower bound. We now use an operational interpretation of the conditional min-entropy due to König et al. [2009]:

\[
H_{\min}(A|BM)_{\rho} = -\log d_A \max_{\Lambda_{BM} \rightarrow A'} F(\Phi_{AA'}, \text{id}_A \otimes \Lambda(\rho_{ABM})),
\]

(5.42)
where $\Phi_{AA'}$ is the maximally entangled state across $AA'$. That is, the min-entropy is directly related to the "amount" of entanglement between $A$ and $E = BM$. The map $\Lambda$ in (5.42) can be understood as a decoding attack $D$ aiming to restore entanglement with Alice. Further, note that $|A'| = |Q|$ and we can equivalently upper bound

$$\max_{D,E} F(A, |D \circ (F \otimes \text{id}_M) \circ E| \circ (\Phi_{AQ})) = \max_{D,E} F_c(D \circ (F \otimes \text{id}_M) \circ E).$$

(5.43)

By the assumption on the storage device $F$, we obtain that for any encoding $E$ and decoding $D$ attack of Bob

$$H_{\min}(A|BM)_\rho \geq - \log 2^n F_c(D \circ (F \otimes \text{id}_M) \circ E) \geq - (n - (2 - \log 3)n - \lambda n - \kappa) = -(\log 3 - 1)n + \lambda n + \kappa.$$

Then, using the uncertainty relation for 3 MUBs per qubit of Corollary 5.3.3 (with $\delta = 0$ and $\delta' = \epsilon^2/8$), we get

$$H_{\min}(X|BM\Theta)_\rho \geq \lambda n - \log \left( 2 \cdot 64/\epsilon^4 + 1 \right) - \log(16/\epsilon^4) - 2 + 8 \log(4/\epsilon) \geq \lambda n.$$ 

Note that ideally, we would want a statement of the form: if

$$\max_{D,E} F_c(D \circ (F \otimes \text{id}_M) \circ E) \lesssim 2^{-\lambda n},$$

(5.44)

then the Protocol 1 implements $(n, \lambda, \epsilon, 1/3)$-WSE. Unfortunately, we have a stronger constraint in equation (5.40) with an additional positive factor of $2 - \log 3$. If we wanted to prove security with the condition (5.44), we would need to prove a stronger uncertainty relation than in Corollary 5.3.3. In particular, observe that if $H_{\min}(A|E) \leq -(\log(3) - 1)n$, our uncertainty relation does not give any useful bound. It would be very interesting to improve it so that we can get a non-trivial lower bound for any $H_{\min}(A|E) \geq -cn$ with $c < 1$. Note that, as in [Mandayam and Wehner 2011], we can get arbitrarily close to proving security under a condition of the form (5.44) by using higher dimensional encodings (i.e., using the uncertainty relation of Corollary 5.3.3 with larger values of $d$), but it becomes hard to implement the protocol with current technology.
Example: bounded storage

We look at the simple case where $F = \text{id}^\otimes N$, known as the bounded storage model. In this case, it is simple to prove that if you try to send more than $N$ qubits of information using $F$, the channel fidelity will decrease exponentially in $n$ [Berta et al., 2011a].

**Lemma 5.4.3.** For $n \geq N$, we have

$$\max_{D, E} F_c(D \circ (\text{id}^\otimes N \otimes \text{id}_M) \circ E) \leq 2^{-n+N},$$

where the maximum is over all quantum channels $E : S((\mathbb{C}^2)^\otimes n) \to S((\mathbb{C}^2)^\otimes N \otimes M)$ and $D : S((\mathbb{C}^2)^\otimes N \otimes M) \to S((\mathbb{C}^2)^\otimes n)$.

**Proof** Consider a decomposition of the encoding and decoding map in terms of their Kraus operators as $E(\rho) = \sum_j E_j \rho E_j^\dagger$ and $D(\rho) = \sum_{k,m} \hat{D}_{k,m} \rho \hat{D}_{k,m}^\dagger$ where $\hat{D}_{k,m} = D_{k,m} \otimes |m\rangle\langle m|$. Note that without loss of generality, the latter has this form since it is processing classical forward communication on $M$. Let $\Pi_{k,m}$ denote the projector onto the subspace that $\hat{D}_{k,m}$ maps to. We can now bound

$$F_c(D \circ (\text{id}^\otimes N \otimes \text{id}_M) \circ E)$$

$$= \frac{1}{2^n} \sum_i \langle ii | \left( \sum_{j, j'} \hat{D}_{k,m} E_j \langle \ell'| \hat{E}_{j'}^\dagger \otimes |\ell| \right) \sum_i |ii\rangle$$

$$= \frac{1}{2^n} \left( \sum_{j, j'} \sum_{k,m} \langle \ell| \hat{D}_{k,m} E_j \langle \ell'| \hat{E}_{j'}^\dagger \hat{D}_{k,m}^\dagger |\ell'\rangle \right)$$

$$= \sum_{j, j'} \left| \text{tr} \left[ \hat{D}_{k,m} E_j \left( \frac{\text{id}}{2^{n/2}} \right) \cdot \left( \frac{\text{id}}{2^{n/2}} \right) \Pi_{m,k} \right] \right|^2$$

$$\leq \sum_{j, j'} \left| \text{tr} \left[ \hat{D}_{k,m} E_j \left( \frac{\text{id}}{2^{n/2}} \right) \hat{E}_{j'}^\dagger \hat{D}_{k,m}^\dagger \right] \text{tr} \left[ \Pi_{k,m} \left( \frac{\text{id}}{2^{n/2}} \right) \right] \right|$$

$$\leq 2^{-n+N} \left| \text{tr} \left[ D \circ E \left( \frac{\text{id}}{2^n} \right) \right] \right|$$

$$= 2^{-n+N},$$

where in the third equality, we used the cyclicity of the trace and fact that $\Pi_{m,k} \hat{D}_{k,m} = \hat{D}_{k,m}$. We used the Cauchy-Schwarz inequality for the first inequality, and the fact that $\text{tr} [\Pi_{k,m}] = \text{rank}[\hat{D}_{k,m}] = \text{rank}[D_{k,m} \otimes |m\rangle\langle m|] \leq 2^N$. For the last equality, we used the fact that $D$ and $E$ are trace preserving. \qed
It then follows from Theorem 5.4.2 that if \( N/n \) is bounded away from \( \log 3 - 1 \), then the described protocol is secure. We note that the parameters obtained here are slightly worse than what was obtained in Mandayam and Wehner [2011], where security was shown to be possible for \( N/n \) bounded away from \( 2/3 \) instead of 0.585. This is due to the fact that the lower bound 0.585 in our uncertainty relation stems from an expression involving the collision entropy rather than the Shannon entropy. We emphasize however, that due to finite size effects our bound is still better in the practically relevant regime of \( n \lesssim 10^6 \) (for the same security parameters).

5.5 Concluding remarks

In this chapter, we considered uncertainty relations that take into account an adversary that is potentially entangled with the system being measured. As in Chapter 3, the measure of uncertainty we used is the distance to the uniform distribution. But in addition, we also asked for the joint state of the outcome together with the system of the adversary to be close to a product state. The advantage of this measure is that we were able to apply techniques similar to the decoupling theorem. We first use a Hölder type inequality to work with the \( \ell_2 \)-norm, which is much easier to handle. Then, we use symmetry properties of the unitaries to obtain bounds on these norms. This allowed us to analyse several constructions of basis, but as we saw in Proposition 5.2.10 this technique cannot be used to show uncertainty relations for small sets of bases. Handling quantum side information using the \( \ell_1 \) norm directly seems like a difficult technical challenge. In the context of CC-extractors, there are constructions that have a small seed [De et al., 2009, König and Terhal 2008, Ta-Shma, 2009]. It would be interesting to use these ideas to construct QQ or QC-extractors with small seed.

We then used one of our uncertainty relations for single-qubit measurements to relate the security of two-party secure function evaluation to the capacity of the storage device to store quantum information reliably. We showed that provided the storage device is “very bad” at storing \( n \) qubits, there is a protocol for performing secure function evaluation in which Alice and Bob communicate \( n \) qubits. This is the first time the security is related to the capacity of the channel to send quantum information. As explained in the discussion following Theorem 5.4.2, this is not totally satisfying, but is hopefully a step towards proving the ideal result which would be that we get security as soon as the storage device is just “bad” at storing quantum information.
Chapter 6

Discussion

6.1 Summary

In this thesis, we considered uncertainty relations for several observables and their applications to quantum information theory. We have first seen how the problem of finding uncertainty relations is closely related to the problem of finding large almost Euclidean subspaces of $\ell_1(\ell_2)$. Even though we did not use any norm embedding result directly, many of the ideas presented here come from the proofs and constructions in the study of the geometry of normed spaces. In particular, we obtained an explicit family of bases that satisfy a strong metric uncertainty relation by adapting a construction of Indyk [2007]. Moreover, using standard techniques from asymptotic geometric analysis, we were able to prove a strong uncertainty relations for random bases.

We used these uncertainty relations to exhibit strong locking effects. In particular, we obtained the first explicit construction of a method for encrypting a random $n$-bit string in an $n$-qubit state using a classical key of size polylogarithmic in $n$. Moreover, our non-explicit results give better key sizes than previous constructions while simultaneously meeting a stronger locking definition. In particular, we showed that an arbitrarily long message can be locked with a constant-sized key. Our results on locking are summarized in Table 4.1. We should emphasize that, even though we presented information locking from a cryptographic point of view, it is not a composable primitive because an eavesdropper could choose to store quantum information about the message instead of measuring. For this reason, a locking scheme has to be used with great care when composed with other cryptographic primitives.

We also used uncertainty relations to construct quantum identification codes. We proved that it is possible to identify a quantum state of $n$ qubits by communicating $n$ classical bits and $O(\log(1/\epsilon))$ quantum bits. We also presented an efficient encoder for this problem that
6.2. Open questions

uses $O(\log^2(n/\epsilon))$ qubits of communication instead. The main weakness of this result is that the decoder uses a classical description of the state $|\varphi\rangle$ that is in general exponential in the number of qubits of $|\varphi\rangle$. One cannot hope to avoid this difficulty because, as shown by Winter [2004], if Bob was to receive a copy of the quantum state $|\varphi\rangle$, the task of quantum identification becomes the same as the task of transmission of quantum information.

We then considered uncertainty relations that hold even in the presence of quantum side information. For this, we defined QC-extractors which are sets of unitary transformations that have the following property: for any state $\rho_{AE}$ for which $H_{\min}(A|E)_\rho$ is sufficiently larger than $-\log d_A$, applying a typical unitary on $A$ followed by a measurement of some prefix of the output gives an outcome that is almost uniformly distributed and independent of $E$. Such a definition fits in the general framework of the decoupling theorem of Dupuis [2010], Dupuis et al. [2010a] and we use similar techniques to analyse the different constructions we propose; see Table 5.1 for a summary. All these constructions lead to strong min-entropy uncertainty relations. We used them to prove the security of two-party function evaluation under a condition on the capacity of the parties’ storage device to maintain quantum information. We also proved von Neumann entropy uncertainty relations with quantum side information for a full set of mutually unbiased bases, thus generalizing the results of Ivanovic [1992], Sanchez [1993].

6.2 Open questions

We expect to see more applications to quantum information theory of the tools used in the theory of pseudorandomness. An interesting open question is whether these techniques can be helpful in constructing explicit subspaces of highly entangled states. Such subspaces are related to one of the central problems in quantum information theory: the classical capacity of a quantum channel. An explicit construction of such spaces would lead to explicit channels that violate additivity of the minimum output entropy [Hastings 2009, Hayden and Winter 2008], but also explicit protocols for superdense coding of quantum states [Harrow et al. 2004]. As shown by Aubrun et al. [2010, 2011], this problem amounts to finding explicit almost Euclidean sections for matrix spaces endowed with Schatten $p$-norms, which corresponds to the \ell_p norm of the singular values. In addition to the applications in quantum information theory, such almost Euclidean sections are closely related to rank minimization problems for which the nuclear norm heuristic allows exact recovery [Dvijotham and Fazel 2010].

Addressing this question is related to finding explicit constructions of $(0, \epsilon)$-QQ-
extractors with output size close to $n/2$ (which is optimal) and with small (say sublinear in $n$) seed size. In fact, by applying the unitaries of the QQ-extractor in superposition, all input pure states get mapped to highly entangled output states. More generally, it would be very interesting to understand what kinds of sets of unitaries other than unitary two-designs satisfy the decoupling theorem. Is it possible to use a number of unitaries that is smaller than the output dimension? Even non-explicit constructions would be interesting. In the special case of QC-extractors, do the metric uncertainty relations of Chapter 5 remain valid in the presence of quantum side information?

From a computational complexity point of view, I think it would be also interesting to study the hardness of some natural problems related to uncertainty relations. For example, given a set of unitaries as an input, can one compute efficiently how good uncertainty relations they define? Does quantum side information make things significantly harder?

We might also wonder whether the decoupling theorem, or the different notions of quantum extractors defined here have applications to complexity theory, just as classical extractors have applications in derandomization for example.

There is also an intriguing general question on the power of the second moment. We know that pairwise independent permutations are good (classical) extractors. We also know that a full set of mutually unbiased bases — which defines a state 2-design — satisfies good uncertainty relations. In addition, the decoupling theorem says that unitary 2-designs satisfy a strong decoupling statement. All these results are based on a second moment argument. Is there a precise way of unifying these results?

On the cryptography side, are there cryptographic applications of locking schemes? For example, suppose that we authenticate the message before encoding it. Then the receiver can check whether an eavesdropper has altered the encoded message. Conditioned on passing the authentication test, is it true that the state held by the eavesdropper is independent of the message? If this is the case, then the security guarantee would be composable and we could use a locking scheme as a key distribution protocol that only uses communication from Alice to Bob.
Appendices
Appendix A

Deferred proofs

A.1 Existence of metric uncertainty relations

In this section, we prove the lemmas used for proving Theorem 3.3.2.

Lemma 3.3.3 (Average value of $\ell^A_1(\ell^B_2)$ on the sphere). Let $|\varphi\rangle^{AB}$ be a random pure state on $AB$. Then,

$$E\left\{\|\varphi\rangle^{AB}\|_{\ell^A_1(\ell^B_2)}\right\} = \frac{\Gamma(d_B + \frac{1}{2})}{\Gamma(d_B)} \frac{\Gamma(d_Ad_B)}{\Gamma(d_Ad_B + \frac{1}{2})} \geq \sqrt{1 - \frac{1}{d_B} \sqrt{d_A}}.$$ 

where $\Gamma$ is the Gamma function $\Gamma(z) = \int_0^\infty u^{z-1}e^{-u}du$ for $z \geq 0$.

Proof  The presentation uses methods described in [Ball 1997].

Observe that the random variable $\|\varphi\rangle^{AB}\|_{12}$ is distributed as the $\ell^A_1(\ell^B_2)$ norm of a real random vector chosen according to the rotation invariant measure on the sphere $S^{2d_Ad_B-1}$. We define for integers $n$ and $m$ the norm $\ell^m_1(\ell^n_2)$ of a real $n + m$-dimensional vector $\{v_{i,j}\}_{i \in [n], j \in [m]}$ as for the complex case (Definition 3.2.3)

$$\|v\|_{\ell^m_1(\ell^n_2)} = \sum_i \sqrt{\sum_j |v_{i,j}|^2}.$$ 

Note that we only specify the dimension of the systems as the systems themselves are not relevant here. In the rest of the proof, we use $\| \cdot \|_{12}$ as a shorthand for $\| \cdot \|_{\ell^A_1(\ell^B_2)}$. Our objective is to evaluate the expected value $E\{\|\Theta\|_{12}\}$ where $\Theta$ has rotation invariant distribution on the real sphere $S^{s-1}$ and $s = 2d$ with $d = d_Ad_B$. For this, we start by relating the $E\{\|Z\|_{12}\}$ and $E\{\|\Theta\|_{12}\}$ where $Z$ has a standard Gaussian distribution on $\mathbb{R}^s$. By
changing to polar coordinates, we get

\[
E \{ \|Z\|_{12} \} = \int_{\mathbb{R}^s} \|x\|_{12} e^{-\frac{1}{2} \sum_{i=1}^{s} x_i^2 / (2\pi)^{s/2}} dx
\]

\[
= \int_0^\infty \int_{S^{s-1}} \|r\|_{12} e^{-r^2 / 2} \cdot \frac{s\pi^{s/2} d\sigma(\theta)}{\Gamma(s/2 + 1)} r^{s-1} dr
\]

where \(\sigma\) is the normalized Haar measure on \(S^{s-1}\). The term \(\frac{s\pi^{s/2}}{\Gamma(s/2 + 1)}\) is the surface area of the sphere in dimension \(s-1\). Using the equality \(\Gamma(z + 1) = z\Gamma(z)\), we have

\[
\frac{s\pi^{s/2}}{\Gamma(s/2 + 1)} = \frac{2\pi^{s/2}}{\Gamma(s/2)}.
\]

Thus,

\[
E \{ \|Z\|_{12} \} = \frac{2\pi^{s/2}}{(2\pi)^{s/2}\Gamma(s/2)} \int_0^\infty r^s e^{-r^2 / 2} dr \cdot \int_{S^{s-1}} \|\theta\|_{12} d\sigma(\theta)
\]

\[
= \frac{1}{2^{s/2-1}\Gamma(s/2)} \int_0^\infty r^s e^{-r^2 / 2} dr \cdot \int_{S^{s-1}} \|\theta\|_{12} d\sigma(\theta)
\]

We then perform a change of variable \(u = r^2 / 2\):

\[
E \{ \|Z\|_{12} \} = \frac{1}{2^{s/2-1}\Gamma(s/2)} \int_0^\infty (2u)^{(s-1)/2} e^{-u} du \cdot \int_{S^{s-1}} \|\theta\|_{12} d\sigma(\theta)
\]

\[
= \frac{2^{(s-1)/2}\Gamma(s/2)}{2^{s/2-1}\Gamma(s/2)} \cdot \int_{S^{s-1}} \|\theta\|_{12} d\sigma(\theta)
\]

\[
= \frac{\sqrt{2}\Gamma(s/2 + 1)}{\Gamma(s/2)} \cdot E \{ \|\Theta\|_{12} \}. \tag{A.1}
\]

Now, we compute

\[
E \{ \|Z\|_{12} \} = \int_{\mathbb{R}^s} \|x\|_{12} e^{-\frac{1}{2} \|x\|_2^2 / (2\pi)^{s/2}} dx
\]

\[
= \sum_{i=1}^{d_A} \int_{\mathbb{R}^s} \|x_i\|_2 e^{-\frac{1}{2} \|x_i\|_2^2 / (2\pi)^{s/2}} dx_i
\]

where we decomposed \(x = (x_1, \ldots, x_{d_A})\) where \(x_i \in \mathbb{R}^{2d_B}\). As all the terms of the sum are equal

\[
E \{ \|Z\|_{12} \} = d_A \int_{\mathbb{R}^{2d_B}} \|x_0\|_2 e^{-\frac{1}{2} \|x_0\|_2^2 / (2\pi)^{s/2}} dx_0 \left( \int_{\mathbb{R}^{2d_B}} \frac{e^{-\frac{1}{2} \|x_1\|_2^2 / (2\pi)^{s/2}} dx_1}{(2\pi)^d_{d_B}} \right)^{d_A-1}
\]

\[
= d_A \frac{\sqrt{2}\Gamma(d_B + 1/2)}{\Gamma(d_B)} \int_{S^{2d_B-1}} \|\theta\|_2 d\sigma(\theta)
\]

\[
= \frac{\sqrt{2}\Gamma(d_B + 1/2)}{\Gamma(d_B)}.
\]
To get the second equality, we use the same argument as for equation (A.1). We conclude using equation (A.1)

\[
E \left\{ \| \varphi \|_{\ell^A_A(\ell^B_B)} \right\} = E \left\{ \| \Theta \|_{12} \right\} = d_A \frac{\Gamma(d_B + \frac{1}{2})}{\Gamma(d_B)} \cdot \frac{\Gamma(d_A d_B)}{\Gamma(d_A d_B + \frac{1}{2})}.
\]

We now prove the inequality in the statement of the lemma. We use the following two facts about the \( \Gamma \) function: \( \log \Gamma \) is convex and for all \( z > 0, \Gamma(z + 1) = z\Gamma(z) \). The first property can be seen by using Hölder’s inequality for example and the second using integration by parts. Using these properties, we have

\[
\log \Gamma \left( x + \frac{1}{2} \right) \leq \frac{1}{2} \log \Gamma(x) + \frac{1}{2} \log \Gamma(x + 1) = \frac{1}{2} \log \left( x\Gamma(x)^2 \right) = \log \left( \sqrt{x} \Gamma(x) \right).
\]

Thus, \( \frac{\Gamma(x + \frac{1}{2})}{\Gamma(x)} \leq \sqrt{x} \). Similarly, we have \( \frac{\Gamma(x)}{\Gamma(x - \frac{1}{2})} \leq \sqrt{x - \frac{1}{2}} \) which implies that \( \frac{\Gamma(x + \frac{1}{2})}{\Gamma(x)} \geq \sqrt{x - \frac{1}{2}} \) when writing \( \Gamma(x + 1/2) = (x - 1/2)\Gamma(x - 1/2) \).

We conclude that

\[
E \left\{ \| \varphi \|_{\ell^A_A(\ell^B_B)} \right\} \geq d_A \cdot d_B - \frac{1}{2} \frac{1}{\sqrt{d_A d_B}} = \sqrt{d_A} \cdot \sqrt{1 - \frac{1}{2d_B}}.
\]

\[\square\]

**Lemma 3.3.4** (Levy’s lemma). Let \( f : \mathbb{C}^d \to \mathbb{R} \) and \( \eta > 0 \) be such that for all pure states \( |\varphi_1\rangle, |\varphi_2\rangle \) in \( \mathbb{C}^d \),

\[
|f(|\varphi_1\rangle) - f(|\varphi_2\rangle)| \leq \eta \| |\varphi_1\rangle - |\varphi_2\rangle \|_2.
\]

Let \( |\varphi\rangle \) be a random pure state in dimension \( d \). Then for all \( 0 \leq \delta \leq \eta \),

\[
\Pr \{ |f(|\varphi\rangle) - E \{ f(|\varphi\rangle) \} | \geq \delta \} \leq 4 \exp \left( -\frac{\delta^2 d}{4c\eta^2} \right)
\]

where \( c \) is a constant. We can take \( c = 9\pi^2 \).
A.1. Existence of metric uncertainty relations

**Proof** We can instead study the concentration of a Lipschitz function on the real sphere $\mathbb{S}^{2d-1}$. Note that the induced function (that we also call $f$) is still $\alpha$-Lipschitz. Concentration on $\mathbb{S}^{2d-1}$ can be proved in a simple way using concentration of the standard Gaussian distribution. This proof is due to Maurey and Pisier and can be found in [Milman and Schechtman, 1986, Appendix V]. Specifically, using [Milman and Schechtman, 1986, Corollary V.2], we get

$$\Pr \{|f(Z) - \mathbb{E}\{f(Z)\}| \geq t\} \leq 2 \exp \left(-\frac{\delta^2(2d)}{18\pi^2\eta^2}\right) + 2 \exp \left(-\frac{2d}{2\pi^2}\right) \leq 4 \exp \left(-\frac{\delta^2d}{9\pi^2\eta^2}\right).$$

In the notation of the proof of [Milman and Schechtman, 1986, Corollary V.2], we have set $\delta = 1/2$. This can be done because using the same arguments as in the proof of Lemma 3.3.3, we can show that the expected $\ell_2$ norm of the standard Gaussian distribution in dimension $n$ at least $\sqrt{2}\sqrt{n - \frac{1}{2}} > \sqrt{n}$ for $n \geq 2$.

We used this version of Levy’s lemma because it has an elementary proof and it gives directly the concentration about the expected value. Different versions involving the median of $f$ and giving better constants can be found in [Milman and Schechtman, 1986, Corollary 2.3] or [Ledoux, 2001, Proposition 1.3] for example.  

**Lemma 3.3.6** ($\delta$-net). Let $\delta \in (0, 1)$. There exists a set $\mathcal{N}$ of pure states in $\mathbb{C}^d$ with $|\mathcal{N}| \leq (3/\delta)^{2d}$ such that for every pure state $|\psi\rangle \in \mathbb{C}^d$ (i.e., $\|\psi\|_2 = 1$), there exists $|\tilde{\psi}\rangle \in \mathcal{N}$ such that

$$\|\psi\rangle - |\tilde{\psi}\rangle\|_2 \leq \delta.$$

**Proof** A proof can be found in [Hayden et al., 2004, Lemma II.4]. We repeat it here for completeness. Let $\mathcal{N}$ be a maximal set of pure states satisfying $\|\psi_1\rangle - |\psi_2\rangle\|_2 \geq \delta$ for all pure states $|\psi_1\rangle, |\psi_2\rangle \in \mathcal{N}$. This set can be constructed iteratively by adding at each step a state that is at distance at least $\delta$ from all states already in the set. First, we show that this procedure terminates by bounding the size of such a set. We do this using a volume argument. For this it is simpler to look at vectors $|\psi\rangle \in \mathcal{N}$ as real vectors in dimension $2d$. The open balls of radius $\delta/2$ centered at each $|\psi\rangle \in \mathcal{N}$ are disjoint and are contained in the open ball of radius $1 + \delta/2$ centered at the origin. Therefore,

$$|\mathcal{N}| \left(\frac{\delta}{2}\right)^{2d} \leq \left(\frac{3}{2}\right)^{2d}.$$
We conclude by observing that such a set has the desired property. In fact, if there exists a state \( |\psi\rangle \in \mathcal{H} \) such that for all \( |\tilde{\psi}\rangle \in \mathcal{N} \), \( |||\psi\rangle - |\tilde{\psi}\rangle||_2 > \delta \), then \( |\psi\rangle \) can be added to \( \mathcal{N} \) and contradict the fact that \( \mathcal{N} \) is maximal.

\[ \square \]

**Lemma 3.3.5** (Concentration of the average). Let \( a, b \geq 1 \), \( \delta \in (0, 1) \) and \( t \) a positive integer. Suppose \( X \) is a random variable with 0 mean satisfying the tail bounds

\[
\Pr \{ X \geq \eta \} \leq a e^{-b\eta^2} \quad \text{and} \quad \Pr \{ X \leq -\eta \} \leq a e^{-b\eta^2}.
\]

Let \( X_1, \ldots, X_t \) be independent copies of \( X \). Then if \( \delta^2 b \geq 16a^2 \pi \),

\[
\Pr \left\{ \left| \frac{1}{t} \sum_{k=1}^{t} X_k \right| \geq \delta \right\} \leq \exp \left( -\frac{\delta^2 bt}{2} \right).
\]

**Proof** For any \( \lambda > 0 \), using Markov’s inequality

\[
\Pr \left\{ \sum_{k=1}^{t} X_k \geq t\delta \right\} = \Pr \left\{ \exp \left( \lambda \sum_{k=1}^{t} X_k \right) \geq \exp (\lambda t\delta) \right\} \leq \mathbb{E} \left\{ \exp \left( \lambda \sum_{k=1}^{t} X_k \right) \right\} e^{-\lambda t \delta} = \mathbb{E} \left\{ e^{\lambda X} \right\} e^{-\lambda t \delta}.
\]

We now bound the moment generating function \( \mathbb{E} \left\{ e^{\lambda X} \right\} \) of \( X \) using the tail bounds.

\[
\mathbb{E} \left\{ e^{\lambda X} \right\} = \int_{0}^{\infty} \Pr \left\{ e^{\lambda X} \geq u \right\} du \leq 1 + \int_{1}^{\infty} a \exp \left( -\frac{b \ln^2 u}{\lambda^2} \right) du = 1 + a \int_{0}^{\infty} \exp \left( -\frac{b z^2}{\lambda^2} \right) e^z dz
\]
A.1. Existence of metric uncertainty relations

by making the change of variable $z = \log u$.

$$
\mathbb{E} \{ e^{\lambda X} \} \leq 1 + a \int_0^\infty \exp \left(-\frac{b}{\lambda^2} \left(z - \frac{\lambda^2}{2b}\right)^2 + \frac{\lambda^2}{4b}\right) dz \\
\leq 1 + a \exp \left(\frac{\lambda^2}{4b}\right) \int_{-\infty}^\infty \exp \left(-\frac{b}{\lambda^2} \left(z - \frac{\lambda^2}{2b}\right)^2\right) dz \\
= 1 + a \exp \left(\frac{\lambda^2}{4b}\right) \frac{\lambda}{\sqrt{2b}} \int_{-\infty}^\infty \exp \left(-\frac{u^2}{2}\right) du \\
= 1 + a \frac{\sqrt{2\pi\lambda}}{\sqrt{2b}} \cdot \exp \left(\frac{\lambda^2}{4b}\right) \\
\leq 2 \max \left(1, a \frac{\sqrt{\pi\lambda}}{\sqrt{b}} \cdot \exp \left(\frac{\lambda^2}{4b}\right)\right).
$$

We choose $\lambda = 2\delta b$ (this is not the optimal choice but it makes expressions simpler).

$$
\Pr \left\{ \sum_{k=1}^t X_k \geq t\delta \right\} \leq \max \left(2^t, \left(2a \frac{\sqrt{\pi\lambda}}{\sqrt{b}}\right)^t \cdot \exp \left(\frac{\lambda^2 t}{4b}\right) \cdot \exp \left(-\lambda t\delta\right)\right) \\
= \max \left\{ \exp \left(-2\delta^2 bt + t \ln 2\right), \exp \left(\delta^2 bt - 2\delta^2 bt + t \ln(4\sqrt{\pi}\delta\sqrt{b})\right)\right\} \\
= \max \left\{ \exp \left((-2\delta^2 b + \ln 2) t\right), \exp \left((-\delta^2 b + \ln(4\sqrt{\pi}\delta\sqrt{b}) t\right)\right\}.
$$

Claim. For all $c \geq 1$ and $x \geq c$

$$
\frac{1}{2} \ln(cx) - x \leq -\frac{x}{2}.
$$

The function $x \mapsto \frac{x}{2} - \frac{1}{2} \ln(cx)$ is increasing for $x \geq 1$. It suffices to show that it is nonnegative for $x = c$. To see that, we differentiate the function $y \mapsto y - \ln(y^2)$ to prove that for all $y \geq 1$, we have $y - \ln(y^2) \geq 0$. This proves the claim.

Using this inequality, we have for $\delta^2 b \geq 16a^2\pi$,

$$
-\delta^2 b + \ln(4\sqrt{\pi}\delta\sqrt{b}) \leq -\frac{\delta^2 b}{2} \quad \text{and} \quad -2\delta^2 b + \ln 2 \leq -\frac{\delta^2 b}{2}.
$$

Finally,

$$
\Pr \left\{ \sum_{k=1}^t X_k \geq t\delta \right\} \leq \exp \left(-\frac{\delta^2 bt}{2}\right).$$
In order to prove the existence of strong permutation extractors with good parameters, we use the construction of Guruswami, Umans, and Vadhan [2009] which is inspired by list decoding. Their main construction is a lossless condenser based on Parvaresh-Vardy codes. Using this condenser, they build an explicit extractor with good parameters. However, this lossless condenser based on Parvaresh-Vardy codes does not seem to be easily extended into a permutation condenser. The same paper also presents a lossy condenser based on Reed-Solomon codes, which can indeed be transformed into a permutation condenser. This permutation condenser can then be used in the extractor construction instead of the lossless condenser giving a strong permutation extractor. In this section, we describe this construction. For completeness, we reproduce most of the proof here, except the results that are used exactly as stated in Guruswami et al. [2009].

A condenser is explicit if for all \( y \in S \), both \( P_y \) and \( P_y^{-1} \) are computable in polynomial time.

The following theorem describes the condenser that will be used as a building block in the extractor construction. It is an analogue of Theorem 7.2 in Guruswami et al. [2009].
Theorem A.2.3. For all positive integers \( n \) and \( \ell \leq n \), as well as \( \alpha, \epsilon \in (0, 1/2) \), there exists an explicit family of permutations \( \{RS_y\}_{y \in S} \) of \( \mathbb{F}_2^n \) that is an
\[
(nt, (\ell + 1)t) \rightarrow_\epsilon (\ell t, (1 - \alpha)\ell t - 4)
\]
strong permutation condenser with \( t = \lceil 1/\alpha \cdot \log(24n^2/\epsilon) \rceil \) and \( \log |S| \leq t \). Moreover, the functions \((x, y) \mapsto RS_y(x)\) and \((x, y) \mapsto RS_y^{-1}(x)\) can be computed by a circuit of size \( O(n \text{ polylog}(n/\epsilon)) \).

Proof. Set \( q = 2^t \) and \( \epsilon_0 = \epsilon/6 \). Consider the function \( C' : \mathbb{F}_q^n \times \mathbb{F}_q \rightarrow \mathbb{F}_q^{\ell+1} \) defined by
\[
C'(f, y) = [y, f(y), f(\zeta y), \ldots, f(\zeta^{\ell-1} y)]
\]
where \( \mathbb{F}_q^n \) is interpreted as the set of polynomials over \( \mathbb{F}_q \) of degree at most \( n - 1 \) and \( \zeta \) is a generator of the multiplicative group \( \mathbb{F}_q^* \). First, we compute the input and output sizes in terms of bits. The inputs can be described using \( \log |\mathbb{F}_q^n| = n \log q = nt \) bits, the seed using \( \log |\mathbb{F}_q| = t \) bits and the output using \( \log |\mathbb{F}_q^{\ell+1}| = (\ell + 1)t \). Using [Guruswami et al., 2009, Theorem 7.1], for any integer \( h \), \( C' \) is a
\[
\left( nt, \log \left( \frac{q^\ell - 1}{\epsilon_0} \right) \right) \rightarrow_{2\epsilon_0} \left( \ell t + t, \log \left( \frac{Ah^{\ell} - 1}{2\epsilon_0} \right) \right)
\]
condenser where \( A \stackrel{\text{def}}{=} \epsilon_0 q - (n - 1)(h - 1)\ell \). We now choose \( h = \lceil q^{1-\alpha} \rceil \). As \( q \geq (4n^2/\epsilon_0)^{1/\alpha} \), we have \( A \geq \epsilon_0 q - n^2h \geq \epsilon_0 q - \epsilon_0 q^{\alpha}/4 \cdot (q^{1-\alpha} + 1) \geq \epsilon_0 q/2 \). Thus, we can compute the bounds we obtain on the condenser \( C' \):
\[
\log \left( \frac{q^\ell - 1}{\epsilon_0} \right) = \ell t + \log(1/\epsilon_0) \leq (\ell + 1)t
\]
and
\[
\log \left( \frac{Ah^{\ell} - 1}{2\epsilon_0} \right) = \log \left( \frac{Ah^{\ell}}{2\epsilon_0} \right) + \log \left( 1 - \frac{1}{Ah^{\ell}} \right)
\geq \log(q/4) + \ell \log h - 1
\geq t + (1 - \alpha)\ell t - 3.
\]
Plugging these values in equation (A.2), we get that \( C' \) is a
\[
(nt, (\ell + 1)t) \rightarrow_{2\epsilon_0} (\ell t + t, (1 - \alpha)\ell t + t - 3))
\]
condenser.
Observe that the seed $y$ is part of the output of the condenser. As we want to construct a strong condenser, we do not consider the seed as part of the output of the condenser. For this, we define $C : \mathbb{F}_q^n \times \mathbb{F}_q \to \mathbb{F}_q^t$ by $C(f, y) = [f(y), \ldots, f(\zeta^{t-1}y)]$. Moreover, as will be clear later when we try to build a permutation condenser, we take the seed to be uniform on $S \overset{\text{def}}{=} \mathbb{F}_q^* = \mathbb{F}_q - \{0\}$ instead of being uniform on the whole field $\mathbb{F}_q$. Note that this increases the error of the condenser by at most $2^{-t} \leq \epsilon_0$ (because one can choose $U_{\mathbb{F}_q^*} = U_{\mathbb{F}_q}$ with probability $1 - 2^{-t}$). Here and in the rest of this proof, we will be using Doeblin’s coupling lemma (see Chapter 2).

Equation (A.3) then implies that if $X$ has min-entropy at least $(\ell + 1)t$ and $U_S$ is uniform on $S$, then the distribution of $(U_S, C(X, U_S))$ is $3\epsilon_0$-close to a distribution with min-entropy at least $(1 - \alpha)\ell t + t - 3$. Let $Y \in S$ and $Z \in \{0, 1\}^{(\ell + 1)t}$ be random variables such that $H_{\text{min}}(Y, Z) \geq (1 - \alpha)\ell t + t - 3$ and $(U_S, C(X, U_S)) = (Y, Z)$ with probability at least $1 - 3\epsilon_0$. If $Y$ was uniformly distributed on $S$, then it would follow directly that for all $y \in S$, $H_{\text{min}}(Z|Y = y) \geq (1 - \alpha)\ell t - 3$. However, $Y$ is not necessarily uniformly distributed. We define a new random variable $Z'$ by

$$Z' = \begin{cases} Z & \text{if } Y = U_S \\ U' & \text{if } Y \neq U_S \end{cases}$$

where $U'$ is uniformly distributed on $\{0, 1\}^{(\ell + 1)t}$ and independent of all the other random variables. We have for any $z \in \{0, 1\}^{(\ell + 1)t}$ and $y \in S$,

$$\Pr \{Z' = z|U_S = y\} = \frac{1}{\Pr \{U_S = y\}} \left(\Pr \{Z' = z, Y = y, Y = U_S\} \right.
+ \Pr \{Z' = z, U_S = y, Y \neq U_S\})
\leq \frac{1}{\Pr \{U_S = y\}} \left(2^{-(1-\alpha)\ell t - t + 3} + 2^{-(\ell + 1)t}, \frac{1}{|S|}\right)
\leq 2 \cdot 2^{-(1-\alpha)\ell t + 3}.$$ 

Moreover, we have $(U_S, C(X, U_S)) = (U_S, Z')$ with probability at least $1 - 6\epsilon_0$.

We conclude that $C$ is a

$$(nt, (\ell + 1)t) \rightarrow_\epsilon (\ell t, (1 - \alpha)\ell t - 4)) \quad (A.4)$$

strong condenser.

To define our permutation condenser, we set the first $n' = \ell t$ bits $RS_y^C(x)$ of $RS_y(x)$ to be $RS_y^C(x) = C(x, y)$. We then define the remaining bits by $RS_y^R(f) = \lfloor f(\zeta^t y), \ldots, f(\zeta^{n-1} y) \rfloor$. As $q \geq n - 1$ and $\zeta$ is a generator of $\mathbb{F}_q^*$, the elements
y, ζy, . . . , ζ^{n-1}y are distinct provided y ≠ 0. So for y ≠ 0, (RS^C_y, RS^R_y)(f) is the evaluation of the polynomial f of degree at most n − 1 in n distinct points. Thus, f → RS_y(f) is a bijection in \( \mathbb{F}_q^n \) for all y ≠ 0. This is why the value 0 for the seed was excluded earlier.

Concerning the computation of the functions RS^C_y and RS^R_y, they only require the evaluation of a polynomial on elements of the finite field \( \mathbb{F}_q \). Computations in the finite field \( \mathbb{F}_q \) can be performed efficiently by finding an irreducible polynomial of degree log q over \( \mathbb{F}_2 \) and doing computations modulo this polynomial. In fact, finding an irreducible polynomial of degree log q over \( \mathbb{F}_2 \) can be done in time polynomial in log q (see for example Shoup [1990] for a deterministic algorithm and Corollary 14.43 in the book von zur Gathen and Gerhard [1999] for a simpler randomized algorithm). Since addition, multiplication and evaluation of a polynomial on elements of the finite field \( \mathbb{F}_q \) can be done in time polynomial in log q (see for example Shoup [1992] shows the existence of a generator \( \zeta \) of the group \( \mathbb{F}_q^* \). For example, Theorem 1.1 in Shoup [1992] shows the existence of a generator \( \zeta \) of the group \( \mathbb{F}_q^* \). For example, Theorem 1.1 in Shoup [1992] shows the existence of a generator \( \zeta \) of the group \( \mathbb{F}_q^* \).

To evaluate RS_y at a polynomial f, we compute the field elements y, ζy, . . . , ζ^{n-1}y, and then evaluate the polynomial f on these points. Using a fast multipoint evaluation, this step can be done in \( O(n \log n) \) number of operations in \( \mathbb{F}_q \) (see Corollary 10.8 in von zur Gathen and Gerhard [1999]). Moreover, given a list \([f(y), . . . , f(ζ^{n-1}y)]\) for y ≠ 0, we can find f by fast interpolation in \( \mathbb{F}_q[X] \) (see Corollary 10.12 in von zur Gathen and Gerhard [1999]). As a result \( RS_y^{-1} \) can also be computed in \( O(n \log n) \) operations in \( \mathbb{F}_q \).

This condenser will be composed with other extractors, the following lemma shows how to compose condensers.

**Lemma A.2.4** (Composition of strong permutation condensers). Let \((P_{1,y_1})_{y_1 \in S_1} \) be an \((n, k) \to (n', k')\) strong permutation condenser and \((P_{2,y_2})_{y_2 \in S_2} \) be an \((n', k') \to (n'', k'')\) strong permutation condenser. Then \((P_y)_{y = (y_1, y_2) \in S_1 \times S_2} = (P^C_y, P^R_y) \) where \( P^C_{y_1,y_2} = P^C_{2,y_2} \circ P^C_{1,y_1} \) and \( P^R_{y_1,y_2} = (P^R_{2,y_2} \circ P^C_{1,y_1}) \cdot P^R_{1,y_1} \) is an \((n, k) \to (n'', k'')\) strong permutation extractor.

**Proof** \( P_y \) is clearly a permutation of \( \{0, 1\}^n \). We only need to check that \( P^C_y \) is a strong condenser. By definition, if \( H_{\min}(X) \geq k \), \((U_{S_1}, P^C_{1,U_{S_1}}(X)) \) is \( \epsilon \)-close to \((U_{S_1}, Z) \) where \( Z|_{U_{S_1} = y_1} \) has min-entropy at least \( k' \). Now putting Z into the condenser \( P^C_2 \), we get that for any \( y_1 \), \((U_{S_2}, P^C_{2,U_{S_2}}(Z|_{U_{S_1}})) \) is \( \epsilon \)-close to \((U_{S_2}, Z_2) \) where \( Z_2|_{U_{S_2} = y_2} \) has min-entropy at least \( k'' \) for any \( y_2 \in S_2 \). Thus, \( Z_2|_{U_{S_1}U_{S_2} = y_1y_2} \) has min-entropy at least \( k'' \). Moreover, by the
triangle inequality, we have \( \Delta \left( (U_{S_1}, U_{S_2}, P^C_{U_{S_1} U_{S_2}}(X)), (U_{S_1}, U_{S_2}, Z_2) \right) \leq 2\epsilon. \)

Next, we present one of the standard extractors that are used as a building block in many constructions.

**Lemma A.2.5** (“Leftover Hash Lemma” extractor \[Impagliazzo et al., 1989\]). For all positive integers \( n \) and \( k \leq n \), and \( \epsilon > 0 \), there exists an explicit family \((P_y)_{y \in S}\) of permutations of \( \{0,1\}^n \) that is an \((n,k) \to \epsilon m\) strong permutation extractor with \( \log |S| = \log(2^n - 1) \) and \( m \geq k - 2 \log(2/\epsilon) \).

**Proof** We view \( \{0,1\}^n \) as the finite field \( \mathbb{F}_{2^n} \) and the set \( S = \mathbb{F}_{2^n}^* \). We then define the permutation \( P_y(x) = x \cdot y \) where the product \( x \cdot y \) is taken in the field \( \mathbb{F}_{2^n} \). The family of functions \( P_y \) is pairwise independent. Applying the Leftover Hash Lemma \[Impagliazzo et al., 1989\], we get that if \( Y \) uniform on \( \mathbb{F}_{2^n} \), the distribution of the first \( \lceil k - 2 \log(1/\epsilon) \rceil \) bits of \( P_Y(X) \) together with \( Y \) is \( \epsilon \)-close to uniform. Now if \( U_S \) is only uniform in \( \mathbb{F}_{2^n}^* \), \((U_S, P_{U_S}(X))\) is \( \epsilon + 2^{-n} \)-close to the uniform distribution. The result follows from the fact that we can suppose \( \epsilon \geq 2^{-n} \) (otherwise, \( k - 2 \log(1/\epsilon) \leq 0 \) and the theorem is true).

The problem with this extractor is that it uses a seed that is as long as the input. Next, we introduce the notion of a block source.

**Definition A.2.6** (Block source). \( X = (X_1, X_2, \ldots, X_s) \) is a \((k_1, k_2, \ldots, k_s)\) block source if for every \( i \in \{1, \ldots, s\} \) and \( x_1, \ldots, x_i-1, X|_{X_1=x_1, \ldots, X_{i-1}=x_{i-1}} \) is a \( k_i\)-source. When \( k_1 = \cdots = k_s = k \), we call \( X \) a \( s \times k \) source.

A block source has more structure than a general source. However, for a source of large min-entropy \( k \) (or equivalently with small entropy deficiency \( \Delta = n - k \)), one does not lose too much entropy by viewing a general source as a block source where each block has entropy deficiency roughly \( \Delta \). See \[Guruswami et al., 2009\] Corollary 5.9] for a precise statement.

**Lemma A.2.7** (\[Guruswami et al., 2009, Lemma 5.4\]). Let \( s \) be a (constant) positive integer. For all positive integers \( n \) and \( \ell \leq n \) and all \( \epsilon > 0 \), setting \( t = \lceil 8s \log(24n^2 \cdot (4s + 1)/\epsilon) \rceil \), there is an explicit family \( \{L_y\}_{y \in S} \) of permutations of \( \{0,1\}^n \) that is an \((n, 2\ell t) \to \epsilon \ell t\) strong permutation extractor with \( \log |S| \leq 2\ell t/s + t \).

**Proof** As the extractor is composed of many building blocks, each generating some error, we define \( \epsilon_0 = \epsilon/(4s + 1) \) where \( \epsilon \) is the target error of the final extractor. The
idea is to first apply the condenser $RS$ of Theorem A.2.3 with $\alpha = \frac{1}{8s}$ to obtain a string $X' = RS^C(X, U_{F^*})$ of length $n' = (2\ell - 1)t$ which is $\epsilon_0$-close to a $k'$-source where

$$k' = \left(1 - \frac{1}{8s}\right) (2\ell - 1)t - 4$$

The entropy deficiency $\Delta$ of this $k'$-source can be bounded by $\Delta = n' - k' \leq \frac{(2\ell - 1)t}{8s} + 4$. Then, we partition $X' = (X'_1, \ldots, X'_{2s})$ (arbitrarily) into $2s$ blocks of size $n'' = \lfloor n'/2s \rfloor$ or $n'' + 1$. Using [Guruswami et al., 2009, Corollary 5.9], $(X'_1, \ldots, X'_{2s})$ is $2s\epsilon_0$-close to some $2s \times k''$-source where $k'' = (n'' - \Delta - \log(1/\epsilon_0))$.

We have $\Delta \leq \ell t/(4s) + 3 \leq \ell t/(3s)$ for $n$ large enough. Thus,

$$k'' \geq \frac{2t}{2s} - \frac{\ell t}{3s} - \log(1/\epsilon_0) = \frac{2}{3s} \ell t - \log(1/\epsilon_0).$$

We can then apply the extractor of Lemma A.2.5 to all the $2s$ blocks using the same seed of size $n'' + 1$. Note that we can reuse the same seed because we have a strong extractor and the seed is independent of all the blocks. This extractor extracts almost all the min-entropy of the sources. More precisely, if we input to this extractor a $2s \times k''$-source, the output distribution is $2s\epsilon_0$-close to $m$ uniform bits where

$$m \geq 2s \cdot (k'' - 2\log(2/\epsilon_0)) \geq \frac{4}{3} \ell t - 6s \log(2/\epsilon_0) \geq \ell t.$$

Overall, the output of this extractor is $\epsilon_0 + 2s\epsilon_0 + 2s\epsilon_0 = \epsilon$-close to the uniform distribution on $m$ bits.

It only remains to show that the extractor we just described is strong and can be extended to a permutation. This follows from Lemma A.2.4 and the fact the condensers (coming from Theorem A.2.3 and Lemma A.2.5) are strong permutation condensers. \qed

**Remark.** As pointed out in [Guruswami et al., 2009], a stronger version of this lemma (i.e., with larger output) can be proved by using the condenser of Theorem A.2.3 and the high min-entropy extractor in [Goldreich and Wigderson, 1997] with a Ramanujan expander (for example, the expander of [Lubotzky et al., 1988]). This construction can also give a strong permutation extractor. However, using this extractor would slightly complicate the exposition and does not really influence the final extractor construction presented in Theorem 3.4.5.

The following lemma basically says that the entropy is conserved by a permutation extractor. It is an adapted version of [Raz et al., 1999, Lemma 26].
Lemma A.2.8. Let \( \{P_y\}_{y \in S} \) be a \((n, k) \rightarrow \epsilon \) \( m \)-strong permutation extractor. Let \( X \) be a \( k \)-source, then \((U_S, P^E_{U_S}(X), P^R_{U_S}(X))\) is \(2\epsilon\)-close to \((U'_S, U'_{\{0,1\}^m}, W)\) where \(U'_S\) and \(U'_{\{0,1\}^m}\) are independent and uniformly distributed over \(S\) and \(\{0,1\}^m\) respectively, and for all \(y \in S, z \in \{0,1\}^m\)

\[
H_{\min}(W|(U'_S, U'_{\{0,1\}^m}) = (y, z)) \geq k - m - 1.
\]

Proof As \( \{P^E_y\} \) is a strong extractor, there exist random variables \(U'_S\) and \(U'_{\{0,1\}^m}\) uniformly distributed on \(S\) and \(\{0,1\}^m\) such that \(\Pr\left\{(U_S, P^E_{U_S}(X)) \neq (U'_S, U'_{\{0,1\}^m})\right\} \leq \epsilon\). Define \(\Gamma = \{(y, z) \in S \times \{0,1\}^m : \Pr\left\{P^E_y(X) = z\right\} < \frac{1}{2} \cdot 2^{-m}\}\). We have for every \((y, z) \notin \Gamma\) and \(x \in \{0,1\}^{n-m}\),

\[
\Pr\left\{P^R_y(X) = x | P^E_y(X) = z\right\} \leq \frac{\Pr\left\{P^R_y(X) = x, P^E_y(X) = z\right\}}{2^{-m-1}} \\
\leq 2^{m+1} \Pr\left\{X = P^{-1}_y(x, z)\right\} \\
\leq 2^{-(k-m-1)}.
\]

We then show that \(\Pr\left\{(U_S, P^E_{U_S}) \in \Gamma\right\} \leq \epsilon\). Using the fact that \(\{P^E_y\}\) is a strong extractor, we have

\[
\left|\Pr\left\{U'_S, U'_{\{0,1\}^m} \in \Gamma\right\} - \Pr\left\{(U_S, P^E_{U_S}) \in \Gamma\right\}\right| \leq \epsilon.
\]

But recall that, by definition of \(\Gamma\), \(\Pr\left\{(U_S, P^E_{U_S}) \in \Gamma\right\} < \frac{1}{2} \Pr\left\{(U'_S, U'_{\{0,1\}^m}) \in \Gamma\right\}\), so we get

\[
\Pr\left\{(U_S, P^E_{U_S}) \in \Gamma\right\} \leq \epsilon.
\]

Finally we define

\[
W = \begin{cases} 
  P^R_{U_S}(X) & \text{if } (U_S, P^E_{U_S}(X)) \notin \Gamma \\
  U^* & \text{if } (U_S, P^E_{U_S}(X)) \in \Gamma
\end{cases}
\]

where \(U^*\) is uniform on \(\{0,1\}^{n-m}\) and independent of all other random variables. We conclude by observing that with probability at least \(1 - 2\epsilon\), we have \((U_S, P^E_{U_S}(X)) = (U'_S, U'_{\{0,1\}^m})\) and \(P^R_{U_S}(X) = W\). \(\square\)

We then combine these results to obtain the desired extractor. The proof of the following theorem closely follows [Guruswami et al. 2009 Theorem 5.10] but using the lossy condenser presented in Theorem [A.2.3] and making small modifications to obtain a permutation extractor.
Theorem A.2.9. For all integers \( n \geq 1 \), all \( \epsilon \in (0, 1/2) \), and all \( k \in [200 \lfloor 200 \log(24n^2/\epsilon) \rfloor, n] \) there is an explicit \((n, k) \rightarrow_{\epsilon} \lfloor k/4 \rfloor \) strong permutation extractor \( \{P_y\}_{y \in S} \) with \( \log |S| \leq 200 \lfloor 200 \log(24n^2/\epsilon) \rfloor \). Moreover, the function \((x, y) \mapsto P_y(x)\) can be computed by a circuit of size \( O(n \log(1/\epsilon)) \).

Proof. If \( n \leq 2 \cdot 10^6 \), we can use the extractor of Lemma A.2.7 with \( s = 200 \) and \( \ell \geq 1 \) such that \( 2\ell t \leq k \leq 2(\ell + 1)t \). This gives an extractor whose seed has size \( \frac{k}{200} \leq 10^4 \leq 200 \lfloor 200 \log(24n^2/\epsilon) \rfloor \) and that extracts \( \ell t \geq \frac{1}{4} \cdot 2(\ell + 1)t \geq \frac{k}{4} \) bits, so the statement still holds true. In the rest of the proof, we assume \( n > 2 \cdot 10^6 \).

The idea of the construction is to build for an integer \( i \geq 0 \) an explicit \((n, 2^i \cdot 8d) \rightarrow_{\epsilon} 2^{i-1} \cdot 8d \) extractor using \( d \) bits of seed by induction on \( i \). Fix \( t(\epsilon) = \lfloor 200 \log(24n^2/\epsilon) \rfloor \) and \( d(\epsilon) = 200t(\epsilon) \). The induction hypothesis for an integer \( i \geq 0 \) is as follows: For all integers \( i' \leq i \) and \( n \) and \( \epsilon > 0 \), there is an explicit

\[
(n, 2^{i'} \cdot 8d(\epsilon)) \rightarrow_{\epsilon} 2^{i'-1} \cdot 8d(\epsilon)
\]

strong permutation extractor with seed size \( d(\epsilon) \). This extractor is called \( \{P_y^{(i')}\}_{y \in S_{i'}} \).

For both \( i = 0 \) and \( i = 1 \), we can use the extractor of Lemma A.2.7 with \( s = 20 \). For \( i \in \{0, 1\} \), this gives an extractor with seed \( \frac{2^i \cdot 8d(\epsilon)/81}{20} + t \leq \frac{16}{20} d(\epsilon) + \frac{16}{20} \cdot 200 \lfloor 200 \log(81) \rfloor \leq d(\epsilon) \).

We now show for \( i \geq 2 \) how to build the extractor \( \{P_y^{(i)}\} \) using the extractors \( \{P_y^{(i')}\} \) for \( i' < i \). Using the induction hypothesis, we construct the following extractor, which will be applied four times to extract the necessary random bits to prove the induction step. The choice of the form of the min-entropy values will become clear later. Set \( \epsilon_0 = \epsilon/20 \).

Claim. There exists an

\[
(n, 2^i \cdot 4.5d(\epsilon_0)) \rightarrow_{5\epsilon_0} 2^i \cdot d(\epsilon_0)
\]

strong permutation extractor \( \{Q_y\}_{y \in T} \) with seed size \( \log |T| \leq \frac{d(\epsilon_0)}{8} \).

To prove the claim, we start by applying the condenser of Theorem A.2.3 with \( \alpha = 1/200 \) and \( \epsilon = \epsilon_0 \) (so we use a seed of size \( t(\epsilon_0) \)). The output \( X' \) of size at most \( 2^i \cdot 4.5d(\epsilon_0) \) is then \( \epsilon_0 \)-close to having min-entropy at least \( (1 - \alpha)2^i \cdot 4.5d(\epsilon_0) - t(\epsilon_0) \). The entropy deficiency of this distribution is \( \alpha 2^i \cdot 4.5d(\epsilon_0) + \frac{d(\epsilon_0)}{200} \leq \frac{2^i \cdot 4.5d(\epsilon_0)}{100} \). We then divide \( X' \) into two equal blocks \( X' = (X'_1, X'_2) \), and we know that it is \( 2\epsilon_0 \)-close to being a \( 2 \times k' \)-source for

\[
k' = \frac{2^i \cdot 4.5d(\epsilon_0)}{2} - \frac{2^i \cdot 4.5d(\epsilon_0)}{100} - \log(1/\epsilon_0) \geq \left( \frac{49}{100} \cdot 2^i \cdot 4.5 - \frac{1}{200} \right) d(\epsilon_0)
\]
Figure A.1: The extractor $Q$ is obtained by first applying the condenser of Theorem A.2.3 and decomposing the output into two parts. The Leftover Hash Lemma extractor (Lemma A.2.7) is applied to the first half and its output is used as a seed for the extractor $\{P_y^{(i-2)}\}$ coming from the induction hypothesis.

as $\log(1/\epsilon_0) \leq t(\epsilon_0) = \frac{d(\epsilon_0)}{200}$. For the extractors we will apply next to this source, we should note that $k' \geq 2d(\epsilon_0)$ and that $2^i \cdot 4d(\epsilon_0) \leq k' < 2^i \cdot 8d(\epsilon_0)$.

We now apply the extractor of Lemma A.2.7 to $X'_1$ (viewed as a $2d(\epsilon_0)$-source) using a seed of size $\frac{2d(\epsilon_0)}{20}$ and obtaining $X''$ that is $\epsilon_0$ close to uniform on $d(\epsilon_0)$ bits. We then use the extractor $\{P_y^{(i-2)}\}$ obtained by induction for $i - 2$ to the $X'_2$ (of size $2^i \cdot 4.5d(\epsilon_0) \leq n$) with seed $X''$ (of size $d(\epsilon_0)$): it is an $(n, 2^{i-2} \cdot 8d(\epsilon_0)) \rightarrow_{\epsilon_0} 2^i \cdot d(\epsilon_0)$ permutation extractor.

The construction is illustrated in Figure A.1. Note that the number of bits of the seed is $\log |T| \leq t(\epsilon_0) + \frac{2d(\epsilon_0)}{20} \leq \frac{d(\epsilon_0)}{8}$. This concludes the proof of the claim.
The permutation extractor $\{Q_y\}$ described in the claim is applied four times with independent seeds in order to extract $2^{i-1} \cdot 8d(\epsilon)$ random bits.

The source $X$ we begin with is a $2^i \cdot 8d(\epsilon_0)$-source. But we have $2^i \cdot 8d(\epsilon) \geq 2^i \cdot 8d(\epsilon_0) - 2^i \cdot 8 \cdot 200^2 \log 20 \geq 2^i \cdot 4.5d(\epsilon_0)$ so that we can apply the permutation extractor $(Q_y)_{y \in T}$ of the claim. We obtain $Q^E_{U_T}(X)$ which is $\epsilon_0$-close to $2^i \cdot d(\epsilon_0)$ random bits. As $Q^E$ is part of a permutation extractor, the remaining entropy is not lost: it is in $Q^R_{U_T}(X)$. More precisely, applying Lemma A.2.8, we get $Q^R_{U_T}(X)$ is $\epsilon_0$-close to a source of min-entropy at least $2^i \cdot 8d(\epsilon) - 2^i \cdot 8d(\epsilon_0) - 1$. As $2^i \cdot 8d(\epsilon) - 2^i \cdot d(\epsilon_0) - 1 \geq 2^i \cdot 4.5d(\epsilon_0)$, we can apply the extractor $\{Q_y\}_{y \in T}$ of the claim to this source. Note that the input size has decreased but this only makes it easier to extract random bits as one can always encode in part of the input space. To apply $Q$, we use a fresh new seed that outputs a bit string that is close to uniform on $2^{i-3} \cdot 8d(\epsilon_0)$ bits and the remaining entropy can be found in the $R$ register. We apply this procedure four times in total as shown in Figure A.2. Note that the reason we can apply it four times is that at the last application $2^i \cdot 8d(\epsilon) - 3 \cdot 2^i-3 \cdot 8d(\epsilon_0) - 3 \geq 2^i \cdot 4.5d(\epsilon_0)$. As the extractor $\{Q_y\}_{y \in T}$ has error at most $5\epsilon_0$, the total error is bounded by $20\epsilon_0 = \epsilon$.

We thus obtain an

$$(n, 2^i \cdot 8d(\epsilon)) \rightarrow 4 \cdot 2^{i-3} \cdot 8d(\epsilon_0)$$

strong permutation extractor with seed set $S = T^4$ so that $\log |S| \leq 4 \cdot \frac{d(\epsilon_0)}{8} \leq d(\epsilon)$. This proves the induction step. To obtain the theorem, we simply choose the smallest $i$ such that $2^i \cdot 8d(\epsilon) \geq k$. \hfill \Box

By a repeated application of the previous theorem, we can extract a larger fraction of the min-entropy.

**Theorem 3.4.5** For all (constant) $\delta \in (0, 1)$, there exists $c > 0$, such that for all positive integers $n$, all $k \in [c \log(n/\epsilon), n]$, and all $\epsilon \in (0, 1/2)$, there is an explicit
\((n,k) \rightarrow \epsilon (1-\delta)k\) strong permutation extractor \(\{P_y\}_{y \in S}\) with \(|S| = O(\log(n/\epsilon))\). Moreover, the functions \((x,y) \mapsto P_y(x)\) and \((x,y) \mapsto P^{-1}_y(x)\) can be computed by circuits of size \(O(n \text{ polylog}(n/\epsilon))\).

**Proof** We start by applying the extractor of Theorem A.2.9. We extract part of the min-entropy of the source and the remaining min-entropy is in the \(R\) system (Lemma A.2.8). This min-entropy can be extracted using once again the extractor of Theorem A.2.9. After \(O(\log(1/\delta))\) applications of the extractor, we obtain the desired result. \(\square\)

### A.3 Various technical results

This section contains various technical results. We start by a lower bound on the key size for an encryption scheme.

**Proposition A.3.1.** Let \(E : \{0,1\}^n \times [t] \rightarrow S(A)\) be an encryption scheme with the following properties: there exists a decoding map \(D_k\) for every \(k \in \{0,1\}^s\) such that \(D_k(E(x,k)) = x\) and for all \(x \neq x'\), we have

\[
\Delta \left( \frac{1}{t} \sum_k E(x,k), \frac{1}{t} \sum_k E(x',k) \right) \leq \epsilon. \tag{A.5}
\]

Then, \(\log t \geq n - 2\) provided \(\epsilon \leq 1/2\).

**Proof** The argument we use is quite similar to [Desrosiers and Dupuis, 2010, Theorem 6]. First by averaging (A.5) over all \(x'\), we obtain

\[
\Delta \left( \rho^{XA}, \rho^{X} \otimes \rho^{A} \right) \leq \epsilon + 2^{-n},
\]

where \(\rho^{XKA} = \frac{1}{t} \sum_{x,k} |x\rangle |x\rangle^{X} \otimes |k\rangle^{K} \otimes E(x,k)^{A}\). Using the relation between the trace distance and fidelity (equation (2.7)), we get

\[
F\left( \rho^{XA}, \rho^{X} \otimes \rho^{A} \right) \geq 1 - \epsilon - 2^{-n}. \tag{A.6}
\]

Now, using the key \(K\), one should be able to recover \(X\) from \(A\): this will allow us to get an upper bound on \(F\left( \rho^{XA}, \rho^{X} \otimes \rho^{A} \right)\). Using Uhlmann’s theorem (Theorem 2.2.1), we can find a purification \(|\rho\rangle^{XKAR}\) of \(\rho^{XKA}\) and a purification \(|\sigma\rangle^{XKAR}\) of \(\rho^{X} \otimes \rho^{A}\) such that

\[
F\left( \rho^{XA}, \rho^{X} \otimes \rho^{A} \right) = F\left( \rho^{XKAR}, \sigma^{XKAR} \right) \leq F\left( \rho^{XKA}, \sigma^{XKA} \right) \leq F\left( D(\rho^{XKA}), D(\sigma^{XKA}) \right) = F\left( \phi^{XX'}, D(\sigma^{XKA}) \right).
\]
Here, \( \mathcal{D} = \sum_k |k\rangle\langle k| \otimes \mathcal{D}_k \) acts on \( KA \) and \( \Phi^{XX'} = \frac{1}{2^n} \sum_x |x\rangle \langle x| \otimes |x\rangle \langle x| \). The two inequalities follow from the monotonicity of the fidelity (equation (2.8)). The last equality comes from the fact that \( D \) decodes \( X \) correctly given \( K \) and \( A \). Note that we can assume that \( \sigma^{X|KA} \) is classical on the \( XK \) system (otherwise, you can simply measure \( XK \) in the computational basis and use the monotonicity of the fidelity). We can then write \( \sigma^{X|KA} = \frac{1}{2^n} \sum_x |x\rangle \langle x| \otimes \sigma_x^{KA} \). In this case \( F(\Phi^{XX'}, D(\sigma^{X|KA})) \) is simply the probability of successfully guessing \( X \) given the system \( KA \) by applying \( D \), the underlying state being \( \sigma^{X|KA} \). In fact, expanding the fidelity, we have

\[
F(\Phi^{XX'}, D(\sigma^{X|KA})) = \left\| \frac{1}{2^n/2} \sum_x |x\rangle \langle x| \otimes |x\rangle \langle x| \cdot \frac{1}{2^n/2} \sum_x |x\rangle \langle x| \otimes \sqrt{D(\sigma_x^{KA})} \right\|_1
\]

\[
= \frac{1}{2^n} \sum_x \left\| |x\rangle \cdot \sqrt{D(\sigma_x^{KA})} \right\|_1
\]

\[
= \frac{1}{2^n} \sum_x \text{tr} \left[ \sqrt{\langle x|D(\sigma_x^{KA})|x\rangle} \right]
\]

\[
= \frac{1}{2^n} \sum_x \langle x|D(\sigma_x^{KA})|x\rangle
\]

\[
\leq P_{\text{guess}}(X|KA)_{\sigma} = 2^{-H_{\text{min}}(X|KA)_{\sigma}},
\]

where we used the operational interpretation of the min-entropy (2.12) in the last line. Now using a chain rule for the min-entropy in [Desrosiers and Dupuis, 2010, Lemma 7], we have \( H_{\text{min}}(X|KA)_{\sigma} \geq H_{\text{min}}(X|A)_{\sigma} - \log t = n - \log t \) (note that it is important here that \( K \) is classical). Combining with (A.6), we get \( 1 - \epsilon - 2^{-n} \leq 2^n/t \), which leads to the desired result.

Next, we state the general decoupling result of [Dupuis, 2010, Dupuis et al. 2010a] for exact unitary 2-designs.

**Lemma A.3.2 ([Dupuis, 2010, Theorem 3.7]).** Let \( A = A_1A_2 \), and consider the map \( \mathcal{T}_{A \rightarrow A_1} \) as defined in Equation (5.3). Then, if \( \{U_1, \ldots, U_t\} \) defines an exact unitary 2-design (Definition 5.2.4), we have for \( \delta \geq 0 \),

\[
\frac{1}{t} \sum_{i=1}^t \left\| \mathcal{T}_{A \rightarrow A_1}(U_i \rho_{AE} U_i^\dagger) - \frac{1}{|A_1|} \otimes \rho_E \right\|_1 \leq \sqrt{\frac{d_{A_1}}{d_A}} 2^{-H_{\text{min}}(A|E)_{\rho}} + 2 \delta.
\]

(A.7)

We also use the fact that a full set of MUBs defines a complex projective 2-design.
Lemma A.3.3 (Klappenecker and Rotteler [2005]). Let \( \{U_1, \ldots, U_{d_A+1}\} \) define a full set of mutually unbiased bases of \( A \). Then

\[
\frac{1}{d_A(d_A + 1)} \sum_{i=1}^{d_A+1} \sum_{a \in [d]} (U_i |a\rangle\langle a| U_i^\dagger)^{\otimes 2} = \frac{2\Pi_{\text{sym}}}{d_A(d_A + 1)},
\]

where \( \Pi_{\text{sym}} \) is the projector onto the symmetric subspace of \( A \otimes A' \) (with \( A' \simeq A \)) spanned by the vectors \(|aa'\rangle + |a'a\rangle\) for \( a, a' \in [A] \). Note that \( \Pi_{\text{sym}} = \frac{\text{id}_{AA'} + F_{AA'}}{2} \).

The following well known ‘swap trick’ is used to prove decoupling statements.

Lemma A.3.4. Let \( M, N \in \mathcal{L}(A) \). Then,

\[
\text{tr}[MN] = \text{tr}[(M_A \otimes N_{A'})(F_{AA'})],
\]

where \( A' \simeq A \) and \( F_{AA'} = \sum_{aa'} |aa'\rangle \langle a'a| \) is the swap operator.

The following is called operator Chernoff bound.

Lemma A.3.5 ([Ahlswede and Winter, 2002, Theorem 19]). Let \( X_1, \ldots, X_t \) be independent and identically distributed operator valued random variables and \( 0 \leq X_i \leq \text{id} \), \( E \{X_i\} = \Gamma \geq \alpha \text{id} \). Then

\[
\Pr \left\{ \frac{1}{t} \sum_{i=1}^{t} X_i \leq (1 + \eta)\Gamma \right\} \geq 1 - d \exp \left( -\frac{t\eta^2 \alpha}{4 \ln 2} \right).
\]
Bibliography

A. Abeyesinghe, I. Devetak, P. Hayden, and A. Winter. The mother of all protocols: Restructuring quantum information’s family tree. *Proceedings of Royal Society A*, 465:2537, 2009. arXiv:quant-ph/0606225v1.

A. Acín, S. Massar, and S. Pironio. Randomness versus nonlocality and entanglement. *Phys. Rev. Lett.*, 108:100402, Mar 2012. arXiv:1107.2754v1.

R. Ahlswede and A. Winter. Strong converse for identification via quantum channels. *IEEE Trans. Inform. Theory*, 48(3):569–579, mar 2002. arXiv:quant-ph/0012127.

R. Alicki and M. Fannes. Continuity of quantum mutual information. 2003. arXiv:quant-ph/0312081.

A. Ambainis. Limits on entropic uncertainty relations. *Quantum Inf. Comput.*, 10(9 & 10):848–858, 2010. arXiv:0909.3720.

A. Ambainis and J. Emerson. Quantum t-designs: t-wise independence in the quantum world. In *Proc. IEEE CCC*, pages 129–140. IEEE, 2007.

A. Ambainis and A. Smith. Small Pseudo-random Families of Matrices: Derandomizing Approximate Quantum Encryption. In *APPROX-RANDOM*, volume 3122 of *LNCS*, pages 249–260. 2004. arXiv:quant-ph/0404075.

A. Ambainis, M. Mosca, A. Tapp, and R. de Wolf. Private quantum channels. In *Proc. ACM STOC*, pages 547–553, 2000.

G. Aubrun, S. Szarek, and E. Werner. Nonadditivity of Rényi entropy and Dvoretzky’s theorem. *J. Math. Phys.*, 51(2):022102, 2010. arXiv:0910.1189.

G. Aubrun, S. Szarek, and E. Werner. Hastings’s Additivity Counterexample via Dvoretzky’s Theorem. *Comm. Math. Phys.*, 305:85–97, 2011. arXiv:1003.4925.
A. Azarchs. Entropic uncertainty relations for incomplete sets of mutually unbiased observables. 2004. quant-ph/0412083.

K. Ball. An elementary introduction to modern convex geometry. Flavors of geometry, 31: 1–58, 1997.

M. A. Ballester and S. Wehner. Entropic uncertainty relations and locking: Tight bounds for mutually unbiased bases. Phys. Rev. A, 75(2):022319, Feb 2007. arXiv:quant-ph/0606244.

S. Bandyopadhyay, P. O. Boykin, V. Roychowdhury, and F. Vatan. A new proof for the existence of mutually unbiased bases. Algorithmica, 34(4):512–528, 2002. arXiv:quant-ph/0103162.

H. Barnum, E. Knill, and M. A. Nielsen. On quantum fidelities and channel capacities. IEEE Trans. Inform. Theory, 46:1317–1329, 2000. arXiv:quant-ph/9809010v1.

A. Ben-Aroya, O. Schwartz, and A. Ta-Shma. Quantum expanders: Motivation and construction. Theory of Computing, 6:47–79, 2010.

C. Bennett, G. Brassard, and J. Robert. Privacy amplification by public discussion. SIAM J. Comput., 17:210, 1988.

C. Bennett, G. Brassard, C. Crepeau, and U. Maurer. Generalized privacy amplification. IEEE Trans. Inform. Theory, 41(6):1915 –1923, Nov 1995.

C. H. Bennett and G. Brassard. Quantum cryptography: Public key distribution and coin tossing. In Proc. International Conference on Computers, Systems and Signal Processing, 1984.

C. H. Bennett, D. DiVincenzo, J. A. Smolin, and W. K. Wootters. Mixed-state entanglement and quantum error correction. Phys. Rev. A, 54(5):3824–3851, Nov 1996. arXiv:quant-ph/9604024.

M. Berta, M. Christandl, R. Colbeck, J. M. Renes, and R. Renner. The uncertainty principle in the presence of quantum memory. Nat. Phys., 6:659, 2010. arXiv:0909.0950v4.

M. Berta, F. Brandao, M. Christandl, and S. Wehner. Entanglement cost of quantum channels. arXiv:1108.5357, 2011a.
M. Berta, M. Christandl, and R. Renner. The quantum reverse Shannon theorem based on one-shot information theory. *Comm. Math. Phys.*, 306:579–615, 2011b. arXiv:0912.3805v2.

M. Berta, F. Furrer, and V. B. Scholz. The smooth entropy formalism on von Neumann algebras. 2011c. arXiv:1107.5460v1.

M. Berta, O. Fawzi, and S. Wehner. Quantum to classical randomness extractors. In *Proc. CRYPTO*, volume 7417 of *LNCS*, pages 776–793, 2012. arXiv:1111.2026.

R. Bhatia. *Matrix Analysis*. Springer, 1997.

I. Bialynicki-Birula and J. Mycielski. Uncertainty relations for information entropy in wave mechanics. *Comm. Math. Phys.*, 44(2):129–132, 1975.

H. Buhrman, R. Cleve, J. Watrous, and R. de Wolf. Quantum fingerprinting. *Phys. Rev. Lett.*, 87(16):167902, 2001. arXiv:quant-ph/0102001.

H. Buhrman, M. Christandl, P. Hayden, H. K. Lo, and S. Wehner. Security of quantum bit string commitment depends on the information measure. *Phys. Rev. Lett.*, 97(25):250501, 2006. arXiv:quant-ph/0609237.

H. Buhrman, M. Christandl, P. Hayden, H. K. Lo, and S. Wehner. Possibility, impossibility, and cheat sensitivity of quantum-bit string commitment. *Phys. Rev. A*, 78(2):22316, 2008. arXiv:quant-ph/0504078.

C. Cachin and U. M. Maurer. Unconditional security against memory-bounded adversaries. In *Proc. CRYPTO*, volume 1294 of *LNCS*, pages 292–306, 1997.

M. Christandl and A. Winter. Uncertainty, monogamy, and locking of quantum correlations. *IEEE Trans. Inform. Theory*, 51:3159–3165, 2005. arXiv:quant-ph/0501090v2.

R. Colbeck. *Quantum and relativistic protocols for secure multi-party computation*. PhD thesis, University of Cambridge, 2006. arXiv:0911.3814v2.

R. Colbeck and A. Kent. Private randomness expansion with untrusted devices. *J. Phys. A - Math. Gen.*, 44:095305, 2011. arXiv:1011.4474v3.

P. J. Coles, L. Yu, V. Gheorghiu, and R. B. Griffiths. Information theoretic treatment of tripartite systems and quantum channels. *Phys. Rev. A*, 83:062338, 2011a. arXiv:1006.4859v5.
P. J. Coles, L. Yu, and M. Zwolak. Relative entropy derivation of the uncertainty principle with quantum side information. 2011b. arXiv:1105.4865v2.

P. J. Coles, R. Colbeck, L. Yu, and M. Zwolak. Uncertainty relations from simple entropic properties. 2012. arXiv:1112.0543v1.

I. Damgård, S. Fehr, L. Salvail, and C. Schaffner. Cryptography in the bounded quantum-storage model. In Proc. IEEE FOCS, pages 449–458, 2005. arXiv:quant-ph/0508222.

I. Damgård, S. Fehr, R. Renner, L. Salvail, and C. Schaffner. A tight high-order entropic quantum uncertainty relation with applications. In Proc. CRYPTO, volume 4622 of LNCS, pages 360–378. 2007. arXiv:quant-ph/0612014.

I. Damgård, T. B. Pedersen, and L. Salvail. On the key-uncertainty of quantum ciphers and the computational security of one-way quantum transmission. In Proc. EUROCRYPT, volume 3027 of LNCS, pages 91–108. 2004. arXiv:quant-ph/0407066.

I. Damgård, T. B. Pedersen, and L. Salvail. A quantum cipher with near optimal key-recycling. In Proc. CRYPTO, volume 3621 of LNCS, pages 494–510. 2005.

C. Dankert, R. Cleve, J. Emerson, and E. Livine. Exact and approximate unitary 2-designs and their application to fidelity estimation. Phys. Rev. A, 80(1):12304, 2009. arXiv:quant-ph/0606161.

G. M. D’Ariano, D. Kretschmann, D. Schlingemann, and R. F. Werner. Reexamination of quantum bit commitment: The possible and the impossible. Phys. Rev. A, 76(3):32328, 2007.

A. De, C. Portmann, T. Vidick, and R. Renner. Trevisan’s extractor in the presence of quantum side information. 2009. arXiv:0912.5514.

S. P. Desrosiers. Entropic security in quantum cryptography. Quantum Inf. Process., 8:331–345, 2009.

S. P. Desrosiers and F. Dupuis. Quantum entropic security and approximate quantum encryption. IEEE Trans. Inform. Theory, 56(7):3455 –3464, Jul 2010. arXiv:0707.0691.

D. Deutsch. Uncertainty in quantum measurements. Phys. Rev. Lett., 50(9):631–633, Feb 1983.
D. P. DiVincenzo, M. Horodecki, D. W. Leung, J. A. Smolin, and B. M. Terhal. Locking classical correlations in quantum states. *Phys. Rev. Lett.*, 92(6):67902, 2004. arXiv:quant-ph/0303088.

Y. Dodis and A. Smith. Entropic security and the encryption of high entropy messages. *Proc. TCC*, pages 556–577, 2005.

W. Doeblin. Exposé de la théorie des chaînes simples constantes de markov á un nombre fini d´états. *Mathématique de lUnion Interbalkanique*, 2(77-105):78–80, 1938.

F. Dupuis. *A decoupling approach to quantum information theory*. PhD thesis, Université de Montreal, 2010. arXiv:1004.1641.

F. Dupuis, M. Berta, J. Wullschleger, and R. Renner. The decoupling theorem. 2010a. arXiv:1012.6044v1.

F. Dupuis, J. Florjanczyk, P. Hayden, and D. Leung. Locking classical information. 2010b. arXiv:1011.1612.

K. Dvijotham and M. Fazel. A nullspace analysis of the nuclear norm heuristic for rank minimization. In *Proc. IEEE ICASSP*, pages 3586–3589. IEEE, 2010.

S. Dziembowski and U. Maurer. On generating the initial key in the bounded-storage model. In *Proc. EUROCRYPT*, volume 3027 of *LNCS*, pages 126–137, 2004.

O. Fawzi, P. Hayden, and P. Sen. From low-distortion norm embeddings to explicit uncertainty relations and efficient information locking. In *Proc. ACM STOC*, pages 773–782, 2011. arXiv:1010.3007v3.

S. Fehr, R. Gelles, and C. Schaffner. Security and Composability of Randomness Expansion from Bell Inequalities. 2011. arXiv:1111.6052v2.

R. Feynman. Simulating physics with computers. *International Journal of Theoretical Physics*, 21:467–488, 1982. 10.1007/BF02650179.

T. Figiel, J. Lindenstrauss, and V. D. Milman. The dimension of almost spherical sections of convex bodies. *Acta Math.*, 139(1):53–94, 1977.

C. A. Fuchs and J. van de Graaf. Cryptographic distinguishability measures for quantum-mechanical states. *IEEE Trans. Inform. Theory*, 45(4):1216–1227, 1999. arXiv:quant-ph/9712042.
F. Furrer, T. Franz, M. Berta, V. B. Scholz, M. Tomamichel, and R. F. Werner. Continuous variable quantum key distribution: Finite-key analysis of composable security against coherent attacks. 2011. arXiv:1112.2179.

D. Gavinsky and T. Ito. Quantum Fingerprints that Keep Secrets. 2010. arXiv:1010.5342.

D. Gavinsky, J. Kempe, I. Kerenidis, R. Raz, and R. de Wolf. Exponential separations for one-way quantum communication complexity, with applications to cryptography. In Proc. ACM STOC, pages 516–525. ACM, 2007. arXiv:quant-ph/0611209.

N. Gisin, G. Ribordy, W. Tittel, and H. Zbinden. Quantum cryptography. Rev. Mod. Phys., 74(1):145–195, 2002. arXiv:quant-ph/0101098.

O. Goldreich. Computational complexity: a conceptual perspective. Cambridge University Press, 2008.

O. Goldreich and A. Wigderson. Tiny families of functions with random properties: A quality-size trade-off for hashing. Random Structures Algorithms, 11(4):315–343, 1997.

D. Gross, K. Audenaert, and J. Eisert. Evenly distributed unitaries: On the structure of unitary designs. J. Math. Phys., 48:052104, 2007. arXiv:quant-ph/0611002v2.

V. Guruswami, C. Umans, and S. Vadhan. Unbalanced expanders and randomness extractors from Parvaresh–Vardy codes. J. ACM, 56(4), 2009.

S. Hallgren, C. Moore, M. Rötteler, A. Russell, and P. Sen. Limitations of quantum coset states for graph isomorphism. J. ACM, 57(6), 2010.

A. Harrow and R. Low. Random quantum circuits are approximate 2-designs. Comm. Math. Phys., 291:257–302, 2009. arXiv:0802.1919v3.

A. Harrow, P. Hayden, and D. Leung. Superdense coding of quantum states. Phys. Rev. Lett., 92(18):187901, 2004. arXiv:quant-ph/0307221.

M. B. Hastings. Superadditivity of communication capacity using entangled inputs. Nat. Phys., 5(4):255–257, 2009.

P. Hayden and A. Winter. Counterexamples to the maximal $p$-norm multiplicativity conjecture for all $p > 1$. Comm. Math. Phys., 284(1):263–280, 2008. arXiv:0807.4753.

P. Hayden and A. Winter. Weak decoupling duality and quantum identification. IEEE Trans. Inform. Theory, PP(99):1, 2012. arXiv:1003.4994.
P. Hayden, D. Leung, P. W. Shor, and A. Winter. Randomizing quantum states: Constructions and applications. *Comm. Math. Phys.*, 250(2):371–391, 2004. arXiv:quant-ph/0307104.

P. Hayden, D. W. Leung, and A. Winter. Aspects of generic entanglement. *Comm. Math. Phys.*, 265(1):95–117, 2006. arXiv:quant-ph/0407049.

P. Hayden, M. Horodecki, J. Yard, and A. Winter. A decoupling approach to the quantum capacity. *Open Systems and Information Dynamics*, 15:7–19, 2008. arXiv:quant-ph/0702005v1.

R. W. Heath, T. Strohmer, and A. J. Paulraj. On quasi-orthogonal signatures for CDMA systems. *IEEE Trans. Inform. Theory*, 52(3):1217–1226, 2006.

W. Heisenberg. ¨Uber den anschaulichen Inhalt der quantentheoretischen Kinematik und Mechanik. *Zeitschrift f¨ur Physik A Hadrons and Nuclei*, 43(3):172–198, 1927.

C. W. Helstrom. Detection theory and quantum mechanics. *Information and Control*, 10: 254–291, 1967.

L. Henderson and V. Vedral. Classical, quantum and total correlations. *J. Phys. A - Math. Gen.*, 34(35):6899, 2001. arXiv:quant-ph/0105028.

I. I. Hirschman. A note on entropy. *American journal of mathematics*, 79(1):152–156, 1957.

K. Horodecki, M. Horodecki, P. Horodecki, and J. Oppenheim. Locking entanglement with a single qubit. *Phys. Rev. Lett.*, 94(20):200501, May 2005a. arXiv:quant-ph/0404096.

M. Horodecki, J. Oppenheim, and A. Winter. Partial quantum information. *Nature*, 436: 673–676, 2005b. arXiv:quant-ph/0505062v1.

M. Horodecki, J. Oppenheim, and A. Winter. Quantum state merging and negative information. *Comm. Math. Phys.*, 269:107, 2006. arXiv:quant-ph/0512247v1.

R. Impagliazzo, L. Levin, and M. Luby. Pseudo-random generation from one-way functions. In *Proc. ACM STOC*, pages 12–24. ACM, 1989.

P. Indyk. Uncertainty principles, extractors, and explicit embeddings of L2 into L1. In *Proc. ACM STOC*, pages 615–620. ACM, 2007.
P. Indyk and S. Szarek. Almost-euclidean subspaces of $\ell_1^n$ via tensor products: A simple approach to randomness reduction. In *APPROX-RANDOM*, volume 6302 of *LNCS*, pages 632–641. 2010. arXiv:1001.0041.

I. D. Ivanovic. An inequality for the sum of entropies of unbiased quantum measurements. *J. Phys. A - Math. Gen.*, 25(7):L363, 1992.

I. Ivanovic. Geometrical description of quantal state determination. *J. Phys. A - Math. Gen.*, 14:3241, 1981.

B. Kashin. Sections of some finite dimensional sets and classes of smooth functions. *Izv. Acad. Nauk SSSR*, 41:334–351, 1977.

J. Kempe and T. Vidick. Parallel repetition of entangled games. 2010. arXiv:1012.4728.

J. Kilian. Founding cryptography on oblivious transfer. In *Proc. ACM STOC*, pages 20–31, 1988.

A. Klappenecker and M. Rotteler. Mutually unbiased bases are complex projective 2-designs. In *Proc. IEEE ISIT*, pages 1740–1744, 2005. arXiv:quant-ph/0502031.

M. Koashi and A. Winter. Monogamy of quantum entanglement and other correlations. *Phys. Rev. A*, 69(2):022309, Feb 2004. arXiv:quant-ph/0310037.

R. König and R. Renner. Sampling of min-entropy relative to quantum knowledge. *IEEE Trans. Inform. Theory*, 57(7):4760–4787, july 2011. arXiv:0712.4291.

R. König and B. M. Terhal. The bounded-storage model in the presence of a quantum adversary. *IEEE Trans. Inform. Theory*, 54:749–762, 2008.

R. König, U. Maurer, and R. Renner. On the power of quantum memory. *IEEE Trans. Inform. Theory*, 51:2391–2401, 2005. arXiv:quant-ph/0305154v3.

R. König, R. Renner, A. Bariska, and U. Maurer. Small accessible quantum information does not imply security. *Phys. Rev. Lett.*, 98(14):140502, Apr 2007. arXiv:quant-ph/0512021.

R. König, R. Renner, and C. Schaffner. The operational meaning of min- and max-entropy. *IEEE Trans. Inform. Theory*, 55:4674–4681, 2009. arXiv:0807.1338v1.

R. König, S. Wehner, and J. Wullschleger. Unconditional security from noisy quantum storage. *IEEE Trans. Inform. Theory*, 58(3):1962–1984, march 2012.
E. Kushilevitz and N. Nisan. *Communication Complexity*. Cambridge University Press, 1997.

U. Larsen. Superspace geometry: the exact uncertainty relationship between complementary aspects. *J. Phys. A - Math. Gen.*, 23:1041, 1990.

M. Ledoux. *The concentration of measure phenomenon*. American Mathematical Society, 2001.

D. Leung. A survey on locking of bipartite correlations. In *Journal of Physics: Conference Series*, volume 143, page 012008. Institute of Physics Publishing, 2009.

H.-K. Lo. Insecurity of quantum secure computations. *Phys. Rev. A*, 56:1154, 1997.

H. K. Lo and H. F. Chau. Is quantum bit commitment really possible? *Phys. Rev. Lett.*, 78 (17):3410–3413, 1997. arXiv:quant-ph/9603004.

A. Lubotzky, R. Phillips, and P. Sarnak. Ramanujan graphs. *Combinatorica*, 8(3):261–277, 1988.

H. Maassen and J. B. M. Uffink. Generalized entropic uncertainty relations. *Phys. Rev. Lett.*, 60(12):1103–1106, Mar 1988.

P. Mandayam and S. Wehner. Achieving the physical limits of the bounded-storage model. *Phys. Rev. A*, 83:022329, 2011. arXiv:1009.1596v2.

U. Maurer. Conditionally-perfect secrecy and a provably-secure randomized cipher. *J. Cryptol.*, 5:53–66, 1992.

D. Mayers. Unconditionally secure quantum bit commitment is impossible. *Phys. Rev. Lett.*, 78(17):3414–3417, 1997. arXiv:quant-ph/9605044.

V. D. Milman. New proof of the theorem of A. Dvoretzky on intersections of convex bodies. *Funct. Anal. Appl.*, 5:288–295, 1971.

V. D. Milman and G. Schechtman. *Asymptotic theory of finite dimensional normed spaces*, volume 1200 of *Lecture Notes in Mathematics*. Springer-Verlag, 1986.

M. Nielsen and I. Chuang. *Quantum computation and quantum information*. Cambridge Series on Information and the Natural Sciences. Cambridge University Press, 2000. ISBN 9780521635035.
N. Nisan and D. Zuckerman. Randomness is linear in space. *J. Comput. Syst. Sci.*, 52(1):43 – 52, 1996.

H. Ollivier and W. Zurek. Quantum discord: A measure of the quantumness of correlations. *Phys. Rev. Lett.*, 88:017901, Dec 2001.

J. Oppenheim and M. Horodecki. How to reuse a one-time pad and other notes on authentication, encryption, and protection of quantum information. *Phys. Rev. A*, 72(4):042309, Oct 2005. arXiv:quant-ph/0306161.

S. Pironio and S. Massar. Security of practical private randomness generation. 2011. arXiv:1111.6056.

S. Pironio, A. Acín, S. Massar, A. de La Giroday, D. Matsukevich, P. Maunz, S. Olmschenk, D. Hayes, L. Luo, T. Manning, et al. Random numbers certified by bell’s theorem. *Nature*, 464(7291):1021–1024, 2010.

G. Pisier. *The volume of convex bodies and Banach space geometry*. Cambridge University Press, 1989.

J. Radhakrishnan and A. Ta-Shma. Bounds for dispersers, extractors, and depth-two superconcentrators. *SIAM J. Discrete Math.*, 13:2, 2000.

J. Radhakrishnan, M. Rötteler, and P. Sen. Random Measurement Bases, Quantum State Distinction and Applications to the Hidden Subgroup Problem. *Algorithmica*, 55(3):490–516, 2009.

R. Raz, O. Reingold, and S. Vadhan. Extracting all the randomness and reducing the error in Trevisan’s extractors. In *Proc. ACM STOC*, pages 149–158. ACM, 1999.

O. Reingold, S. Vadhan, and A. Wigderson. Entropy waves, the zig-zag graph product, and new constant-degree expanders and extractors. In *Proc. IEEE FOCS*, pages 3 –13, 2000.

J. M. Renes and J.-C. Boileau. Conjectured strong complementary information tradeoff. *Phys. Rev. Lett.*, 103:020402, Jul 2009. arXiv:0806.3984.

R. Renner. Security of quantum key distribution. *Int. J. Quantum Inf.*, 6:1, 2008. arXiv:quant-ph/0512258v2.

R. Renner and R. König. Universally composable privacy amplification against quantum adversaries. *Proc. TCC*, pages 407–425, 2005. arXiv:quant-ph/0403133v2.
H. P. Robertson. The uncertainty principle. *Physical Review*, 34(1):163–164, 1929.

A. Russell and H. Wang. How to fool an unbounded adversary with a short key. In *Proc. EUROCRYPT*, volume 2332 of *LNCS*, pages 133–148. 2002.

J. Sanchez. Entropic uncertainty and certainty relations for complementary observables. *Phys. Lett. A*, 173(3):233 – 239, 1993.

J. Sanchez-Ruiz. Improved bounds in the entropic uncertainty and certainty relations for complementary observables. *Phys. Lett. A*, 201:125, 1995.

V. Scarani, H. Bechmann-Pasquinucci, N. Cerf, M. Dušek, N. Lütkenhaus, and M. Peev. The security of practical quantum key distribution. *Rev. Mod. Phys.*, 81:1301–1350, Sep 2009. doi: 10.1103/RevModPhys.81.1301. URL http://link.aps.org/doi/10.1103/RevModPhys.81.1301

C. Schaffner. *Cryptography in the Bounded-Quantum-Storage Model*. PhD thesis, BRICS, University of Aarhus, March 2007. arXiv:0709.0289.

C. Schaffner, B. Terhal, and S. Wehner. Robust cryptography in the noisy-quantum-storage model. *Quantum Inf. Comput.*, 9:11, 2008. arXiv:0807.1333v3.

P. Schönenmann. A generalized solution of the orthogonal procrustes problem. *Psychometrika*, 31(1):1–10, 1966.

R. Shaltiel. Recent developments in explicit constructions of extractors. *Bull. EATCS*, 77:67–95, 2002.

C. Shannon. Communication Theory of Secrecy Systems. *Bell System Technical Journal*, 28:656–715, 1949.

P. Shor. Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer. *SIAM J. Comput.*, 26(5):1484–1509, Oct. 1997. arXiv:quant-ph/9508027.

V. Shoup. New algorithms for finding irreducible polynomials over finite fields. *Math. Comp.*, 54(189):435–447, 1990.

V. Shoup. Searching for primitive roots in finite fields. *Math. Comp.*, 58(197):pp. 369–380, 1992.
R. W. Spekkens and T. Rudolph. Degrees of concealment and bindingness in quantum bit commitment protocols. *Phys. Rev. A*, 65(1):12310, 2001. arXiv:quant-ph/0106019.

S. Szarek. Convexity, complexity, and high dimensions. In *International Congress of Mathematicians*, volume 2, pages 1599–1621, 2006.

O. Szehr, F. Dupuis, M. Tomamichel, and R. Renner. Decoupling with unitary almost two-designs. arXiv:1109.4348, 2011.

A. Ta-Shma. Short seed extractors against quantum storage. In *Proc. ACM STOC*, pages 401–408. ACM, 2009.

M. Tomamichel. *A Framework for Non-Asymptotic Quantum Information Theory*. PhD thesis, 2012. arXiv:1203.2142.

M. Tomamichel and R. Renner. Uncertainty relation for smooth entropies. *Phys. Rev. Lett.*, 106:110506, Mar 2011. arXiv:1009.2015.

M. Tomamichel, R. Colbeck, and R. Renner. A fully quantum asymptotic equipartition property. *IEEE Trans. Inform. Theory*, 55:5840–5847, 2009. arXiv:0811.1221v3.

M. Tomamichel, R. Colbeck, and R. Renner. Duality between smooth min- and max-entropies. *IEEE Trans. Inform. Theory*, 56:4674, 2010. arXiv:0907.5238v2.

M. Tomamichel, C. Schaffner, A. Smith, and R. Renner. Leftover hashing against quantum side information. *IEEE Trans. Inform. Theory*, 57(8):5524 –5535, aug. 2011. arXiv:1002.2436v1.

M. Tomamichel, C. Lim, N. Gisin, and R. Renner. Tight finite-key analysis for quantum cryptography. *Nat. Comm.*, 3:634, 2012. arXiv:1103.4130.

J. Tropp. *Topics in Sparse Approximation*. PhD thesis, University of Texas at Austin, 2004.

A. Uhlmann. The “transition probability” in the state space of a *-algebra. *Reports on Mathematical Physics*, 9(2):273–279, 1976.

S. Vadhan. The unified theory of pseudorandomness: guest column. *ACM SIGACT News*, 38(3):39–54, 2007.

U. Vazirani and T. Vidick. Certifiable Quantum Dice -Or, testable exponential randomness expansion. 2011. arXiv:1111.6054v1.
J. von zur Gathen and J. Gerhard. *Modern computer algebra*. Cambridge University Press, 1999.

S. Wehner and A. Winter. Entropic uncertainty relations—a survey. *New J. Phys.*, 12:025009, 2010. arXiv:0907.3704.

S. Wehner, C. Schaffner, and B. Terhal. Cryptography from noisy storage. *Phys. Rev. Lett.*, 100:220502, 2008. arXiv:0711.2895v3.

S. Wiesner. Conjugate coding. *ACM SIGACT News*, 15(1):78–88, 1983.

M. Wilde. *From Classical to Quantum Shannon Theory*. 2011. arXiv:1106.1445.

A. Winter. Quantum and classical message identification via quantum channels. *Quantum Inf. Comput.*, 4(6&7):563–578, 2004. arXiv:quant-ph/0401060.

W. K. Wootters and B. D. Fields. Optimal state-determination by mutually unbiased measurements. *Ann. Physics*, 191(2):363 – 381, 1989.

D. Zuckerman. On unapproximable versions of np-complete problems. *SIAM J. Comput.*, 25(6):1293–1304, 1996a.

D. Zuckerman. Simulating BPP using a general weak random source. *Algorithmica*, 16(4):367–391, 1996b.