Emerging Paradigms in Quantum Error Correction and Quantum Cryptography

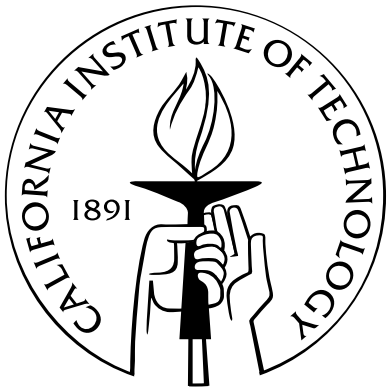
Thesis by

Prabha Mandayam Doddamane

In Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy



California Institute of Technology

Pasadena, California

2011

(Defended April 8, 2011)

To

Amma and *Appa*

Acknowledgements

As I get ready to submit my doctoral thesis, it is a pleasure to finally acknowledge and thank everyone who has been a part of this incredible journey.

First and foremost, thanks to my adviser Prof. John Preskill for his help and advice these past six years. Being his student and interacting with him has been a rare privilege and a great learning experience for me. Apart from having learnt some of the fundamental concepts of quantum information processing from him, I have also learnt a lot from his approach to scientific research, in particular, the importance of motivating and communicating the physical intuition behind one's research. But most of all, thank you John, for your patience and understanding during my initial blundering forays into research. And thanks also, for letting me be a part of the great institution that IQI is. With its array of post-docs working in diverse areas, and a continuous stream of visitors, I couldn't have asked for a better place in which to start my research career in quantum information.

A huge thanks to my collaborators David Poulin, Hui Khoon Ng and Stephanie Wehner. David was instrumental in getting me started on my first research problem, in the area of approximate quantum error correction (AQEC). I am greatly indebted to him for guiding me during that initial phase and for investing all that time during his busy post-doc days. In fact, the intuition behind the AQEC result presented in this thesis, are largely the outcome of discussions with David. Hui Khoon, who was a graduate student IQI, has been a great friend and colleague. What started out as a coffee-time discussion between students with similar interests, eventually led to my first paper on AQEC. Thanks Hui, for all the times we shared, chatting about work and life and everything else in between! Stephanie joined IQI during the final years of my graduate studies, leading to a very enjoyable and productive collaboration. I have of course benefitted immensely from her rich knowledge and experience. I have also learnt a lot from her single-minded approach to research, and her focus and dedication in solving a problem. Thanks Stephanie, for being the great friend

and mentor that you are. And thanks also, for introducing me to the wonderful world of quantum cryptography! A special word of thanks to Niranjan Balachandran whose expertise in combinatorics and graph theory helped us formalize some of the proofs in Chapter 3.

Thanks to Profs. Alexei Kitaev, Oskar Painter and Gil Refael for serving on my thesis committee. Thanks also to Kovid, Panos, Graeme, Greg, Ersen, Issac, Jeongwan, Robert, Robin, Liang and other colleagues at IQI.

I would also like to acknowledge here, some of the teachers who played a motivational role during my student days in India, especially Prof. Arul Lakshminarayanan and Prof. M.V.Satyanarayana. Arul, with whom I worked on my Masters' thesis, has been a great friend and guide. MVS, whose fierce passion for physics has inspired many a student, has been an honest critique and a constant well-wisher.

This thesis is dedicated to my parents, for it was with them that this journey really began. Back in high school, it was through my mother Vijayalakshmi that I discovered the joy of doing mathematics. And it was from my father Srinivas, who has been an academic all his life, that I developed an interest in research and teaching. Indeed, my fascination for quantum mechanics dates back to many simulating discussions with my father during my undergraduate days—he is among the best teachers of the subject I have seen to date! I am grateful to my parents, not only for shaping my approach to science, but for the larger lessons that I have had a chance to learn from them, through the values they exude in their everyday lives.

To my sister Nitya, my oldest friend and confidante, thanks for sharing in all the ups and downs of these PhD years. To Uday, Shweta, Sushree, JK, Mansi, Setu, Mayank, Shankar, Varun, Shriharsh, Naresh, Chithra, Arundhati, Sameer, Pinkesh, Vikram, Tejaswi and all my other friends at Caltech, thank you for making this place a home away from home for me. Thank you for all the music, food, concerts, plays and above all, the companionship, which truly enriched my graduate school experience. Thanks also to my friends from India—Jeevisha, Ashok, Devi and Roshni—who have stood by me through thick and thin. I am thankful to my husband's parents for their keen interest in my progress and their support and encouragement.

To my husband Krishna, who has been an inexhaustible source of strength these last six years, I present this thesis as a culmination of our combined efforts and aspirations. I am grateful for everything we share, and look forward to the many milestones we will cross together in the future.

Abstract

We study two novel paradigms in quantum error correction and quantum cryptography—*approximate* quantum error correction and *noisy-storage cryptography*—which explore alternate approaches for dealing with quantum noise. Approximate quantum error correction seeks to relax the constraint of perfect error correction and construct codes that might be better adapted to correct for specific noise models. Noisy-storage cryptography relies on the power of quantum noise to execute two-party cryptographic tasks securely.

Motivated by examples of approximately correcting codes, which make use of fewer physical resources than perfect codes and still obtain comparable levels of fidelity, we study the problem of finding and characterizing such codes in general. We construct for the first time a universal, near-optimal recovery map for approximate quantum error correction (AQEC), with optimality defined in terms of worst-case fidelity. Using the analytical form of this recovery, we also obtain easily verifiable conditions for AQEC. This in turn leads to a simple algorithm for identifying good approximate codes, without having to perform a difficult optimization over all recovery maps for every possible encoding.

Noisy-storage cryptography envisions a setting where two-party cryptographic protocols can be securely implemented based solely on the assumption that the quantum storage device possessed by either party is noisy and bounded. Here, we construct two-party protocols (using higher-dimensional states) that are secure even when a dishonest player can store all but a small fraction of the information transmitted during the protocol, in his noiseless quantum memory. We also show that when his memory is noisy, security can be extended to a larger class of noisy quantum memories. Our result demonstrates that the physical limits of the quantum noisy-storage model are indeed achievable, albeit asymptotically.

We also describe our investigations on obtaining strong entropic uncertainty relations using symmetric complementary bases. Uncertainty relations are an important and useful resource in analyzing the security of quantum cryptographic protocols, in addition to being of interest from a foundational standpoint. We present a novel construction of sets of symmetric, complementary bases in dimension $d = 2^n$, which are cyclically permuted under the action of a unitary transformation. We also obtain new lower bounds for uncertainty relations in terms of the min-entropy, which are tight for specific instances of our construction.

Contents

Acknowledgements **iv**

Abstract **vi**

1 Introduction **1**

1.1 Adapting to Quantum Noise: Approximate Quantum Error Correction 3

1.2 Using the Power of Quantum Noise: Noisy-Storage Cryptography 5

1.3 A Mathematical Interlude: Uncertainty Relations for Complementary Aspects 7

1.4 Thesis Organization 8

2 Approximate Quantum Error Correction Using the Transpose Channel **9**

2.1 Quantum Error Correction 10

2.1.1 Quantum Channels 12

2.1.2 Perfect QEC Conditions 14

2.2 Approximate Quantum Error Correction 17

2.2.1 The Approximate [4,1] Code 18

2.2.2 AQEC as an Optimization Problem 19

2.3 The Transpose Channel 22

2.3.1 The Transpose Channel for Perfect QEC 24

2.3.2 Near-Optimality of the Transpose Channel 25

2.4 The Transpose channel and QEC Conditions 28

2.4.1 Alternative Form of the Perfect QEC Conditions 28

2.4.2 AQEC Conditions 29

2.5 Finding AQEC Codes 32

2.6 Example: Amplitude Damping Channel . . . 33
2.7 Conclusions and Open Problems . . . 36

3 Symmetric Complementary Aspects and Entropic Uncertainty Relations 38
3.1 Preliminaries . . . 39
3.1.1 Measures of Entropy . . . 39
3.1.2 Clifford Algebras . . . 41
3.2 EURs and MUBs: An Overview . . . 42
3.2.1 Entropic Uncertainty Relations . . . 43
3.2.2 Mutually Unbiased Bases . . . 45
3.2.3 MUBs and Strong Uncertainty Relations . . . 48
3.3 Min-entropic Uncertainty Relations . . . 50
3.3.1 Symmetries . . . 51
3.3.2 Discrete Wigner Function . . . 53
3.4 New Lower Bounds on the Average Min-entropy . . . 54
3.4.1 Proof of Theorem 3.4.1 . . . 56
3.4.2 Proof of Theorem 3.4.2 . . . 57
3.5 Construction of Symmetric MUBs . . . 60
3.5.1 Examples . . . 62
3.6 Tight Lower Bounds for Symmetric MUBs . . . 63
3.7 Conclusions and Open Questions . . . 64

4 Achieving the Physical Limits of the Bounded-Storage Model 67
4.1 Preliminaries . . . 69
4.1.1 Quantifying Adversarial Information . . . 69
4.1.2 Noisy-Storage Model . . . 70
4.1.3 Characterizing the Noisy Quantum Storage . . . 72
4.1.4 Security, Storage Rate, and Channel Capacity . . . 73
4.1.5 Techniques . . . 74
4.2 Weak String Erasure . . . 75
4.2.1 Nonuniform Weak String Erasure . . . 75

4.2.2 Security against Dishonest Bob . . . 79
4.2.3 Security against Dishonest Alice . . . 82
4.3 Cryptographic Tools . . . 82
4.3.1 Privacy Amplification . . . 82
4.3.2 Min-entropy Sampling . . . 84
4.3.3 Interactive Hashing . . . 85
4.4 Oblivious Transfer . . . 86
4.4.1 Security for Bob . . . 90
4.4.2 Security for Alice . . . 92
4.4.3 Correctness . . . 92
4.5 Conclusion . . . 94

A The AQEC Algorithm for Qubit Codes **96**
A.1 Computing the Maximum Eigenvalue of Δ_{sum} . . . 96
A.2 Computing the Fidelity Loss for the Transpose Channel . . . 98

B Constructing Maximally Commuting Classes of Clifford Generators **102**
B.1 Mathematical Tools . . . 103
B.1.1 Length-2 Operators . . . 104
B.1.2 Higher-Length Operators . . . 104
B.2 Constructing $\mathbf{2n+1}$ Prime Classes . . . 106
B.3 Constructing $\mathbf{L|n}$ Classes for Prime Values of L . . . 109

Bibliography **115**

Chapter 1

Introduction

It has been over two decades since the idea of quantum information processing has captured the imagination of physicists, computer scientists, and information theorists alike. Two important discoveries that provided critical impetus for the growth of the field in its nascent stages have been quantum error correction and quantum cryptography. The discovery of quantum error correcting codes in the mid-nineties [22, 120, 123] demonstrated that it is indeed possible to perform quantum information processing reliably in the presence of noise. The early quantum cryptographic protocols discovered in the pervious decade [11, 13, 42] demonstrated the usefulness of quantum systems in performing fundamentally unbreakable cryptographic tasks.

Naïvely, the problem of quantum error correction appears daunting in the face of several conceptual challenges. The *no-cloning theorem* [133] rules out the possibility of constructing quantum repetition codes analogous to classical repetition codes. Also, since quantum errors are continuous, it is difficult to measure and identify the errors with precision. Furthermore, since the measurement process actually destroys or modifies quantum information, we cannot directly employ the standard classical technique of observing the output of the channel and selecting the decoding procedure on that basis. However, discovery of necessary and sufficient conditions for quantum error correction in general noise models [15, 43, 70] demonstrated how quantum codes can be constructed in spite of these challenges. On the other hand, for cryptographic protocols, these very properties prove to be useful in achieving security against an eavesdropper.

Taking off from these early results, there have been several important developments in both these areas. Significant milestones in quantum error correction include the discovery of topological codes [65, 66], the stabilizer formalism to describe quantum error correcting codes [51], the fault

tolerance threshold theorem [2, 39, 71, 121], and more recent ideas of dynamical decoupling [63, 126] and subsystem or operator quantum error correction [82]. We refer the reader to [102, 103] for a detailed overview of some of these important developments.

Quantum cryptography has also come a long way since the first quantum key distribution (QKD) protocol. Known as the *BB84 protocol* in honor of its creators [11], this became a prototype for later quantum cryptographic protocols. Starting from experimental implementations over a distance of 32.5 cm in 1989 [12], today we have experimental setups that have successfully demonstrated QKD over distances of hundreds of kilometers [59, 116]. An alternate view of QKD based on quantum entanglement [42] motivated newer ideas like quantum privacy amplification [35] and made the task of proving the security of QKD protocols much easier. Since then, several security proofs have been constructed [87, 95, 107, 122] leading to stronger QKD protocols that are unconditionally secure against general attacks. However, moving beyond key distribution has proved to be a big challenge in the quantum setting.

Early results showed that unconditional security was not possible for non-QKD protocols such as quantum bit commitment [33, 96] and quantum oblivious transfer [86]. Bit commitment and oblivious transfer are protocols that help to realize practical two-party functions such as online auctions or secure identification, where the participating parties do not trust each other. We will describe oblivious transfer in Section 4.4 and refer to [127] for a complete introduction to these and other two-party cryptographic tasks. In fact, it was shown in [86] that it is impossible to implement any such two-party quantum cryptographic protocol securely, without imposing some restrictions on the dishonest party.

Quantum cryptography also admits phenomena which do not have a classical analog, like the phenomenon of information locking [38]. The development of quantum cryptography has also spurred tremendous interest in the study of quantum information measures, in particular the properties of Rényi entropies and extensions thereof [74,109], which lie at the heart of most cryptographic security proofs today.

In this thesis, we focus our attention on two new paradigms that have emerged in the recent past. First is the concept of *approximate* quantum error correction which was first introduced in the work of Leung et al. [85], who showed via an example, that approximate codes might sometimes perform better than the perfect codes arising from standard constructions. Second is the idea

of *noisy storage cryptography* [76, 115] which envisions a setting where two-party cryptographic protocols can be securely implemented based solely on the assumption that the quantum storage device possessed by the adversary is noisy and bounded. In what follows, we introduce and motivate these two ideas in greater detail.

1.1 Adapting to Quantum Noise: Approximate Quantum Error Correction

We have already referred to some of the important theoretical advances in the theory of quantum error correction (QEC). While the standard paradigm of QEC is well-understood and rests on a solid mathematical foundation, the discovery of approximate codes [46,85] has shown that this standard framework might be somewhat restrictive. Approximate quantum error correction (AQEC) relaxes the constraint of perfect recovery and instead seeks codes that recover the input state with high enough fidelity. A typical quantum error correcting code is designed to perfectly correct only some subset of the errors that constitute the noise channel, in particular the ones that occur with a higher probability. Every perfect code is thus an *approximate* code for the least probable errors of the channel. These standard codes are, however, designed based on conditions that demand perfect correction of the complete noise channel. The theory of perfect quantum error correction is discussed in some detail in Section 2.1.2.

In practice, experimental setups for creating and storing qubits are usually characterized by a certain dominant noise process. For example, in many setups based on quantum optics, amplitude damping noise [26,49] is most dominant. More recently, it was observed that in some superconducting qubit systems, dephasing noise is stronger than other noise processes by a factor of 10^3 [1,3]. Standard code constructions however, do not take advantage of this specific knowledge. Rather, standard error correction proceeds by first discretizing errors in terms of the Pauli operators and then constructing codes to correct these Pauli errors perfectly. Approximate error correction, on the other hand, incorporates our knowledge of the noise model and allows us to find codes optimized for specific noise models. Thus, while the shortest perfect quantum code requires at least five qubits to encode a single qubit [84], approximate codes constructed for specific noise channels [47,85] use only four qubits to achieve comparable fidelity. This illustrates a key advantage of relaxing the

requirement for perfect QEC—one might be able to encode the same amount of information into fewer qubits while retaining a nearly identical level of protection from the noise.

These observations indicate the importance of developing a general theory of approximate quantum error correction. There are two possible approaches to characterize approximate codes. The Leung et al. approximate code was constructed as a simple perturbation of a perfect code, raising the possibility that perturbing the perfect error correction conditions might yield conditions for approximate correction [90]. Alternately, approximate quantum error correction can be formulated as an optimization problem [46,135]. Given a noise channel and the information we need to encode, AQEC is the problem of finding the optimal encoding and recovery maps, with optimality defined in terms of a chosen measure of accuracy of recovery.

In our work presented in Chapter 2, we combine both these aspects of approximate error correction via a universal recovery map, namely the *transpose channel*. On the one hand, by defining optimality in terms of the *worst-case fidelity*, we demonstrate that the transpose channel is a near-optimal recovery map for any noise channel. We also show that the perfect error correcting conditions can be rewritten in terms of the transpose channel. Combining this with our fidelity bound, we are able to write down simple conditions for approximate quantum error correction, which are indeed obtained as perturbations of the perfect QEC conditions. Earlier studies [8] had in fact shown that the transpose channel is a near-optimal recovery for an average fidelity measure based on the entanglement fidelity. Indeed, while the problem of finding optimal codes for the entanglement fidelity has been studied in the recent past [17, 45, 124], we present for the first time an analytical description of a recovery map that is close to optimal for the worst-case fidelity. The worst-case fidelity involves minimizing the fidelity between the input and output states over all input states in contrast to the average fidelity which is simply an average over input states. Optimizing for the worst-case measure thus provides a stronger assurance that *all* the information in the input space is well protected.

Our conditions for AQEC also lead to an easily implementable algorithm to identify approximate codes, for a given noise channel and fidelity threshold. Using this, we obtain good four-qubit codes for the amplitude damping channel. We thus formulate a simple, analytical approach to characterizing and finding approximate quantum codes. This is of course a first step toward a complete theory of approximate quantum error correction, with potential applications in designing

a fault tolerant quantum computing architecture.

1.2 Using the Power of Quantum Noise: Noisy-Storage Cryptography

As mentioned above, one of the early negative results in quantum cryptography was the discovery that secure two-party quantum cryptography is not possible without additional assumptions [86]. In the classical setting, the usual assumptions that go into realizing secure two-party protocols involve mathematical hardness results and a restriction on the computing power of a cheating party. An interesting physical assumption that also leads to security in the classical setting is to assume a restriction on the amount of classical storage a cheating party can use [21, 94]. However, in this classical *bounded-storage model*, the cheating party requires only quadratically more storage than the honest party to break the security. Furthermore, a tight bound on classical storage is not easily enforceable in today's context. Since storing quantum states for any significant length of time still remains a hard problem, the quantum analog of the classical bounded-storage model might be a more realistic prospect.

In this setting of bounded quantum storage, protocols for secure implementation of quantum bit commitment and oblivious transfer were demonstrated under the assumption that a cheating party cannot store any quantum information at all [14, 29]. Recently, this bound was improved, showing that quantum oblivious transfer can be securely implemented if the cheating party can store no more than a fourth of the qubits transmitted during the protocol [30, 31]. An alternate physical situation to consider is of course the case where the cheating party's quantum storage is noisy, while allowing for a larger storage size. In this quantum noisy-storage setting, a protocol for secure oblivious transfer was constructed [115,128] with the additional constraint that the cheating party could only perform product measurements on the qubits received during the protocol.

It turns out that both these models can be realized as special cases of a generalized *quantum noisy-storage model* introduced by König et al. [76], which incorporates both the amount of storage and noise. A formal description of this model is provided in Section 4.1.2. In this more general setting, protocols for secure bit commitment and oblivious transfer were constructed, which were shown to be secure for reasonable values of the noise parameter (characterizing the noisy storage)

and the storage rate of the cheating party. When specialized to the case of bounded quantum storage (when the storage is noiseless) these protocols were shown to be secure so long as the cheating party cannot store more than half the qubits transmitted. This is quite an improvement over the earlier bound obtained for the bounded-storage model.

In the bounded storage model, it is intuitively clear that if a cheating party were able to store *all* of the information transmitted during a protocol, security cannot be achieved. This is equivalent to saying that the storage rate must be less than one, if a protocol is to be secure. Conversely, one can ask the question as to whether it is always possible to achieve security if the storage rate is strictly less than unity. The protocols constructed so far do not answer this question conclusively. While the best protocol achieves security so long as the storage rate is less than a half in the bounded-storage model, the question as to whether the physical limit of one is achievable remained unresolved. We address this question in our work presented in Chapter 4, and demonstrate that it is indeed possible to achieve security if the cheating party's storage rate is strictly less than one. To achieve this physical limit, it turns out that we have to look beyond qubits and implement our protocols using higher-dimensional quantum states, that is, qu*d*its.

In the setting of the general noisy-storage model [76], we demonstrate a protocol for quantum oblivious transfer using qu*d*its, where again security depends on the noise parameter and the storage rate. Since oblivious transfer is known to be universal for two-party cryptography [64], this implies that our protocol can in principle be used to realize any two-party cryptographic task securely. In the noisy-storage setting, our protocol improves the results of [76] by extending security to a much larger class of noise channels. Most importantly, when the cheating party has a noiseless memory, our protocols are secure so long as he can store all but an arbitrarily small fraction of the qu*d*its transmitted during the protocol. Even though our result is an asymptotic one—in the sense that we achieve security for storage rates close to one only for very large values of dimension d—it is an important demonstration of the concept that the physical limit of the bounded-storage model is indeed achievable. We hope that the techniques used in achieving our bounds might provide some insight into proving similar bounds even for smaller systems or qubits encoded into higher dimensions.

1.3 A Mathematical Interlude: Uncertainty Relations for Complementary Aspects

An important technique that is employed in our security proof, and in fact in a vast majority of the existing security proofs in quantum cryptography, is the use of entropic uncertainty relations to bound the information held by a cheating party [130]. Before describing our work on the noisy-storage model, we take a brief detour in Chapter 3 to present our results on obtaining strong entropic uncertainty relations by making use of mutually unbiased bases.

Entropic uncertainty relations and mutually unbiased bases are formally defined in Sections 3.2.1 and 3.2.2 respectively. Entropic uncertainty relations (EURs) provide lower bounds on the average entropy of probability distributions corresponding to the outcomes of different measurements. They thus provide a natural way to quantify incompatibility among multiple measurements [34]. For two observables, it is well known that this incompatibility is maximum when the measurement bases are complementary or mutually unbiased [89]. For more than two measurement settings, while being mutually unbiased is a necessary condition to obtain strong uncertainty relations, it was recently shown that there do exist small sets of mutually unbiased bases (MUBs) that satisfy trivial uncertainty relations [6]. It remains an important open question to identify and construct sets of complementary bases satisfying strong uncertainty relations. We refer to [130] for a survey of the entropic uncertainty relations in different measurement scenarios and to [40] for a review of the known results on the existence and constructions of mutually unbiased bases in different dimensions.

In our work, we investigate the possibility of constructing symmetric sets of MUBs using the generators of the Clifford algebra in dimension $d = 2^n$. The symmetry of interest here is the existence of a unitary transformation that cyclically permutes the different basis sets. It has been shown that the maximal set of $d + 1$ bases in d dimensions (when d is a prime power) has such a symmetry structure, and that minimum uncertainty states in the Hilbert space are in fact invariant under the corresponding unitary transformation [132]. In Chapter 3, we present an explicit construction of sets of up to $2n+1$ MUBs in dimension $d = 2^n$ which are indeed cyclically permuted under the action of unitary transformation. This unitary can be understood as a generalization of the Fourier transform (which exchanges two MUBs) to multiple complementary aspects. We then apply our transformation to the study of uncertainty relations in terms of the min-entropy,

which also gives a lower bound on the Shannon entropy. We prove a lower bound for min-entropic uncertainty relations for any set of MUBs, and show that symmetry plays a central role in obtaining tight bounds. In fact, we obtain for the first time a tight bound for four MUBs in dimension $d = 4$, which is attained by an eigenstate of the permuting unitary transformation.

As mentioned earlier, the security of protocols in recent cryptographic models like the bounded-storage and noisy-storage models is directly related to entropic uncertainty relations. EURs also figure prominently in the security analysis of quantum key distribution [72, 106] and information locking protocols [38]. A better understanding of the interplay between complementarity and uncertainty relations is thus of interest not only from a foundational standpoint, but has practical implications for analyzing and improving existing cryptographic protocols.

1.4 Thesis Organization

We have presented here a brief overview of the problems investigated in this thesis. The rest of the thesis is organized as follows. Chapter 2 briefly introduces the standard paradigm of quantum error correction and describes our approach to approximate quantum error correction based on the transpose channel. It also discusses an algorithm to search for approximate codes and presents numerical results based on this algorithm. For the practically relevant case of qubit codes, a further simplification of our algorithm is presented in Appendix A. Chapter 3 describes our novel construction of symmetric mutually unbiased bases, as well as the new lower bounds we obtain for min-entropic uncertainty relations. For better readability, we only describe the basic idea behind our construction in the chapter, and relegate the technical details of the construction and related proofs to Appendix B. Chapter 4 describes our work on the noisy-storage model, after providing a brief introduction to some of the basic techniques used in two-party quantum cryptography. Chapters 2 and 3 are self contained and can be read independently. Chapter 4 refers to both these chapters for some preliminary concepts—in particular, it refers to Sec. 2.1.1 in Chapter 2 for the mathematical formalism of quantum noise operations, and to Secs. 3.1 and 3.2.1 in Chapter 3 for a description of entropic measures and uncertainty relations respectively.

Chapter 2

Approximate Quantum Error Correction Using the Transpose Channel

Quantum error correction (QEC) is one of the cornerstones of quantum information and quantum computing. Since quantum effects are extremely fragile and susceptible to damage by environmental noise, QEC plays an important role in making the theoretical idea of quantum information processing a physically realizable prospect. Many tasks in quantum communication or computation would indeed become impossible without the use of error-correcting techniques to protect the information from noise. The idea behind QEC is a very simple one—information is stored in a particular part of the system Hilbert space, cleverly chosen depending on the noise process affecting the system, such that a recovery operation can be applied to retrieve the information.

A vast majority of existing work on error correction focuses on the standard paradigm of *perfect* error correction, where the recovery operation either perfectly corrects the full noise channel, or perfectly corrects the errors conditioned on the fact that fewer than some t errors occurred. Recently, examples of *approximately* correcting quantum codes have been constructed, that recover the information with fidelity comparable to that of perfect QEC codes, suggesting that the requirement for perfect recovery may be too stringent for certain tasks. While the smallest known perfect quantum code requires at least five qubits to encode a single qubit [15, 84], approximate codes constructed for specific noise channels [47, 85] use only four qubits to achieve comparable fidelity. This illustrates a key advantage of relaxing the requirement for perfect QEC—one might be able to encode the same amount of information into fewer qubits while retaining a nearly identical level of

protection from the noise. While examples of good approximate QEC (AQEC) codes are known, characterizing these approximately correctable codes has remained an open problem.

In this chapter,[1] we formulate a simple approach to characterizing and finding AQEC codes. We demonstrate for the first time that there exists a universal, near-optimal recovery map for AQEC codes—**the transpose channel**—where optimality is defined in terms of the worst-case fidelity. Using the transpose channel, which is constructed as a generalization of the recovery channel defined in [8], we obtain a set of conditions for AQEC, which forms the basis of a simple algorithm for finding AQEC codes. Our analytical approach is a departure from earlier work which relies on exhaustive numerical search for the optimal recovery map, with optimality defined in terms of *average* or entanglement fidelity rather than the *worst-case* fidelity. Furthermore, for the practically useful case of codes encoding a single qubit of information, our algorithm is particularly easy to implement.

The rest of the chapter is organized as follows. In Section 2.1 we briefly review the standard paradigm of quantum error correction. In Section 2.2 we introduce the notion of approximate error correction with an example, and formulate the problem of finding AQEC codes as an optimization problem. In Section 2.3, we define the transpose channel, examine its role in standard QEC theory, and prove that it is nearly optimal for AQEC codes. An alternative form of the perfect QEC conditions based on the transpose channel is described in Section 2.4, a perturbation of which leads to a set of AQEC conditions. The algorithm for finding AQEC codes is described in Section 2.5. In Section 2.6, we consider the example of amplitude damping noise and use it to compare our procedure with earlier work on approximate codes. Section 2.7 contains our conclusions and some open problems.

2.1 Quantum Error Correction

The basic idea of quantum error correction is to *encode* the information that we wish to protect into a larger quantum system. Specifically, information is stored in quantum states, which are represented by *density operators* $\rho \in \mathcal{B}(\mathcal{H})$, in a finite-dimensional Hilbert space $\mathcal{H}$. The operator ρ is a positive semi-definite, trace-1, linear operator in $\mathcal{H}$. If the state of the system is exactly

[1] The work described in this chapter has been done in collaboration with Hui Khoon Ng. The original results presented here have been published in [97]. We would like to thank David Poulin for introducing us to the problem of approximate quantum error correction, and for many insightful discussions.

known, it is described by a *pure state* $|\psi\rangle \in \mathcal{H}$, where $|\psi\rangle$ is a unit vector in $\mathcal{H}$. Density operators corresponding to such pure states are given by $\rho = |\psi\rangle\langle\psi|$. Finally, if a system is known to be in state $|\psi_i\rangle$ with probability p_i, where $\{|\psi_i\rangle, i = 1, \ldots, N\}$ is a set of N pure states, it is said to be in a *mixed state* and the density operator describing the state of the system is given by $\rho = \sum_i p_i |\psi_i\rangle\langle\psi_i|$.

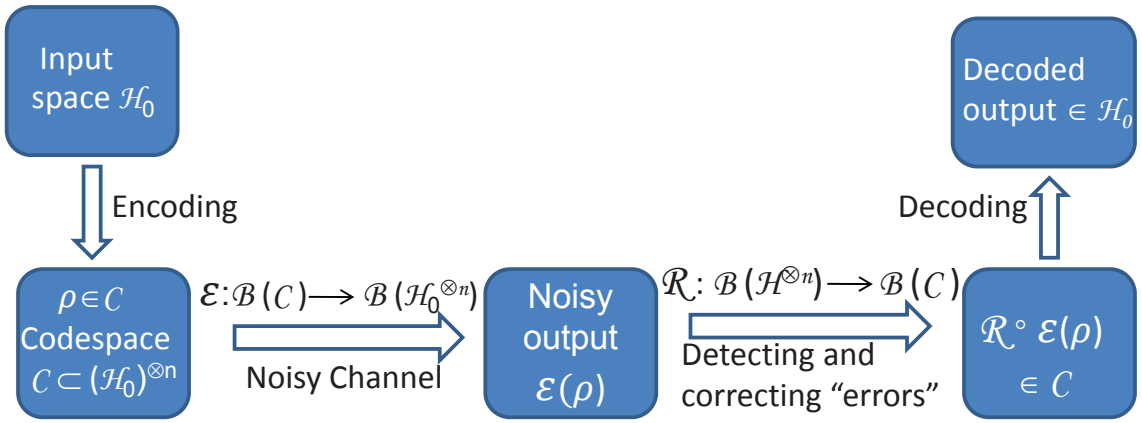


Figure 2.1: Schematic representation of quantum error correction.

Given the system Hilbert space $\mathcal{H}$, we seek to encode a qu*d*it of information—information carried by a d-dimensional Hilbert space $\mathcal{H}_0$—where d is no greater than the dimension of $\mathcal{H}$. Often the system Hilbert space $\mathcal{H}$ is simply an n-fold tensor product of the system we seek to encode, that is, $\mathcal{H} = (\mathcal{H}_0)^{\otimes n}$. The qu*d*it is encoded into a d-dimensional subspace $\mathcal{C}$ of $\mathcal{H}$. When $d = 2$, this reduces to the problem of encoding a single qubit of information, which is of utmost practical relevance today. We refer to $\mathcal{C}$ as a *subspace code*, as opposed to subsystem codes [82] or more general codes in the sense of [19]. Since we are only concerned with subspace codes in this chapter, we will often use "code" to denote the subspace $\mathcal{C}$. Formally, the information is encoded into $\mathcal{C}$ via an encoding map $\mathcal{W} : \mathcal{H}_0 \to \mathcal{C}$, whose action on any orthonormal basis $\{|\phi_i^{(0)}\rangle\}$ for $\mathcal{H}_0$ is $\mathcal{W} : |\phi_i^{(0)}\rangle \mapsto |\phi_i\rangle \in \mathcal{C}$ such that $\langle\phi_i|\phi_j\rangle = \delta_{ij}\ \forall i, j$. One can extend this encoding map on the vector space $\mathcal{H}_0$ to a completely positive (CP), trace-preserving (TP) map on operators, also denoted by $\mathcal{W}$.

After encoding the information into $\mathcal{C}$ using $\mathcal{W}$, we consider the action of noise. This noise is also described by a CPTP map $\mathcal{E} : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{H})$. $\mathcal{E}$ can describe, for example, the Markovian

noise acting on the system over some timestep, or the effect of a single use of a noisy channel for communication. After the action of $\mathcal{E}$, we perform a CPTP recovery map $\mathcal{R} : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{C})$ to undo the effects of the noise, and then decode using $\mathcal{W}^{-1}$. Note that, when a single qu*d*it (qubit) is encoded into n qu*d*its (qubits), the corresponding noise channel is also an n-fold tensor product space of a single qu*d*it (qubit) noise channel $\mathcal{E}$, denoted as $\mathcal{E}^{\otimes n}$. The key steps in the error correction process described here are summarized schematically in Fig. 2.1. We will now proceed to describe the quantum noise model in greater detail with examples of some physically motivated noise processes.

2.1.1 Quantum Channels

Formally, any quantum operation on a system can be described by a completely positive, trace preserving (CPTP) map on density operators in the system Hilbert space. A map $\mathcal{E}$ acting on density operators $\rho \in \mathcal{B}(\mathcal{H}_A)$ is said to completely positive if and only if, (a) $\mathcal{E}(\rho) > 0$, $\forall\ \rho > 0 \in \mathcal{B}(\mathcal{H}_A)$, and (b) $\mathcal{E} \otimes \mathbb{I}_B$ is a positive map on $\mathcal{B}(\mathcal{H}_A \otimes \mathcal{H}_B)$, for any possible extension $\mathcal{H}_A \otimes \mathcal{H}_B$ of the system Hilbert space $\mathcal{H}_A$, where $\mathbb{I}_B$ is the identity map on $\mathcal{H}_B$. Condition (a) is the simple statement that if $\rho \in \mathcal{H}_A$ is a valid density operator, then so is $\mathcal{E}(\rho)$ (up to normalization). The additional requirement of *complete* positivity in condition (b) is the physical requirement that $(\mathcal{E} \otimes \mathbb{I}_B)(\rho_{AB})$ also be a valid density operator for any joint state $\rho_{AB} \in \mathcal{B}(\mathcal{H}_A \otimes \mathcal{H}_B)$, where $\mathcal{E}$ acts only on the subsystem $\mathcal{H}_A$.

Since $\mathrm{tr}[\mathcal{E}(\rho)]$ is the probability that the process represented by $\mathcal{E}$ occurred, given the initial state ρ, conservation of probability requires that $\mathrm{tr}[\mathcal{E}(\rho)] = 1$ for all ρ. When $\mathcal{E}$ describes processes where some extra information is obtained by a measurement, then $\mathcal{E}$ could be non-trace-preserving, that is, $\mathrm{tr}[\mathcal{E}(\rho)] \leq 1$. But since we deal with deterministic noise processes here, the corresponding maps are indeed trace-preserving.

It turns out that these physically motivated requirements lead to an elegant mathematical description of quantum operations. The celebrated result due to Choi and Kraus [25, 79, 80] states that a map $\mathcal{E} : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{H})$ is completely positive if and only if there exists a set of operators $\{E_i\}_{i=1}^N$—referred to as a **Kraus representation** of $\mathcal{E}$—such that the action of $\mathcal{E}$ on $\rho \in \mathcal{B}(\mathcal{H})$ is given by $\mathcal{E}(\rho) = \sum_{i=1}^N E_i \rho E_i^\dagger$. Henceforth, we will denote $\mathcal{E}$ with a particular Kraus representation as $\mathcal{E} \sim \{E_i\}_{i=1}^N$, and refer to set of operators $\{E_i\}$ as a set of Kraus operators corresponding

to the map $\mathcal{E}$. Further, the Kraus representation of a CP map is not unique—any two Kraus representations $\{E_i\}_i^N$ and $\{F_k\}_i^N$ such that $F_k = \sum_{ik} u_{ik} E_i$, for some $N \times N$ unitary matrix (u_{ij}), describe the same map [98, Theorem 8.2]. Finally, the fact that $\mathcal{E}$ is trace-preserving implies that the Kraus operators of $\mathcal{E}$ satisfy

$$\sum_i E_i^\dagger E_i = \mathbb{I},$$

where $\mathbb{I}$ is the identity operator for the domain of $\mathcal{E}$.

To summarize, quantum noise processes or quantum channels are modeled as CPTP maps on the system Hilbert space, with the individual errors given by the Kraus operators in the operator-sum representation described above. We conclude this section with some concrete examples of single qubit quantum channels.

(i) **Bit flip channel:** The simplest quantum channel we can construct is one that simply flips the state of a qubit from $|0\rangle$ to $|1\rangle$ with probability $1-p$. This is a straightforward generalization of the corresponding classical channel which flips the value of a bit with probability $1-p$, to the quantum setting. In the $\{|0\rangle, |1\rangle\}$ basis, the Kraus operators of this channel are given by

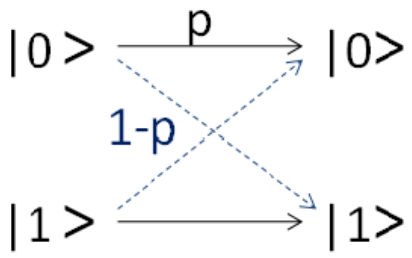


Figure 2.2: The quantum bit flip channel.

$$E_0 = \sqrt{p}\,\mathbb{I} = \sqrt{p}\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}; \qquad E_1 = \sqrt{1-p}\,X = \sqrt{1-p}\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.$$

The quantum bit flip channel thus describes the physical process when the system is either affected by a Pauli X operation with probability $1-p$, or left unaffected with probability p.

(ii) **Depolarizing channel:** This describes the scenario where a single qubit is replaced by the maximally mixed (*depolarized*) state with probability p. Formally, the action of the channel

on any state ρ can be described in terms of the Pauli operators as

$$\mathcal{E}_{\text{Dep}}(\rho) = (1 - \frac{3p}{4})\rho + \frac{p}{4}(X\rho X + Y\rho Y + Z\rho Z) \ . \tag{2.1}$$

The Kraus operators corresponding to this channel are therefore given by $\mathcal{E}_{\text{Dep}}(\rho) \sim \{\sqrt{1-3p/4}\mathbb{I}, \sqrt{p}X/2, \sqrt{p}Y/2, \sqrt{p}Z/2\}$.

(iii) **Amplitude damping channel:** This is an important quantum channel that describes energy dissipation and characterizes the effects due to loss of energy from a quantum system. The single qubit amplitude damping channel $\mathcal{E}_{\text{AD}}$ has Kraus operators

$$E_0^{\text{AD}} = \begin{pmatrix} 1 & 0 \\ 0 & \sqrt{1-\gamma} \end{pmatrix}, \quad \text{and} \quad E_1^{\text{AD}} = \begin{pmatrix} 0 & \sqrt{\gamma} \\ 0 & 0 \end{pmatrix}, \tag{2.2}$$

written in some qubit basis $\{|0\rangle, |1\rangle\}$. $\mathcal{E}_{\text{AD}}$ can be thought of as describing energy dissipation in a two-level system, where $|0\rangle$ is the ground state and $|1\rangle$ is some excited state. γ is then the probability of a transition from the excited state to the ground state. In the Pauli basis,

$$E_0^{\text{AD}} = \frac{1}{2}[(1+\sqrt{1-\gamma})\ \mathbb{I} + (1-\sqrt{1-\gamma})\ Z] \ , \ E_1^{\text{AD}} = \frac{\sqrt{\gamma}}{2}[X + iY] \ ,$$

showing that no linear combination of E_0 and E_1 can give an operation element proportional to the identity operator. The fact that the operator elements of the amplitude damping channel cannot be realized as scaled Pauli operators makes it different from the other two channels described here.

2.1.2 Perfect QEC Conditions

The problem of quantum error correction is to pick the right encoding and recovery operation for a given noise channel $\mathcal{E}$, such that the recovery operation either perfectly corrects the full CPTP noise channel, or perfectly corrects the errors conditioned on the fact that fewer than t errors occurred. We will henceforth refer to this standard paradigm of error correction as *perfect* QEC. As depicted in Fig. 2.1, it is common to fix the encoding Hilbert space as the n-fold tensor product space of the system Hilbert space, so that the problem of perfect QEC reduces to the problem of choosing the right pair $(\mathcal{C}, \mathcal{R})$ of codespace and recovery map such that the action of the map $\mathcal{E}$ followed by

$\mathcal{R}$ leaves states in the codespace $\mathcal{C}$ unaffected. Formally, the action of a channel $\mathcal{E}$ is said to be *perfectly correctible* on a codespace $\mathbb{C}$ if and only if there exists a quantum channel $\mathcal{R}$ such that $\mathcal{R} \circ \mathcal{E}(\rho) \ = \ \rho \, \forall \, \rho \ \in \ \mathbb{C}$.

An important characterization of perfect QEC codes is given by the perfect QEC conditions [15, 43, 70], which we briefly review here. For a given code $\mathcal{C}$, let P be the projector onto $\mathcal{C}$. The QEC conditions can then be stated as follows.

Theorem 2.1.1 (**Perfect QEC conditions** [15, 43, 70])**.** *A CPTP recovery map $\mathcal{R}$ that perfectly corrects the action of a noise channel $\mathcal{E} \sim \{E_i\}_{i=1}^{N}$ on a subspace code $\mathcal{C}$ exists, if and only if*

$$\forall i, j, \quad P E_i^\dagger E_j P = \alpha_{ij} P, \tag{2.3}$$

for some complex matrix α.

Note that this is simply a condition on the existence of a perfect QEC code for a given $\mathcal{E}$, and stipulates no knowledge of the recovery map $\mathcal{R}$.

It is more insightful to rewrite (2.3) in a "diagonal" form. From (2.3), it is clear that α must be a Hermitian matrix. Therefore, there exists a unitary u and a diagonal matrix d such that $\alpha = udu^\dagger$. The set of operators defined by $F_k \equiv \sum_i u_{ik} E_i$ constitutes an alternate Kraus representation for $\mathcal{E}$, so that $\mathcal{E} \sim \{F_k\}$. With this choice of Kraus representation, the perfect QEC condition takes the following form:

$$\forall k, l, \quad P F_k^\dagger F_l P = \delta_{kl} d_{kk} P, \tag{2.4}$$

where d_{kk} are the diagonal entries of d, or equivalently, the eigenvalues of α. Notice that $d_{kk} \geq 0, \ \forall k$, since the left-hand side of (2.4) is positive semi-definite when $k = l$. α is hence a positive semi-definite matrix ($\alpha \geq 0$).

The diagonal form of the perfect QEC condition makes it easier to appreciate the intuition behind Theorem 2.1.1. Using polar decomposition, we can express the F_k's as $F_k P = \sqrt{d_{kk}} U_k P$, for unitary U_k. The effect of the Kraus operators on a perfect code is therefore unitary, in the sense that the operator F_k simply rotates the codespace $\mathcal{C}$ into the subspace defined by the projector $P_k \equiv U_k P U_k^\dagger$. Equation (2.4) further guarantees that these projectors are orthogonal, that is,

$$P_k P_l = U_k P U_k^\dagger U_l P U_l^\dagger = \frac{U_k P F_k^\dagger F_l P U_l^\dagger}{d_{kk} d_{ll}} = 0, \ \forall \ k \neq l,$$

thus implying that the individual errors due to the action of the channel can be reliably distinguished by a projective measurement. The individual error operators of a noise channel thus map a perfectly correctible codespace to mutually orthogonal subspaces of the encoding Hilbert space, as shown in Fig. 2.3.

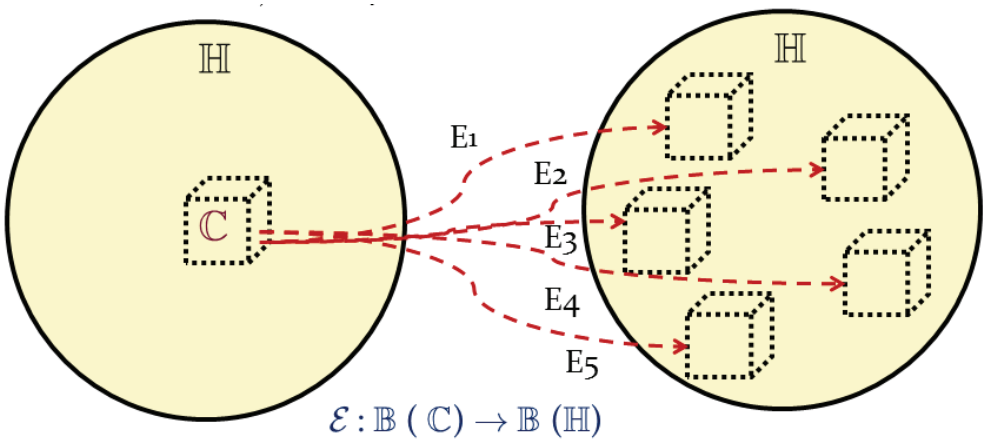


Figure 2.3: Action of noise on perfectly correctible code.

The recovery map for correcting the errors when (2.3) is satisfied, is now easy to construct. Let us denote this recovery as $\mathcal{R}_{\text{perf}}$. To write down $\mathcal{R}_{\text{perf}}$, we use the polar decomposition to obtain the unitaries U_k such that $F_k P = \sqrt{d_{kk}} U_k P$. Then, $\mathcal{R}_{\text{perf}} : \mathcal{B}(\mathcal{P}_{\mathcal{E}}) \to \mathcal{B}(\mathcal{C})$ is given by

$$\mathcal{R}_{\text{perf}} \sim \{P U_k^\dagger\}.$$

One can check that $\mathcal{R}_{\text{perf}}$ is TP on its domain $\mathcal{B}(\mathcal{P}_{\mathcal{E}})$, and that it perfectly corrects the code in the sense that for any $\rho \in \mathcal{B}(\mathcal{C})$,

$$(\mathcal{R}_{\text{perf}} \circ \mathcal{E})(\rho) = \Big(\sum_k d_{kk} \Big) \rho. \tag{2.5}$$

$\sum_k d_{kk}$ is just the trace of $\mathcal{E}(\rho)$ for any $\rho \in \mathcal{C}$. This sum is independent of ρ because of the QEC conditions (2.4), and is exactly equal to 1 if and only if $\mathcal{E}$ is TP on $\mathcal{C}$. Equation (2.5) thus implies that $\mathcal{R}_{\text{perf}}$ recovers the original code state, up to any reduction in trace due to the possible non-TP nature of $\mathcal{E}$.

The perfect QEC condition can be used to identify good error-correcting codes since it is easily

checkable for a given codespace $\mathcal{C}$. Furthermore, (2.3) is *linear*. If $\mathcal{E} \sim \{E_i\}$ is correctible on codespace $\mathcal{C}$, then any channel whose operator elements are linear combinations of $\{E_i\}$ is also correctible. Since the Pauli matrices form a basis for 2×2 matrices, this linearity property has the important consequence that for correcting single qubit errors, it suffices to check that a given code satisfies the condition for the "Pauli errors," namely the operators $(\mathbb{I}, X, Y, Z)$.

It turns out that the smallest code capable of perfectly correcting an arbitrary error on any single qubit of the system, requires five qubits. This can be shown to follow from the linearity of the error-correcting condition (2.3) and the assumption that errors act *independently* on different qubits. We refer to [98, Section 10.3.4] for detailed proofs that any general quantum code that seeks to correct single qubit errors perfectly, must encode into atleast 5 qubits. The **five-qubit code** [15, 84] is thus the shortest known perfect QEC code. We will henceforth refer to this as the as the $[[5, 1, 3]]$ code, where the first entry in the brackets corresponds to the number of qubits in the system, and the second entry is the number of qubits of information encoded in the system. The third entry is the *distance* of the code, defined as $d = 2t + 1$ where t is the maximum number of qubits that the code can perfectly correct. Since the five-qubit code is capable of correcting any error on a qubit, its distance parameter is equal to 3. This code satisfies the perfect QEC conditions for any noise channel $\mathcal{E}^{\otimes 5}$, but with terms corresponding to more than a single-qubit (Pauli) error discarded.

2.2 Approximate Quantum Error Correction

The vast majority of existing work on error correction focuses on *perfect* QEC described above. However, the example of a code designed for correcting errors affected by weak amplitude damping noise presented in [85] suggests that the requirement for perfect recovery may be too stringent for certain tasks. While perfect QEC requires at least five qubits to encode a single qubit, the code in [85] uses only four qubits to achieve comparable fidelity. This illustrates a key advantage of relaxing the requirement for perfect QEC—one might be able to encode the same amount of information into fewer qubits while retaining a nearly identical level of protection from the noise process. The four-qubit code is also specially designed for the channel in question, a departure from standard QEC codes that seek to perfectly correct up to some t *arbitrary* errors on the system. This adaptation of the code to the noise channel, an idea emphasized later in [45], is a crucial factor

behind the success of their code. Such *approximate* QEC (AQEC) codes reveal the possibility of designing codes that are better tailored to the particular information processing task at hand. Before proceeding to analyze the problem of characterizing such approximately correcting codes, it will help to gain some intuition into the working of this four-qubit code.

2.2.1 The Approximate [4,1] Code

The four-qubit code constructed by Leung et al. [85] protects a single qubit of information against amplitude damping noise by encoding into four physical qubits. Assuming that the noise acts independently on the qubits, the four-qubit noise channel is just four copies of $\mathcal{E}_{\mathrm{AD}}$, that is, $\mathcal{E}_{\mathrm{AD}}^{\otimes 4}$. The four-qubit subspace code constructed in [85] is the span of the following two states:

$$|0_L\rangle \equiv \frac{1}{\sqrt{2}}\left(|0000\rangle + |1111\rangle\right), \quad \text{and} \quad |1_L\rangle \equiv \frac{1}{\sqrt{2}}\left(|0011\rangle + |1100\rangle\right). \tag{2.6}$$

$|0_L\rangle$ and $|1_L\rangle$ respectively represent the $|0\rangle$ and $|1\rangle$ states of the single qubit of information we want to encode in the four-qubit Hilbert space. We denote this as the $[4,1]$ code, where as before, the first entry in the brackets corresponds to the number of qubits in the system, and the second entry is the number of qubits of information encoded in the system. It was shown in [85] that this code satisfies the perfect QEC conditions for $\mathcal{E}_{\mathrm{AD}}^{\otimes 4}$, except for small corrections of order γ^2. If $P_L = |0_L\rangle\langle 0_L| + |1_L\rangle\langle 1_L|$ denotes the projector onto the codespace, the Kraus operators $\{E_i^{\mathrm{AD}}\}$ corresponding to the four-qubit channel $\mathcal{E}_{\mathrm{AD}}^{\otimes 4}$ satisfy

$$P_L (E_i^{\mathrm{AD}})^\dagger E_j^{\mathrm{AD}} P_L = 0 \ , \ i \neq j; P_L (E_i^{\mathrm{AD}})^\dagger E_i^{\mathrm{AD}} P_L = P_L D_i P_L, \tag{2.7}$$

where

$$D_i = \begin{pmatrix} d_i^{(1)} & 0 \\ 0 & d_i^{(2)} \end{pmatrix}, \ |d_i^{(1)} - d_i^{(2)}| \leq \mathrm{O}(\gamma^2).$$

Comparing (2.7) with the perfect QEC condition (2.3), we see that here, while the Kraus operators map the codespace (defined in (2.6)) to mutually orthogonal subspaces, these subspaces are *not* unitary transformations of the codespace. This is schematically represented in Fig. 2.4.

As in the case of perfect QEC, a recovery operation similar to $\mathcal{R}_{\mathrm{perf}}$ can be constructed to approximately correct for single qubit errors in the $[4,1]$ code. Polar decomposition of the Kraus operators $E_k^{\mathrm{AD}} P_L = U_k^L \sqrt{P_L D_k P_L}$ yields the unitaries $\{U_k^L\}$. The *Leung recovery* map is then

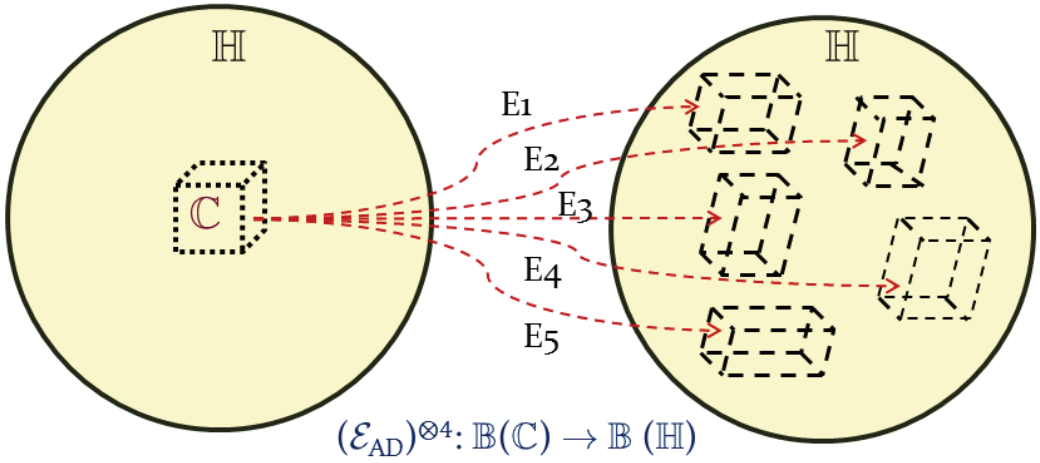


Figure 2.4: Action of $\mathcal{E}_{\mathrm{AD}}\otimes 4$ on the $[4,1]$ code.

given by $R_L \equiv \{P_L(U_k^L)^\dagger\}$. We compare the performance of the $[4,1]$ code with that of the perfect $[[5,1,3]]$ code and other approximate codes in Section 2.6.

2.2.2 AQEC as an Optimization Problem

While the analysis in [85] is based on investigating small perturbations of the perfect QEC conditions, recent work has focused on solving AQEC as an optimization problem. Indeed the challenge of AQEC is to find the optimal encoding and recovery maps, given a noise channel and the information we want to encode (qubit or higher-dimensional object), with optimality defined in terms of a chosen measure of faithfulness between the input state and the recovered state.

The *fidelity* between the input qu*d*it state and the decoded state after noise and recovery quantifies how well the information is protected from the noise. The fidelity between any two states ρ and σ is given by

$$F(\rho,\sigma) \equiv \mathrm{tr}\sqrt{\rho^{1/2}\sigma\rho^{1/2}},$$

which for a pure state $\rho \equiv |\psi\rangle\langle\psi|$, can be written as

$$F(|\psi\rangle,\sigma) \equiv F(|\psi\rangle\langle\psi|,\sigma) = \sqrt{\langle\psi|\sigma|\psi\rangle}.$$

For any ρ and σ, $F(\rho,\sigma)$ takes value between 0 and 1. $F=0$ if and only if ρ and σ have orthogonal support, and $F=1$ if and only if $\rho=\sigma$. The fidelity hence gives a measure of how close two states are.

We say that a code $\mathcal{C}$, together with its encoding and recovery maps, is effective at protecting the information from the noise $\mathcal{E}$ if the *worst-case fidelity*

$$\min_{\rho\in\mathcal{S}(\mathcal{H}_0)} F\big[\rho,\left(\mathcal{W}^{-1}\circ\mathcal{R}\circ\mathcal{E}\circ\mathcal{W}\right)(\rho)\big] \tag{2.8}$$

is close to 1. Here, $\mathcal{S}(\mathcal{H}_0)$ denotes the set of all states, pure or mixed, of the qu*d*it. In practice, it suffices to minimize over pure states in $\mathcal{S}(\mathcal{H}_0)$ only, since the fidelity measure is jointly concave in its arguments. Concavity implies that for any probability distribution p_i and density operators ρ_i and σ_i,

$$F\Big(\sum_i p_i\rho_i,\sum_i p_i\sigma_i\Big)\geq\sum_i p_i F\left(\rho_i,\sigma_i\right), \tag{2.9}$$

where $\sum_i p_i=1$. Suppose $\rho_i=|\psi_i\rangle\langle\psi_i|$ for some set of pure states $|\psi_i\rangle\in\mathcal{H}_0$, and define $\rho\equiv\sum_i p_i|\psi_i\rangle\langle\psi_i|$. Then, for any CPTP map Φ, (2.9) implies

$$\begin{aligned}
F\left[\rho,\Phi(\rho)\right] &\geq \sum_i p_i F\left[|\psi_i\rangle,\Phi\left(|\psi_i\rangle\langle\psi_i|\right)\right]\\
&\geq \Big(\sum_i p_i\Big)\min_{|\psi\rangle\in\mathcal{H}_0} F\left[|\psi\rangle,\Phi(|\psi\rangle\langle\psi|)\right]\\
&= \min_{|\psi\rangle\in\mathcal{H}_0} F\left[|\psi\rangle,\Phi(|\psi\rangle\langle\psi|)\right].
\end{aligned}$$

Since this is true for all states $\rho\in\mathcal{S}(\mathcal{H}_0)$, the minimum fidelity is clearly attained on a pure state. Setting $\Phi\equiv\mathcal{W}^{-1}\circ\mathcal{R}\circ\mathcal{E}\circ\mathcal{W}$, we see that the minimization in (2.8) needs to be performed only over pure states.

Equation (2.8) defines the worst-case fidelity for a given encoding map (or equivalently a given code $\mathcal{C}\subset\mathcal{H}$) and a given recovery map. In reality, one wants to maximize the error correction capability by choosing $\mathcal{W}$ and $\mathcal{R}$ such that the worst-case fidelity is as close to 1 as possible. The problem of AQEC can thus be phrased as the following triple optimization problem:

$$\max_{\mathcal{W}}\ \max_{\mathcal{R}}\ \min_{|\psi\rangle\in\mathcal{H}_0} F\left[|\psi\rangle,\left(\mathcal{W}^{-1}\circ\mathcal{R}\circ\mathcal{E}\circ\mathcal{W}\right)(|\psi\rangle\langle\psi|)\right]. \tag{2.10}$$

If the quantity in (2.10) attains the maximum possible value of 1, that is, if there exist $\mathcal{W}$ and $\mathcal{R}$ such that the worst-case fidelity is 1, then we have perfect QEC. In reality, one should also allow $\mathcal{H}$ to vary, and choose the smallest possible $\mathcal{H}$ that can accommodate a code with good fidelity performance. For example, in the case of a system consisting of n quantum registers, one would like to minimize n to reduce resource requirements. Choosing a Hilbert space that is too small might however reduce the worst-case fidelity of possible codes, so one would need to seek an optimal balance between having a small n and having high fidelity.

A simple approach to estimate (2.10) is to fix either the encoding or the recovery map, and then perform the optimization over the remaining two variables—the recovery or the encoding map, and the input state. Past work on finding optimal AQEC codes [46, 48, 77, 78, 105] has for the most part focused on the simpler problem of optimizing for measures based on *entanglement fidelity* [117], which characterize the performance of the code averaged over some input ensemble (including the case of a trivial ensemble comprising a single state). This eliminates the minimization over all input states required for the worst-case fidelity measure. The task of finding the optimal encoding or recovery map is then tractable via convex-optimization methods, but the resulting recovery is now optimal for an averaged measure of fidelity. Recovery maps which are near-optimal for the average entanglement fidelity have been constructed analytically [8, 124]. Conditions for AQEC based on the worst-case entanglement fidelity have also been formulated recently [17].

For many communication or computational tasks, however, one would prefer an assurance that *all* the information stored in the code is wellprotected. In such cases, the *worst-case fidelity* defined above (2.8) is the appropriate measure for determining the optimality of encoding and recovery maps. The resulting double-optimization problem for a given encoding map was examined using semi-definite programming in [135]. This method however requires a relaxation of one of the constraints in the problem, so the recovery map found is typically suboptimal. Furthermore, the numerically computed recovery map is difficult to describe and understand analytically.

In what follows, we will assume that a system of fixed size is available for encoding the information, and we search for good codes within the Hilbert space of that system. The simplest approach for finding the best code involves an exhaustive search over all possible encodings, which amounts to randomly choosing a d-dimensional subspace $\mathcal{C}$ in $\mathcal{H}$. For each $\mathcal{C}$, we still need to optimize over $\mathcal{R}$ and $|\psi\rangle \in \mathcal{H}_0$ to obtain the largest worst-case fidelity. From the form of $\mathcal{W}$, it is evident that,

given $\mathcal{W}$ and $\mathcal{R}$, the worst-case fidelity can as well be computed over states in $\mathcal{C}$ instead of $\mathcal{H}_0$. Therefore, the relevant optimization problem is

$$\max_{\mathcal{R}} \min_{|\psi\rangle \in \mathcal{C}} F\left[|\psi\rangle, (\mathcal{R} \circ \mathcal{E})\left(|\psi\rangle\langle\psi|\right)\right]. \tag{2.11}$$

Estimating (2.11) for a given code space $\mathcal{C}$ is a difficult problem since it involves a double optimization. In our work, we approach the problem stated in (2.11) using a universal recovery map that is analytically very simple to write down, and provably near-optimal, with optimality defined with respect to the worst-case fidelity.

Before proceeding further, let us define some useful terminology. We will often make use of the square of the fidelity, which we denote as $F^2(\cdot,\cdot) \equiv [F(\cdot,\cdot)]^2$. Whenever it is unambiguous, we will also refer to F^2 as the fidelity. It is also convenient to define the *fidelity loss* $\eta_{\mathcal{R}}$, for a given code $\mathcal{C}$ and a recovery map $\mathcal{R}$, as the deviation of the square of the worst-case fidelity from 1, that is,

$$\eta_{\mathcal{R}} \equiv 1 - \min_{|\psi\rangle \in \mathcal{C}} F^2\left[|\psi\rangle, (\mathcal{R} \circ \mathcal{E})(|\psi\rangle\langle\psi|)\right]. \tag{2.12}$$

The fidelity loss for the optimal recovery map $\mathcal{R}_{\text{op}}$ is denoted by η_{op}, and is given by $\eta_{\text{op}} = \min_{\mathcal{R}} \eta_{\mathcal{R}}$ for a given $\mathcal{C}$ (which is just a restatement of (2.11)). We refer to η_{op} as the *optimal fidelity loss*. A code $\mathcal{C}$ for $\mathcal{E}$ is said to be ϵ-*correctable* if it has $\eta_{\text{op}} \leq \epsilon$ for some $\epsilon \in [0,1]$. ϵ-correctable codes with $\epsilon \ll 1$ are said to be approximately correctable, and have states with fidelity at least $\sqrt{1-\epsilon} \simeq 1 - \epsilon/2$ after the action of the noise and recovery.

2.3 The Transpose Channel

Using the worst-case fidelity measure to define optimality, and assuming a fixed encoding, we now demonstrate a universal recovery map—the transpose channel—which gives a worst-case fidelity that can be suboptimal, but cannot be too far from that of the optimal recovery. First, we establish that this transpose channel is exactly the standard recovery map for perfect QEC codes characterized by the QEC condition (2.3). Then, we show that the transpose channel is nearly optimal even in the case of AQEC codes.

We begin this section with a description of the transpose channel. For a given code $\mathcal{C}$, let P be the projector onto $\mathcal{C}$ (a subspace). Let $\mathcal{P}_{\mathcal{E}} \equiv \text{supp}[\mathcal{E}(P)]$, and $P_{\mathcal{E}}$ denote the projector onto $\mathcal{P}_{\mathcal{E}}$.

Let $\{E_i\}_{i=1}^N$ be a Kraus representation for $\mathcal{E}$. The transpose channel $\mathcal{R}_P : \mathcal{B}(\mathcal{P}_\mathcal{E}) \to \mathcal{B}(\mathcal{C})$ for the given $\mathcal{C}$ is defined as the following CPTP map:

$$\mathcal{R}_P(\cdot) \equiv \sum_{i=1}^{N} PE_i^\dagger \mathcal{E}(P)^{-1/2}(\cdot)\mathcal{E}(P)^{-1/2} E_i P,$$

i.e., $\mathcal{R}_P \sim \{PE_i^\dagger \mathcal{E}(P)^{-1/2}\}_{i=1}^N$. The inverse of $\mathcal{E}(P)$ is taken on its support $\mathcal{P}_\mathcal{E}$. $\mathcal{R}_P$ has this universal form for any channel $\mathcal{E}$ and any code $\mathcal{C}$, and depends on $\mathcal{C}$ only through P. $\mathcal{R}_P$ is a special case of a recovery map introduced in [8] for reversing the effects of a quantum channel on a given initial state. In fact the $\mathcal{R}_P$ is exactly the recovery map for the initial state P/d, where d is the dimension of $\mathcal{C}$. In [19], $\mathcal{R}_P$ was shown to be useful for correcting information carried by codes preserved according to an operationally motivated notion. The term transpose channel owes its origin to [99], where this channel was first defined in an information-theoretic context. It was shown [101] that the transpose channel has the property of being the unique noise channel that saturates Uhlmanns theorem on the monotonicity of relative entropy – a fact that was later used to characterize states that saturate the strong subadditivity of quantum entropy [55].

Observe that the Kraus operators of $\mathcal{R}_P$ satisfy

$$\sum_i (PE_i^\dagger \mathcal{E}(P)^{-1/2})^\dagger (PE_i^\dagger \mathcal{E}(P)^{-1/2}) = P_\mathcal{E},$$

so $\mathcal{R}_P$ is TP on its domain, $\mathcal{B}(\mathcal{P}_\mathcal{E})$. Note that we can always add an additional projector $(\mathbb{I} - P_\mathcal{E})$ – corresponding to doing nothing on the complement of $\mathcal{P}_\mathcal{E}$—to the Kraus operators of $\mathcal{R}_P$, thus rendering it TP on the full $\mathcal{H}$ and making it a valid physical operation on the system. However, since we assume that the information is encoded completely within the code space, the action of $\mathcal{R}_P$ outside $\mathcal{P}_\mathcal{E}$ is irrelevant, so we can ignore this extension outside $\mathcal{P}_\mathcal{E}$.

We can understand the transpose channel as being composed of three CP maps: $\mathcal{R}_P = \mathcal{P} \circ \mathcal{E}^\dagger \circ \mathcal{N}$, where $\mathcal{P}$ is the projection $P(\cdot)P$ onto $\mathcal{C}$, and $\mathcal{N}$ is the normalization map $\mathcal{N}(\cdot) = \mathcal{E}(P)^{-1/2}(\cdot)\mathcal{E}(P)^{-1/2}$. In this form, $\mathcal{R}_P$ is manifestly independent of the choice of Kraus representation for $\mathcal{E}$. Without the map $\mathcal{N}$, $\mathcal{R}_P$ is just the adjoint map $\mathcal{E}^\dagger \sim \{E_i^\dagger\}$ with an additional projection to ensure that we end up in $\mathcal{B}(\mathcal{C})$. However, $\mathcal{P} \circ \mathcal{E}^\dagger$ is not TP, and $\mathcal{N}$ is added precisely to remedy that.

While we will mainly use $\mathcal{R}_P$ to discuss AQEC codes, understanding the relevance of $\mathcal{R}_P$ to

perfect QEC codes provides the intuition behind the AQEC conditions presented later.

2.3.1 The Transpose Channel for Perfect QEC

A natural question to ask here is how the transpose channel $\mathcal{R}_P$ relates to the recovery $\mathcal{R}_{\text{perf}}$ for a given $\mathcal{E}$ and $\mathcal{C}$ that satisfy the QEC conditions. Here, we show that they are exactly the same map, as previously noted in [8].

Lemma 2.3.1 ([8]). *Given a channel $\mathcal{E} \sim \{E_i^\dagger\}$ on codespace $\mathcal{C}$ satisfying the perfect QEC conditions* (2.3)*, $\mathcal{R}_P = \mathcal{R}_{perf}$.*

Proof. First, recall that the Kraus operators of $\mathcal{R}_P$ are given by $PF_k^\dagger \mathcal{E}(P)^{-1/2}$. Observe that $\mathcal{E}(P) = \sum_k (F_k P)(PF_k^\dagger) = \sum_k d_{kk} U_k P P U_k^\dagger = \sum_k d_{kk} P_k$, where $P_k \equiv U_k P U_k^\dagger$. Equation (2.4) tells us that $PU_k^\dagger U_l P = \delta_{kl} P$, implying that the P_k's are orthogonal projectors satisfying $P_k P_l = \delta_{kl} P_k$. Hence, $\mathcal{E}(P)^{-1/2} = \sum_k P_k / \sqrt{d_{kk}}$, where the inverse is taken on the support $P_{\mathcal{E}} = \sum_k P_k$. Then, we can write

$$\begin{aligned} PF_k^\dagger \mathcal{E}(P)^{-1/2} &= PF_k^\dagger \sum_l \frac{P_l}{\sqrt{d_{ll}}} \\ &= \sum_l \sqrt{\frac{d_{kk}}{d_{ll}}} PU_k^\dagger U_l P U_l^\dagger = PU_k^\dagger, \end{aligned} \tag{2.13}$$

which are exactly the Kraus operators of $\mathcal{R}_{\text{perf}}$. Thus, we see that when the perfect QEC conditions are satisfied, $\mathcal{R}_P$ is exactly the optimal recovery map that perfectly corrects $\mathcal{E}$ on $\mathcal{C}$. ■

Note that Theorem 2.1.1 and Lemma 2.3.1 remain true even for an $\mathcal{E}$ that is not TP. Traditionally, perfect QEC is discussed for a noise channel $\mathcal{E}$ that is CP but not necessarily TP. The non-TP case is particularly relevant when we deal with a system of n quantum registers, with each register independently affected by some noise $\mathcal{E}_1$. Then, instead of requiring the code to correct the entire noise channel $\mathcal{E}_1^{\otimes n}$, one often looks for codes that perfectly correct the noise up to some maximum number t of quantum registers with errors. In this case, we take $\mathcal{E}$ as the channel describing noise where at most t registers have errors, instead of the full noise channel $\mathcal{E}_1^{\otimes n}$. Such an $\mathcal{E}$ is not TP, since we have discarded the part of $\mathcal{E}_1^{\otimes n}$ that corresponds to having errors in more than t registers. This gives rise to the notion of the *distance* of a code, inherited from the theory of classical codes, which is given by $2t + 1$ for a code that corrects a maximum of t errors.

Actually, a perfectly correctable code for such a non-TP noise channel can be viewed as an approximately correctable code for the original n-register noise channel $\mathcal{E}_1^{\otimes n}$, which *is* TP. In our AQEC discussion, the code we look for is approximately correctable on the channel anyway, so $\mathcal{E}$ is always assumed to be TP, which is the physically relevant scenario. Note that the analysis in the remainder of the paper does apply for a special type of non-TP maps—$\mathcal{E} \sim \{E_i\}$ satisfying $\sum_i PE_i^\dagger E_i P = aP$, where P is the projector onto the code space and $0 \leq a \leq 1$. Our analysis applies in this case, except that one would have to add the proportionality factor a to our expressions.

2.3.2 Near-Optimality of the Transpose Channel

For AQEC codes, while the transpose channel $\mathcal{R}_P$ need not be the optimal recovery map $\mathcal{R}_{\rm op}$, we show that $\mathcal{R}_P$ does nearly as well as $\mathcal{R}_{\rm op}$. This is our central result and forms the basis of much of the discussion that follows.

Theorem 2.3.2. *Given a subspace code $\mathcal{C}$ of dimension d and optimal fidelity loss η_{op}, for any $|\psi\rangle \in \mathcal{C}$,*

$$\begin{aligned} &F^2\left[|\psi\rangle, (\mathcal{R}_{op} \circ \mathcal{E})(|\psi\rangle\langle\psi|)\right] \\ &\leq \sqrt{1 + (d-1)\eta_{op}}\; F\left[|\psi\rangle, (\mathcal{R}_P \circ \mathcal{E})(|\psi\rangle\langle\psi|)\right]. \end{aligned} \tag{2.14}$$

Proof. Let $\{R_j\}$ be the Kraus operators of $\mathcal{R}_{\rm op} : \mathcal{B}(\mathcal{P}_\mathcal{E}) \to \mathcal{B}(\mathcal{C})$. For any $|\psi\rangle \in \mathcal{C}$, following [8], we have,

$$F^2\left[|\psi\rangle, (\mathcal{R}_{\rm op} \circ \mathcal{E})(|\psi\rangle\langle\psi|)\right] \leq \sqrt{\Big[\sum_i |\langle E_i^\dagger \mathcal{E}(P)^{-1/2} E_i\rangle|^2\Big]\Big[\sum_j |\langle R_j^\dagger \mathcal{E}(P)^{1/2} R_j^\dagger\rangle|^2\Big]}, \tag{2.15}$$

where $\langle . \rangle$ denotes the expectation value with respect to the state $|\psi\rangle$. Since $\mathcal{R}_{\rm op}$ is TP, we have that $\sum_j |\langle R_j^\dagger \mathcal{E}(P)^{1/2} R_j^\dagger\rangle|^2 \leq \langle \sum_j R_j^\dagger \mathcal{E}(P) R_j\rangle = \langle (\mathcal{R}_{\rm op} \circ \mathcal{E})(P)\rangle$.

Now, we choose a basis $\{|\psi_i\rangle\}_{i=1}^d$ for $\mathcal{C}$ such that $|\psi_1\rangle \equiv |\psi\rangle$. Let $\rho_i \equiv (\mathcal{R}_{\rm op} \circ \mathcal{E})(|\psi_i\rangle\langle\psi_i|) = \sum_{kl} \alpha_{kl}^{(i)} |\psi_k\rangle\langle\psi_l|$, where the coefficients $\alpha_{kl}^{(i)}$ satisfy the normalization condition $\sum_k \alpha_{kk}^{(i)} = 1$ and $\alpha_{kk}^{(i)} \geq 0,\ \forall\, k$. From the definition of the optimal fidelity loss $\eta_{\rm op}$ (2.12), we know that $\alpha_{ii}^{(i)} = \langle\psi_i|\rho_i|\psi_i\rangle = F^2\left[|\psi_i\rangle, (\mathcal{R}_{\rm op} \circ \mathcal{E})(|\psi_i\rangle\langle\psi_i|)\right] \geq 1 - \eta_{\rm op}$. This, together with the normalization condition, implies that $\sum_{k\neq i} \alpha_{kk}^{(i)} \leq \eta_{\rm op}$, which in turn tells us that $\alpha_{kk}^{(i)} \leq \eta_{\rm op},\ \forall\ k \neq i$. Since $|\psi\rangle = |\psi_1\rangle$ by

construction, we get

$$\begin{aligned}\langle\psi|(\mathcal{R}_{\rm op}\circ\mathcal{E})(P)|\psi\rangle &= \langle\psi_1|\sum_{i=1}^{d}\rho_i|\psi_1\rangle \\ &= \alpha_{11}^{(1)} + \sum_{i=2}^{d}\alpha_{11}^{(i)} \leq 1 + (d-1)\eta_{\rm op}.\end{aligned}$$

Putting this back into (2.15), and noting that $\left[\sum_i |\langle E_i^\dagger \mathcal{E}(P)^{-1/2} E_i\rangle|^2\right]^{1/2} \leq F(|\psi\rangle, \mathcal{R}_P\circ\mathcal{E})$ gives

$$F^2\left[|\psi\rangle, (\mathcal{R}_{\rm op}\circ\mathcal{E})(|\psi\rangle\langle\psi|\right] \leq \sqrt{1+(d-1)\eta_{\rm op}}F\left[|\psi\rangle, (\mathcal{R}_P\circ\mathcal{E})(|\psi\rangle\langle\psi|)\right], \tag{2.16}$$

which proves the theorem. ■

Let η_P denote the fidelity loss for code $\mathcal{C}$ with the transpose channel $\mathcal{R}_P$ as the recovery map. Then, Theorem 2.3.2 implies the following corollary.

Corollary 2.3.3. *η_P satisfies $\eta_{op} \leq \eta_P \leq \eta_{op} f(\eta_{op}; d)$, where $f(\eta; d)$ is the function*

$$f(\eta; d) \equiv \frac{(d+1)-\eta}{1+(d-1)\eta} = (d+1) + O(\eta). \tag{2.17}$$

Proof. That $\eta_P \geq \eta_{\rm op}$ is true by the definition of $\eta_{\rm op}$. To show that $\eta_P \leq \eta_{\rm op} f(\eta_{\rm op}; d)$, define for any $|\psi\rangle \in \mathcal{C}$, $\eta_{P,\psi}$ such that $F^2\left[|\psi\rangle, (\mathcal{R}_P\circ\mathcal{E})(|\psi\rangle\langle\psi|)\right] \equiv 1-\eta_{P,\psi}$. η_P is then just $\eta_P \equiv \max_\psi \eta_{P,\psi}$. From Theorem 2.3.2, we see that

$$\begin{aligned}1-\eta_{\rm op} &\leq F^2\left[|\psi\rangle, (\mathcal{R}_{\rm op}\circ\mathcal{E})\left(|\psi\rangle\langle\psi|\right)\right] \\ &\leq \sqrt{1+(d-1)\eta_{\rm op}}\; F\left[|\psi\rangle, (\mathcal{R}_P\circ\mathcal{E})(|\psi\rangle\langle\psi|)\right] \\ &= \sqrt{\left[1+(d-1)\eta_{\rm op}\right](1-\eta_{P,\psi})}.\end{aligned}$$

Rearranging gives $\eta_{P,\psi} \leq \eta_{\rm op} f(\eta_{\rm op}; d)$. Since this holds for all $\eta_{P,\psi}$, it also holds for η_P. ■

The inequality $\eta_P \leq \eta_{\rm op} f(\eta_{\rm op}; d)$ makes precise our statement that $\mathcal{R}_P$ is near-optimal. The recovery $\mathcal{R}_P$ works nearly as well as the optimal recovery, since its fidelity loss picks up at most an additional factor of $(d+1)$ (ignoring the $O(\eta)$ corrections). For the most practically relevant case of a code encoding a single qubit, this is a factor of 3 which is not too large. Observe also

that when $\eta_{\text{op}} = 0$, the inequality in Corollary 2.3.3 reduces to $\eta_P = \eta_{\text{op}}$, reaffirming that $\mathcal{R}_P$ is the optimal recovery in the case of perfect QEC.

We do not know if the upper bound on η_P in Corollary 2.3.3 is tight. However, the appearance of the dimension d of the code in the bound is unavoidable, as can be seen from the following example. Consider a noise channel $\mathcal{E} \sim \{E_i\}$ such that the action of $\mathcal{E}$ on a code $\mathcal{C}$ can be described by the set of Kraus operators $\{E_i P\} = \{\sqrt{1-p}\ P, \sqrt{p}\ |0\rangle\langle 0|, \sqrt{p}\ |0\rangle\langle 1|, \ldots, \sqrt{p}\ |0\rangle\langle d-1|\}$, for $0 \leq p \ll 1$. As usual, P is the projector onto $\mathcal{C}$ and d is the dimension of $\mathcal{C}$. $\mathcal{E}$ mostly acts like the identity channel on $\mathcal{C}$, but has a small component that maps a small part of every code state onto the state $|0\rangle$. For $d \geq 3$, one can show that the worst-case fidelity, when using the transpose channel as the recovery, occurs for the state $|0\rangle$. The corresponding fidelity loss is

$$\eta_P = \frac{(d-1)p}{1+(d-1)p}. \tag{2.18}$$

On the other hand, since $\mathcal{E}$ is nearly the identity channel, we can perhaps not do any recovery, i.e., make the identity channel the recovery map. In this case, we find that the fidelity loss is $\eta_0 \equiv p$ which is always smaller than η_P for small p. Since the optimal fidelity loss η_{op} must always be smaller than η_0, we have that $\eta_P/\eta_{\text{op}} \geq \eta_P/\eta_0 = (d-1)/[1+(d-1)p]$, which grows as d increases, for fixed p. Therefore, we see that there is an increasing separation between η_P and η_{op} as d increases.

In the next section, we will see that this approach to AQEC using the transpose channel can be viewed as a perturbation from the perfect QEC case. The factor of d appearing in our bounds can perhaps be understood as quantifying the number of degrees of freedom in which the approximate case can deviate from the perfect case. Note, however, that as d gets large, $f(\eta; d)$ approaches $1/\eta$. In this case, the inequality in Corollary 2.3.3 simply becomes the trivial statement $\eta_{\text{op}} \leq \eta_P \leq 1$. While we will often only be interested in codes with small values of d, this demonstrates the weakness in the bounds derived here for large values of d.

Finally, note that Corollary 2.3.3 provides a necessary and sufficient condition for $\mathcal{C}$ to be approximately correctable—$\mathcal{C}$ is approximately correctable if and only if η_P is small. In the next section, we will use this corollary to derive a set of AQEC conditions, much like those in Theorem 2.1.1 for perfect QEC.

2.4 The Transpose channel and QEC Conditions

One of the key tools in perfect QEC are the QEC conditions stated in Theorem 2.1.1. Similar conditions characterizing AQEC codes would be very useful. A natural approach to getting a set of AQEC conditions is to perturb the perfect QEC conditions to allow for small deviations. For example, the four-qubit code for the amplitude damping channel described in Section 2.2.1 was shown to obey a set of perturbed QEC conditions. More recent studies [90] have looked at small perturbations of the perfect QEC conditions for general CPTP channels. However, the analysis in [90] is complicated, and one wonders if there is a simpler approach using the transpose channel.

In this section, we prove a simple set of AQEC conditions based on Corollary 2.3.3. Drawing from our earlier observation that the transpose channel is the optimal recovery map for perfect QEC codes in Lemma 2.3.1, we rewrite the condition (2.3) for perfect QEC in such a way that the role of the transpose channel is apparent. From this, we derive a necessary and a sufficient condition for AQEC founded upon the transpose channel, as a natural generalization of the perfect QEC conditions. While AQEC conditions have been derived in the past from information-theoretic perspectives [16, 20, 69, 118], our conditions are algebraic, and lead to a simple and universal algorithm to find AQEC codes that does not require optimizing over all recovery maps for each encoding map.

2.4.1 Alternative Form of the Perfect QEC Conditions

The role of the transpose channel in perfect QEC becomes a lot more transparent once we realize that the QEC conditions in Theorem 2.1.1 can be written as follows.

Theorem 2.4.1 (**Alternative perfect QEC conditions).** *A code $\mathcal{C}$ satisfies the perfect QEC conditions of Theorem 2.1.1 if and only if it satisfies*

$$\forall i, j, \quad PE_i^\dagger \mathcal{E}(P)^{-1/2} E_j P = \beta_{ij} P, \tag{2.19}$$

where $\beta \equiv \sqrt{\alpha}$, for α from Theorem 2.1.1.

Proof. For a code $\mathcal{C}$ that satisfies the perfect QEC conditions, using (2.13) and $PU_k^\dagger U_l P = \delta_{kl} P$, we have

$$PF_k^\dagger \mathcal{E}(P)^{-1/2} F_l P = \sqrt{d_{ll}} P U_k^\dagger U_l P = \delta_{kl} \sqrt{d_{kk}} P. \tag{2.20}$$

This diagonal form can be rotated to any other Kraus representation by using the appropriate unitary u, such that $F_k = \sum_i u_{ik} E_i$ and $\alpha = udu^\dagger$. Then, defining $\beta \equiv \sqrt{\alpha}$, we get (2.19), thus showing that a code $\mathcal{C}$ satisfying the perfect QEC conditions, also satisfies (2.19).

Conversely, suppose we start with the "diagonal" form of (2.19) as in (2.20), which can be accomplished by choosing a unitary u so that β is diagonal with entries $\sqrt{d_{kk}}$. Since $\mathcal{E}$ is CP, $\mathcal{E}(P) \geq 0$ and hence $\mathcal{E}(P)^{-1/2} \geq 0$. Therefore, we can take square root of (2.20) and write $\mathcal{E}(P)^{-1/4} F_k P = (d_{kk})^{1/4} V_k P$, for some unitary V_k, which implies that

$$F_k P = (d_{kk})^{1/4} \, \mathcal{E}(P)^{1/4} V_k P. \tag{2.21}$$

Note that the inverse of $\mathcal{E}(P)$ is taken on its support, so that $\mathcal{E}(P)^{1/4}\mathcal{E}(P)^{-1/4} = P_{\mathcal{E}}$. Putting (2.21) back into (2.20) then gives $PV_k^\dagger V_l P = \delta_{kl} P$. Furthermore,

$$\mathcal{E}(P) = \sum_k (F_k P)(P F_k^\dagger) = \mathcal{E}(P)^{1/4} \Big(\sum_k \sqrt{d_{kk}} V_k P V_k^\dagger \Big) \mathcal{E}(P)^{1/4},$$

which implies $\mathcal{E}(P)^{1/2} = \sum_k \sqrt{d_{kk}} V_k P V_k^\dagger$. A simple calculation now shows $P F_k^\dagger F_l P = \delta_{kl} d_{kk} P$, which is exactly the diagonal form of the perfect QEC conditions (2.4). Applying an appropriate u to rotate to the desired Kraus representation gives (2.3). ■

It may be observed that the left-hand side of (2.19) is simply a Kraus operator of the map $\mathcal{R}_P \circ \mathcal{E}$. In other words, the QEC conditions given in Theorem 2.4.1, as also the original version given in Theorem 2.1.1, simply express the fact that $\mathcal{C}$ is perfectly correctable if and only if $\mathcal{R}_P \circ \mathcal{E} \propto \hat{\mathcal{P}}$, where $\hat{\mathcal{P}}$ is the projection $P(\cdot)P$, which acts trivially on the code $\mathcal{C}$. The proportionality factor is $\sum_{ij} \beta_{ij}^2 = \sum_{ij} \alpha_{ij} = \sum_k d_{kk}$.

2.4.2 AQEC Conditions

We now obtain conditions for AQEC by perturbing the alternative form (2.19) of the perfect QEC conditions. The perturbation is added as a small operator on the right-hand side of (2.19) for each i, j, but in order to make a precise statement, we also need to relate the size of these perturbations to how well the given code can be corrected. This is not difficult since we have already characterized the performance of the transpose channel as a recovery map in Theorem 2.3.2, or equivalently in

Corollary 2.3.3.

Theorem 2.4.2 (AQEC conditions). *Suppose we have a CPTP channel $\mathcal{E} \sim \{E_i\}$, and a d-dimensional subspace code $\mathcal{C}$ with projector P. Let $\Delta_{ij} \in \mathcal{B}(\mathcal{C})$ be traceless operators such that*

$$PE_i^\dagger \mathcal{E}(P)^{-1/2} E_j P = \beta_{ij} P + \Delta_{ij}, \tag{2.22}$$

where $\beta_{ij} \in \mathbb{C}$. Then, for $\epsilon \in [0,1]$, there exists $\eta \in [0,1]$ such that

(i) $\mathcal{C}$ is ϵ-correctable if $\eta \leq \epsilon$;

(ii) $\mathcal{C}$ is ϵ-correctable only if $\eta \leq \epsilon f(\epsilon; d)$, *where f is the function*

$$f(\epsilon; d) \equiv \frac{(d+1) - \epsilon}{1 + (d-1)\epsilon} = (d+1) + O(\epsilon). \tag{2.23}$$

Proof. The left-hand side of (2.22) are Kraus operators of $\mathcal{R}_P \circ \mathcal{E}$. This, along with the fact that $\mathcal{R}_P \circ \mathcal{E}$ is trace-preserving, implies that for a noise channel $\mathcal{E}$ satisfying (2.22), the fidelity under the transpose channel recovery is given by

$$F^2 \left[|\psi\rangle, (\mathcal{R}_P \circ \mathcal{E})(|\psi\rangle\langle\psi|) \right] = 1 - \sum_{ij} \left[\langle\psi|\Delta_{ij}^\dagger \Delta_{ij}|\psi\rangle - |\langle\psi|\Delta_{ij}|\psi\rangle|^2 \right]. \tag{2.24}$$

Recalling the definition of the fidelity loss η_P, for the transpose channel, we get

$$\eta_P = \max_{|\psi\rangle \in \mathcal{C}} \sum_{ij} \left[\langle\psi|\Delta_{ij}^\dagger \Delta_{ij}|\psi\rangle - |\langle\psi|\Delta_{ij}|\psi\rangle|^2 \right]. \tag{2.25}$$

Setting $\eta = \eta_P$, conditions (i) and (ii) follow directly from Corollary 2.3.3. ■

It is clear that the expression for η_P is a non-negative quantity, since the fidelity in (2.24) is bounded by 1. Furthermore, (2.25) elucidates how the fidelity loss arises from the presence of the Δ_{ij} operators. If $\Delta_{ij} = 0 \ \forall i, j$, we have perfect QEC.

The AQEC conditions, like the perfect QEC conditions, provide a way to check if a code is approximately correctable, without requiring knowledge of the optimal recovery. More precisely, given a maximum tolerable fidelity loss ϵ for some information processing task at hand, one can check if a code $\mathcal{C}$ is ϵ-correctable, as follows. We first compute η_P, which can be done once $\mathcal{C}$ and

$\mathcal{E}$ are known. If $\eta_P \leq \epsilon$, then $\mathcal{C}$ is a good code. If however, η_P violates the inequality in Condition (ii), we know that $\mathcal{C}$ is not good enough for our purposes. Of course, there is a gap—for η_P taking values $\epsilon \leq \eta_P \leq \epsilon f(\epsilon; d)$, we cannot use these conditions to determine whether $\mathcal{C}$ is within our tolerable fidelity loss, but this gap is small for small d. We do not know if the gap can be shrunk by replacing η_P with the fidelity loss for a recovery map other than the transpose channel, but we believe it is unlikely to vanish completely.

For a general $\mathcal{C}$, the fidelity loss η_P may be difficult to compute as it requires a maximization over all states in the code space. However, there is a quick way to check for sufficiency by relaxing condition (i) of Theorem 2.4.2 slightly.

Corollary 2.4.3. *$\mathcal{C}$ is ϵ-correctable for some $\epsilon \in [0, 1]$ if*

$$\| \Delta_{sum} \| \leq \epsilon, \tag{2.26}$$

where $\Delta_{sum} \equiv \sum_{ij} \Delta_{ij}^\dagger \Delta_{ij}$ and $\| \cdot \|$ denotes the operator norm.

Proof. Observe that $\sum_{ij} [\langle\psi|\Delta_{ij}^\dagger \Delta_{ij}|\psi\rangle - |\langle\psi|\Delta_{ij}^\dagger|\psi\rangle|^2] \leq \sum_{ij} \langle\psi|\Delta_{ij}^\dagger \Delta_{ij}|\psi\rangle = \langle\psi|\Delta_{\text{sum}}|\psi\rangle$. From the definition of the operator norm, it is easy to see that $\max_{|\psi\rangle \in \mathcal{C}} \langle\psi|\Delta_{\text{sum}}|\psi\rangle = \|\Delta_{\text{sum}}\|$. Hence, $\eta_P \leq \|\Delta_{\text{sum}}\|$, and the condition $\eta_P \leq \epsilon$ in statement (i) of the AQEC conditions (Corollary 2.3.2) is certainly satisfied if $\|\Delta_{\text{sum}}\| \leq \epsilon$. ■

This sufficiency condition (2.26) and the AQEC conditions of Theorem 2.4.2 form the basis of a simple algorithm to find good AQEC codes, presented in the next section. Since Δ_{sum} is a positive semi-definite operator, its operator norm is given by its maximum eigenvalue, which is easily computable. In fact, for codes encoding a single qubit, we show in Appendix A that $\|\Delta_{\text{sum}}\| = 1 - \sum_{ij} |\beta_{ij}|^2$. Note that for a given code $\mathcal{C}$ and noise channel $\mathcal{E}$, β_{ij} is easily computed, since $\beta_{ij} = (1/d) \operatorname{tr}[P E_i^\dagger \mathcal{E}(P)^{-1/2} E_j P]$. Furthermore, we also show in Appendix A that for the case of qubit codes, η_P can be computed easily with simple eigenanalysis. In fact our method of computing the worst-case fidelity for a CPTP qubit map described in Section A.2 might be useful in contexts beyond our present discussion.

2.5 Finding AQEC Codes

Consider the practical problem of finding a d-dimensional code, given some maximum tolerable fidelity loss ϵ, such that, every code state must have fidelity $F \geq \sqrt{1-\epsilon}$, after passing through the noise channel and recovery map. The following algorithm provides a simple procedure to search for such ϵ-correctable codes, for a given noise channel and system Hilbert space.

Algorithm

Step 1. Pick a d-dimensional subspace $\mathcal{C} \subseteq \mathcal{H}$. This can be done, for example, by randomly picking d linearly independent vectors from $\mathcal{H}$ and defining $\mathcal{C}$ as their linear span.

Step 2. Compute $\forall\ i, j$,

$$\Delta_{ij} \equiv P E_i^\dagger \mathcal{E}(P)^{-1/2} E_j P - \beta_{ij} P, \tag{2.27a}$$

$$\beta_{ij} \equiv \frac{1}{d}\ \mathrm{tr}(P E_i^\dagger \mathcal{E}(P)^{-1/2} E_j P). \tag{2.27b}$$

Find the maximum eigenvalue $\lambda_{\max}$ of $\Delta_{\mathrm{sum}} \equiv \sum_{ij} \Delta_{ij}^\dagger \Delta_{ij}$. If $\lambda_{\max} \leq \epsilon$, then we are done, since $\mathcal{C}$ is an ϵ-correctable code.

Step 3. If not, compute the fidelity loss η_P for the recovery map $\mathcal{R}_P$, as given in (2.25). If $\eta_P \leq \epsilon$, then again $\mathcal{C}$ is an ϵ-correctable code.

Step 4. If not, check if $\eta_P > \epsilon f(\epsilon; d)$. If true, $\mathcal{C}$ is *not* ϵ-correctable. We return to Step 1 and try again with a different $\mathcal{C}$.

Step 5. If $\epsilon < \eta_P \leq \epsilon f(\epsilon; d)$, we do not know if $\mathcal{C}$ is ϵ-correctable, but we can still choose to discard this $\mathcal{C}$ and return to Step 1 to try again with a different $\mathcal{C}$.

If this algorithm finds a code that works well enough, one can then try to optimize performance by looking for the optimal recovery map. While looking for this optimal recovery can be a difficult process that requires exhaustive search, with our algorithm we only need to do this possibly expensive computation *once*, for the code generated by our algorithm is guaranteed to be ϵ-correctable. Otherwise, one can always use the transpose channel itself as a good recovery.

There is of course the possibility that the algorithm yields no code within our fidelity loss requirements. This does not immediately imply that $\mathcal{H}$ does not contain an ϵ-correctable code, because of the presence of the gap stated in Step 5. However, figuring out whether any of the codes that fall in this gap is a good enough code is the same as finding the optimal recovery map for that code, a problem which we currently do not know how to solve efficiently.

2.6 Example: Amplitude Damping Channel

In this section we compare the performance of the transpose channel with that of other AQEC schemes, for the case of amplitude damping noise. The single-qubit amplitude damping channel $\mathcal{E}_{\mathrm{AD}}$ is the CPTP channel described in (2.2), parameterized by the damping parameter γ. Recall that γ corresponds to the probability of a transition from the excited state to the ground state. In Fig. 2.5 we plot the worst-case fidelity for different AQEC codes as function of γ.

Clearly, in the absence of any encoding or recovery, the worst-case fidelity for a single qubit undergoing $\mathcal{E}_{\mathrm{AD}}$ decreases as $1-\gamma$ (see Fig. 2.5, line labeled "no error correction"). The $[4,1]$ code due to Leung et al described in Section 2.2.1, in combination with the Leung recovery, increases the fidelity significantly as compared to the no error correction case. In the same figure, we have also plotted the worst-case fidelity using the transpose channel $\mathcal{R}_P$ as the recovery operation instead of the Leung recovery, for the same [4,1] code. From the plot, we can see that using the transpose channel as the recovery map gives a higher fidelity than the original Leung recovery.

For comparison, we have also looked at a recovery map for the $[4,1]$ code constructed by Fletcher et al. in [47]. Their recovery, which we refer to as the *Fletcher recovery*, was originally optimized for an averaged measure of fidelity. We have instead computed the worst-case fidelity for this recovery, [2] and this is plotted in Fig. 2.5. For small values of γ, the Fletcher recovery gives the best performance compared to the other recovery maps, despite being optimized for an averaged measure of fidelity. However, it is only marginally better than the transpose channel recovery.

We have also compared the performance of the $[4,1]$ approximate code under these different recovery maps with that of the smallest known perfect code, namely the $[[5,1,3]]$ code [15, 84].

[2] The recovery map we use here is from Table I of [47]. Their recovery map actually depends on two parameters α and β which can be numerically optimized, for each value of γ, for the best recovery map. For simplicity, we set $\alpha = \beta = 1/\sqrt{2}$ in our plot, which corresponds to the "code-projected recovery" in [47] with comparable performance as the fully optimized recovery.

Using the corresponding $\mathcal{R}_{\text{perf}}$ as the recovery for the [[5,1,3]] code, we have computed the worst-case fidelity for different values of γ. As the plot in Fig. 2.5 shows, the [[5,1,3]] code performs better than the [4,1] code with Leung recovery, but the [4,1] code uses one fewer qubit to encode the same amount of information. The [4,1] code with the transpose channel as recovery has nearly identical worst-case fidelity as the [[5,1,3]] code, while the one with Fletcher recovery does slightly better than the [[5,1,3]] code for small values of γ.

These observations clearly demonstrate the benefit of going beyond the codes described by the perfect QEC conditions. Furthermore, while the [[5,1,3]] code is capable of perfectly correcting any single qubit error on a system subjected to *any* noise channel, the comparison with the [4,1] code with its various recovery maps clearly show the gain that one might achieve by adapting the codes and recovery to the noise channel in question.

Finally, we have also randomly generated codes that encode a single qubit into four physical qubits, for the amplitude damping channel. We computed the worst-case fidelity for each code using the transpose channel as the recovery map. We tried about 500 randomly selected codes, taking less than half an hour on a typical laptop computer. The worst-case fidelity for the best code we found is given in Fig. 2.5 (line marked "random 4-qubit code, $\mathcal{R}_P$ recovery"). For small values of γ, this random code does not do as well as the other codes discussed so far for the amplitude damping channel, but it still does significantly better than the case without error correction. Furthermore, for $\gamma \gtrsim 0.35$, our randomly generated code actually outperforms all the other codes. For comparison, we have also plotted the worst-case fidelity for this randomly generated code in the absence of the transpose channel recovery, i.e., with the identity channel as the recovery map (line marked "random 4-qubit code, Id recovery"). In all this, one should keep in mind the ease with which the performance of the randomly generated code was achieved, due to the fact that the transpose channel is a near-optimal recovery map for *any* code.

One can even consider the possibility of looking for two- and three-qubit codes. We randomly generated codes encoding a single qubit of information into two and three physical qubits for the amplitude damping channel. Because the transpose channel is near-optimal for any code, it can be used a good recovery map for the codes we generate, thus eliminating the need to search for a good recovery for every randomly selected code. The worst-case fidelity for the best codes we found are plotted in Fig. 2.6. For comparison, we have also plotted the worst-case fidelities for the

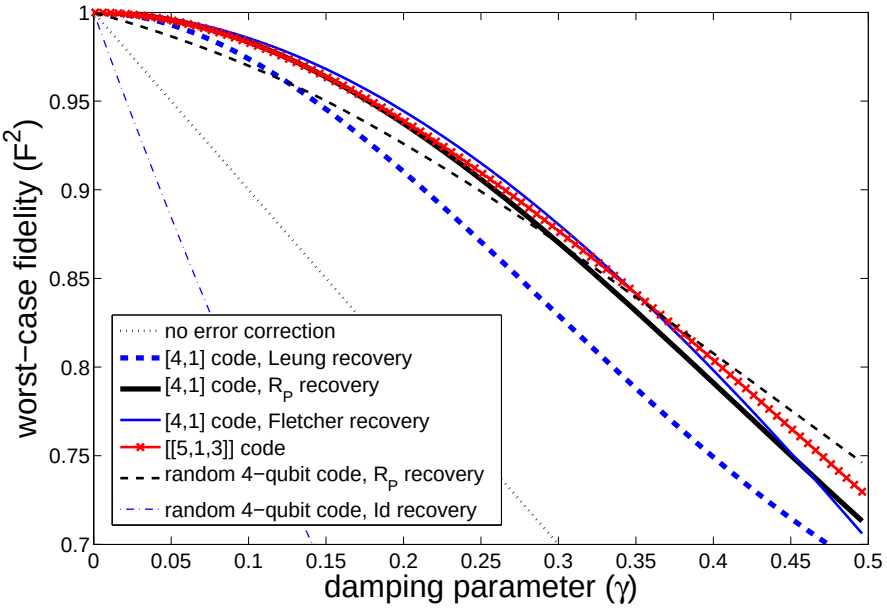


Figure 2.5: Codes for the amplitude damping channel, for $0 \leq \gamma \leq 0.5$.

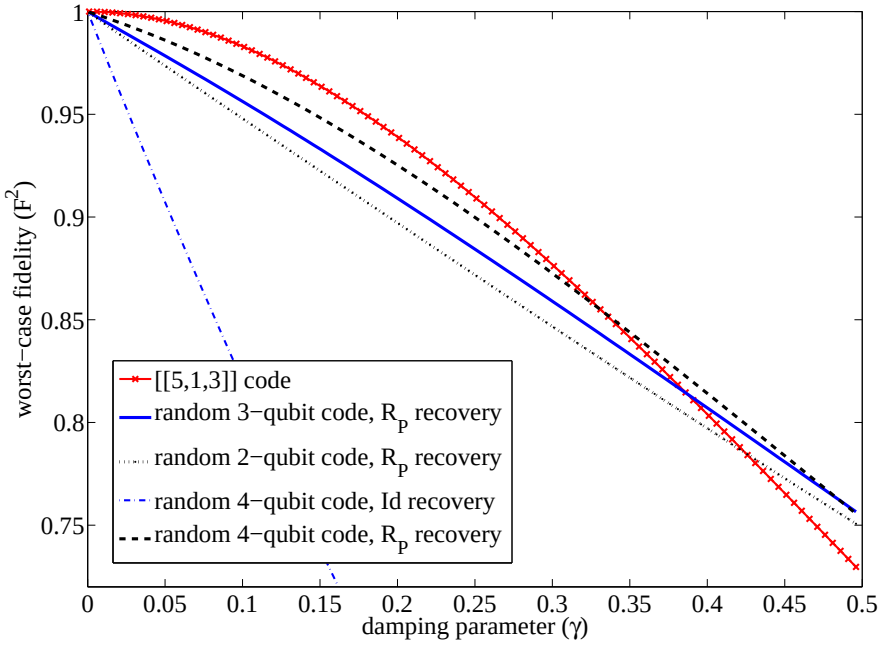


Figure 2.6: Randomly generated two-, three-, and four-qubit codes using the transpose channel as the recovery map. For comparison, we have also plotted the worst-case fidelity for the [[5,1,3]] code, and that of the randomly generated four-qubit code with no recovery (i.e., identity channel as recovery).

randomly generated four-qubit code mentioned in the previous paragraph. From the figure, we see that while the worst-case fidelity decreases as the number of physical qubits decreases, the two- and three-qubit codes in fact do not perform too badly compared to the four-qubit code or the [[5,1,3]] code. Such codes may be of relevance whenever the desire to lower resource requirements trumps the need for the best possible worst-case fidelity.

2.7 Conclusions and Open Problems

In this chapter, we have demonstrated the crucial role the transpose channel plays in perfect QEC and used it to formulate a simple approach to characterize and find AQEC codes. Compared to previous work based on numerically generated recovery maps specific to the noise channel in question, the universal and analytically simple form of the transpose channel makes it particularly useful for developing a better understanding of AQEC. Further, the near-optimality of the transpose channel leads to a simple algorithm for identifying codes that satisfy some maximum fidelity loss requirements, without having to perform a difficult optimization over all recovery maps for every possible encoding. Our approach, founded upon the worst-case fidelity rather than an averaged measure of fidelity, also ensures that the code found is able to protect *all* information that can be stored in the code with some minimum fidelity.

There are many interesting open problems. An immediate question is whether the gap present in our AQEC conditions between the necessary and sufficient conditions (arising from the inequality in Corollary 2.3.3) can be reduced, either by improving the bound in Theorem 2.3.2, or by using a different recovery map that might perform better than the transpose channel. It would also be very interesting to find a similarly simple and universal recovery map, such that the dimension of the code does not appear in the worst-case fidelity. It might also be of interest to extend our efficient method of computing the worst-case fidelity to higher-dimensional codes and more general channels. Furthermore, we expect that the transpose channel can also be used to study approximate codes more general than subspace codes, for example, OQEC codes which also admit a description based on conditions like the QEC conditions [82].

Another important problem is to figure out whether the transpose channel can be easily implemented using measurements and gates. In the case of perfect QEC, the transpose channel (or equivalently $\mathcal{R}_{\text{perf}}$) can be implemented simply using syndrome measurements and conditional gates

(see for example, [98, Section 10.3]). In order for AQEC codes to be useful for computational or communication tasks, it must be possible to implement the recovery operation using physical operations that are not overly complicated or demanding in resources. That this is possible in the perfect QEC case could offer some clues to implementing the transpose channel for AQEC codes.

AQEC provides a new and mostly unexplored arena of possibilities for the design of codes to protect information from noise in quantum information processing tasks. Our work provides an analytical characterization of AQEC and further analytical understanding will undoubtedly prove invaluable toward unlocking the full potential of AQEC.

Chapter 3

Symmetric Complementary Aspects and Entropic Uncertainty Relations

Entropic uncertainty relations provide a natural way to quantify incompatibility between multiple measurements, by lower bounding the average entropy of the probability distributions corresponding to the outcomes of different measurements. For two observables, it is well known that this incompatibility is maximizum when the measurement bases are *complementary* or *mutually unbiased*. For more than two measurement settings, being mutually unbiased is a necessary condition to obtain strong uncertainty relations, but not sufficient. It remains an important open question to identify and construct sets of complementary bases satisfying strong uncertainty relations.

Uncertainty relations figure prominently in the analysis of quantum cryptographic protocols such as quantum key distribution [72, 106], the phenomenon of information locking [38], and in characterizing entanglement and separability [53]. In particular, the security of recent cryptographic models like the bounded-storage and noisy-storage models [30, 31, 76, 115, 128] is derived on the basis of an entropic uncertainty relation. A better understanding of the interplay between complementarity and uncertainty relations is thus of interest not only from a foundational standpoint; it has practical implications for analyzing and improving existing cryptographic protocols.

Here,[1] we construct special sets of up to $2n + 1$ mutually unbiased bases (MUBs) in dimension $d = 2^n$ which have particularly nice symmetry properties derived from the Clifford algebra. More precisely, we show that there exists a unitary transformation that cyclically permutes such bases.

[1]The work described in this chapter was done in collaboration with Stephanie Wehner. The original results presented here have been published in [92]. The proofs presented in Appendix B are based on several useful discussions with Niranjan Balachandran. We are grateful to David Gross for pointing us to the relevant literature for the discrete phase space construction. We also thank Lukasz Fidkowski and John Preskill for interesting discussions.

This unitary can be understood as a generalization of the Fourier transform, which exchanges two MUBs, to multiple complementary aspects. We then obtain a lower bound for min-entropic uncertainty relations for any set of MUBs, and show that symmetry plays a central role in obtaining tight bounds. For example, we obtain for the first time a tight bound for four MUBs in dimension $d = 4$, which is attained by an eigenstate of our complementarity transform. Finally, we discuss our work in relation to other symmetries obtained by transformations in discrete phase space, and note that the extrema of discrete Wigner functions [50] are directly related to min-entropic uncertainty relations for MUBs.

This chapter is organized as follows. We begin with a formal introduction to the different measures of entropy used in this chapter, in Section 3.1.1. In Section 3.1.2 we summarize some of the properties of Clifford algebras which are used in our construction of MUBs. Section 3.2 contains an overview of entropic uncertainty relations (EURs) and mutually unbiased bases (MUBs). In Section 3.3 we focus on the problem of obtaining uncertainty relations for the min-entropy, and also discuss the related problem of finding the extrema of the discrete Wigner function (Section 3.3.2). In Section 3.4 we describe our new lower bound for the min-entropy of *any* set of MUBs. Section 3.5 describes our construction of *symmetric* MUBs in dimension $d = 2^n$ using the generators of the Clifford algebra. Finally, in Section 3.6 we discuss the role of symmetry in obtaining tight lower bounds, with examples in dimensions $d = 4$ and $d = 8$.

3.1 Preliminaries

3.1.1 Measures of Entropy

We first provide a short introduction to the entropic measures used in this chapter. Let $\mathcal{M} = \{ M^b \,|\, M^b \in \mathcal{B}(\mathcal{H}) \}_{b=1}^{d}$ be a measurement in the d-dimensional Hilbert space $\mathcal{H}$, with a finite set of outcomes labeled by b, where $M^b > 0$, and $\sum_b M^b = \mathbb{I}$. For any quantum state ρ, the measurement $\mathcal{M}_j$ induces a probability distribution P_j over the outcomes $P_j(b) = \mathrm{tr}(M_j^b \rho)$. The *Rényi entropy* [110] of order α ($\alpha > 0$, $\alpha \neq 1$), of the probability distribution obtained by measuring

$\mathcal{M}$ on a state $\rho \in \mathcal{S}(\mathcal{H})$, denoted by $H_\alpha(\mathcal{M}_j|\rho)$, is given by

$$H_\alpha(\mathcal{M}|\rho) = \frac{1}{1-\alpha} \log \left[\left(\sum_{b=1}^{d} (\mathrm{tr}[M^b \rho])^\alpha \right)^{\frac{1}{\alpha-1}} \right]. \tag{3.1}$$

The Shannon entropy [119] forms a special case of the Rényi entropy, which is obtained by taking the limit $\alpha \to 1$, that is,

$$H_1(\mathcal{M}|\rho) = \lim_{\alpha \to 1} H_\alpha(\mathcal{M}|\rho) = -\sum_{b=1}^{d} \mathrm{tr}[M^b \rho] \log \mathrm{tr}[M^b \rho]. \tag{3.2}$$

The Shannon entropy is usually written as $H(\cdot)$, omitting the subscript. Other special cases of importance are

(a) the *min-entropy*, when $\alpha \to \infty$

$$H_\infty(\mathcal{M}|\rho) = -\log \left(\max_b \mathrm{tr}[M^b \rho] \right) , \tag{3.3}$$

and

(b) the *collision entropy*, when $\alpha = 2$

$$H_2(\mathcal{M}|\rho) = -\log \sum_{b=1}^{d} (\mathrm{tr}[M^b \rho])^2 .$$

The Rényi entropies are monotonically decreasing in α, that is,

$$H_0(.) = \log d \geq H(\cdot) \geq H_2(\cdot) \geq H_\infty(\cdot) \geq 0 .$$

A lower bound on H_α thus provides us with a bound on H_β as well, whenever $\alpha \geq \beta$.

We have defined these entropic quantities here for a general measurement. In this chapter, however, we are mainly interested in the case where $\mathcal{M}$ is a basis for $\mathcal{H}$ and the operators $\{M^b\}$ are rank-1 projectors of the form $M^b = |b\rangle\langle b|$.

3.1.2 Clifford Algebras

We now provide a brief introduction to Clifford algebras and their properties, which we make use of in our construction of MUBs. For any integer n, the Clifford algebra of dimension $d = 2^n$ is the real, associative algebra generated by operators $\Gamma_0, \ldots, \Gamma_{2n-1}$ satisfying the anticommutation relations

$$\{\Gamma_i, \Gamma_j\} = 2\delta_{ij}, \text{ for } i \neq j.$$

This Clifford algebra has a unique representation by Hermitian matrices on n qubits (up to unitary equivalence) that can be obtained via the famous Jordan-Wigner transformation [62]:

$$\begin{aligned}\Gamma_{2j+1} &= Y^{\otimes(j-1)} \otimes Z \otimes \mathbb{I}^{\otimes(n-j)},\\ \Gamma_{2j} &= Y^{\otimes(j-1)} \otimes X \otimes \mathbb{I}^{\otimes(n-j)},\end{aligned}$$

for $j = 0, \ldots, n-1$, where X, Y and Z to denote the Pauli matrices. Furthermore, we define

$$\Gamma_{2n} := i\Gamma_0 \ldots \Gamma_{2n-1} \ .$$

Note that in dimension $d = 2$, these are just the familiar Pauli matrices, $\Gamma_0 = X$, $\Gamma_1 = Z$ and $\Gamma_2 = Y$.

Of particular importance to us will be the fact that we can view the operators $\Gamma_0, \ldots, \Gamma_{2n-1}$, as $2n$ orthogonal vectors forming a basis for $\mathbb{R}^{2n}$. In particular, for any orthonormal transformation $T \in \mathrm{O}(2n)$ which when applied to the vector $v = (v^{(0)}, \ldots, v^{(2n-1)}) \in \mathbb{R}^{2n}$ gives $\tilde{v} = (\tilde{v}^{(1)}, \ldots, \tilde{v}^{(2n-1)}) = T(v)$, there exists a unitary $U(T) \in \mathcal{B}(\mathcal{H})$ acting on the underlying Hilbert space $\mathcal{H}$ of dimension $d = 2^n$ [88], such that,

$$U(T) \left(\sum_j v_j \Gamma_j \right) U(T)^\dagger = \sum_j \tilde{v}_j \Gamma_j \ .$$

We refer to [127, Appendix C] for a description of how to obtain explicit constructions of $U(T)$.

The orthonormal transformation we are interested in is the one that cyclically permutes the basis vectors. As described above we can find a corresponding unitary $U = U(T)$ which cyclically permutes the basis vectors $\Gamma_0, \Gamma_2, \ldots, \Gamma_{L-1}$. This transformation of the form $U(T)\Gamma_j U(T)^\dagger \to \Gamma_k$,

is particularly simple to obtain. It can be built up from successive rotations in the plane spanned by only two "vectors" Γ_j and Γ_k. We first construct a unitary that corresponds to a rotation around an angle $\pi/2$ in the plane spanned by Γ_j and Γ_k, bringing Γ_j to Γ_k. This is simply a reflection around the plane orthogonal to the midvector between Γ_j and Γ_k, followed by a reflection around the plane orthogonal to Γ_k. Using the geometric properties of the Clifford algebra, this can be shown to correspond to the unitary

$$R_{j\to k} = \Gamma_k(\Gamma_j + \Gamma_k)/\sqrt{2} \ .$$

To obtain the desired unitary, we now compose a number of such rotations. Let $\hat{R}_{j,k} = R_{j\to k}$ if k is odd, and $\hat{R}_{j,k} = R_{k\to j}$ if k is even. Furthermore, define the operator $F = \mathbb{I}$ if L is odd, and $F = \Gamma_{2n}\Gamma_{L-1}$ if L is even. Note that $\Gamma_{2n}\Gamma_{L-1}$ is the unitary that flips the sign of Γ_{L-1}, but leaves all Γ_j for $j \neq 2n$ and $j \neq (L-1)$ invariant. We may then write

$$U(T) = F\hat{R}_{0,1}\hat{R}_{0,2}\ldots\hat{R}_{0,L-1} \ .$$

This unitary thus transforms $\Gamma_0 \to \Gamma_1 \to \ldots \to \Gamma_{L-1} \to \Gamma_0$, but leaves all other generators Γ_j for $j \geq L$ invariant. A similar unitary can be found for any transformation $T \in \mathrm{SO}(2n+1)$ [129], but is more difficult to construct explicitly.

Finally, we will also make use of the property that the set of d^2 operators, consisting of the Clifford generators and their products,

$$\mathcal{S} = \{\mathbb{I}, \Gamma_j, i\Gamma_i\Gamma_j, \Gamma_i\Gamma_j\Gamma_k, \ldots, i\Gamma_1\ldots\Gamma_{2n}\} \ , \tag{3.4}$$

forms an orthogonal basis [2] for $d \times d$ Hermitian matrices in $d = 2^n$ [36].

3.2 EURs and MUBs: An Overview

The uncertainty relation, first proposed by Heisenberg [57] for two conjugate observables, is one of the central principles of quantum mechanics. Indeed, it forms one of the most significant features

[2] Orthogonal with respect to the Hilbert-Schmidt inner product, which is given by $\mathrm{tr}[A^\dagger B]$, for a pair of operators A and B.

of quantum theory showing that the quantum world does differ fundamentally from the classical world. Uncertainty relations today are probably best known in the form given by Robertson [111], who extended Heisenberg's result to two arbitrary observables A and B. Robertson's relation states that if we prepare many copies of the state $|\psi\rangle$, and measure each copy individually using either A or B, we have

$$\Delta A \Delta B \geq \frac{1}{2}|\langle\psi|[A,B]|\psi\rangle|, \tag{3.5}$$

where $\Delta X = \sqrt{\langle\psi|X^2|\psi\rangle - \langle\psi|X|\psi\rangle^2}$ for $X \in \{A, B\}$ is the standard deviation resulting from measuring observable X on $|\psi\rangle$. The essence of (3.5) is that quantum mechanics does not allow us to simultaneously specify definite outcomes for two non-commuting observables when measuring the same state. The largest possible lower bound in Robertson's inequality (3.5) is $1/2$, which happens if and only if A and B are related by a Fourier transform, that is, they are conjugate observables.

However, nature typically allows us to perform more than two measurements on any given system, leading to the question of how we can determine "incompatibility" between multiple measurements. Clearly, due to its use of the commutator relation, the lower bound of (3.5) most directly relates to the case of *two* measurements. Is there a natural way of quantifying uncertainty for multiple measurements? And if so, what measurements might be most "incompatible"?

3.2.1 Entropic Uncertainty Relations

A natural measure that captures relations among probability distributions over the outcomes of different measurements is the entropy of such distributions. The first *entropic uncertainty relation* was proposed by Hirschmann [58] for position and momentum observables. This relation was later improved by Beckner [10] and Bialynicki-Birula and Mycielski [18], where the latter show, for n canonical pairs of position and momentum coordinates X_i and P_i,

$$H(X_1 \ldots X_n|\rho) + H(P_1 \ldots P_n|\rho) \geq n \log(e\pi),$$

where $H(Q_1 \ldots Q_n|\rho)$ is the differential Shannon entropy of the joint distribution of the coordinates $Q_1, \ldots, Q_n$, when measured on the state ρ. They further show that Heisenberg's uncertainty

relation (3.5) is in fact implied by this entropic uncertainty relation, thereby showing that using entropic quantities might provide a more general way of quantifying uncertainty.

In recent times, the study of entropic uncertainty relations has gained impetus from the work of Deutsch [34], who argued that entropy is a more desirable measure to quantify "uncertainty" than the standard deviation, for the following reason. The lower bound in (3.5) is trivial when $|\psi\rangle$ happens to give zero expectation on $[A, B]$. Hence, it would be useful to have a way of measuring "incompatibility" which depends only on the measurements A and B and not on the state.

Definition 3.2.1 (**Entropic Uncertainty Relations**). *For a set of L measurements $\{\mathcal{M}_0, \ldots, \mathcal{M}_{L-1}\}$, an entropic uncertainty relation is a lower bound of the form*

$$\frac{1}{L}\sum_{j=0}^{L-1} H_\alpha(\mathcal{M}_j|\rho) \geq c_{\alpha,\{\mathcal{M}_j\}} \ , \ \forall \rho \in \mathcal{S}(\mathcal{H}), \tag{3.6}$$

where $c_{\alpha,\{\mathcal{M}_j\}}$ is a constant that depends only on the choice of measurements $\{\mathcal{M}_j\}$ and choice of the entropy function (α), but is independent of the choice of state ρ in $\mathcal{H}$.

We call the state ρ that minimizes the average sum of entropies a *maximally certain state*. When H_α is the Shannon entropy $\mathcal{H}$, the largest bound we can hope to obtain for any choice of L measurements in a d-dimensional Hilbert space is

$$c_{1,L} = \frac{L-1}{L} \ \log d \ , \tag{3.7}$$

which is attained when the state ρ is an eigenstate of one of the measurements. If (3.7) is indeed a lower bound to (3.6), we will call the measurements *maximally incompatible with respect to the Shannon entropy.*

Deutsch [34] himself showed that

$$\frac{1}{2}\left(H_\infty(\mathcal{A}||\psi\rangle) + H_\infty(\mathcal{B}||\psi\rangle)\right) \geq -\log\left(\frac{1+c(\mathcal{A},\mathcal{B})}{2}\right), \tag{3.8}$$

where $c(\mathcal{A},\mathcal{B}) := \max\{|\langle a|b\rangle| \mid |a\rangle \in \mathcal{A}, |b\rangle \in \mathcal{B}\}$, and $H_\infty(\mathcal{A}||\psi\rangle)$ is the min-entropy arising from measuring the pure state $|\psi\rangle$ using the basis $\mathcal{A}$. If $\mathcal{A}$ and $\mathcal{B}$ are related by a Fourier transform, then (3.8) becomes

$$\frac{1}{2}\left(H_\infty(\mathcal{A}||\psi\rangle) + H_\infty(\mathcal{B}||\psi\rangle)\right) \geq -\log\left(\frac{1}{2} + \frac{1}{2\sqrt{d}}\right), \tag{3.9}$$

and this minimum value is achieved by a state that is invariant under the Fourier transform. Since the Shannon entropy obeys $H(\cdot) \geq H_\infty(\cdot)$, Deutsch's bound also holds for the Shannon entropy.

Better lower bounds have since been obtained for the Shannon entropy by Maassen and Uffink [89] following a conjecture of Kraus [81]. The Maassen-Uffink bound states that for any two orthonormal bases $\mathcal{A} \equiv \{|a_1\rangle, \ldots, |a_d\rangle\}$ and $\mathcal{B} \equiv \{|b_1\rangle, \ldots, |b_d\rangle\}$, in a d-dimensional space $\mathcal{H}$,

$$\frac{1}{2}\left(H(\mathcal{A}||\psi\rangle) + H(\mathcal{B}||\psi\rangle)\right) \geq -\log c(\mathcal{A},\mathcal{B}), \quad \forall |\psi\rangle \in \mathcal{H}, \tag{3.10}$$

where $c(\mathcal{A},\mathcal{B}) := \max_{a,b} |\langle a|b\rangle|$. The lower bound can at most take a value of $\frac{1}{2}\log d$, which implies

$$\frac{1}{2}\left(H(\mathcal{A}||\psi\rangle) + H(\mathcal{B}||\psi\rangle)\right) \geq \frac{1}{2}\log d. \tag{3.11}$$

This lower bound is attained if for all basis vectors $|a\rangle$ of basis $\mathcal{A}$ and all vectors $|b\rangle$ of basis $\mathcal{B}$,

$$|\langle a|b\rangle|^2 = \frac{1}{d}\,. \tag{3.12}$$

Any two bases satisfying this property are called *mutually unbiased bases*, or *complementary aspects*, and the unitary that exchanges two mutually unbiased bases can be understood as a Fourier transform. While the Maassen-Uffink bound is not tight for any two observables in general, it is indeed tight for two mutually unbiased bases.

In the light of this and Robertson's uncertainty relation (3.5), it seems that bases which are related by the Fourier transform should play a special role in our understanding of quantum mechanics, in the sense that they are the measurements which are most "incompatible".

3.2.2 Mutually Unbiased Bases

Let us now define the notion of MUBs more formally and state some of the known constructions and existence results.

Definition 3.2.2 (Mutually Unbiased Bases). *Let* $\mathcal{B}_1 = \{|0^{(1)}\rangle, \ldots, |(d-1)^{(1)}\rangle\}$ *and* $\mathcal{B}_2 = \{|0^{(2)}\rangle, \ldots, |(d-1)^{(2)}\rangle\}$ *be two orthonormal bases in* $\mathbb{C}^d$*. They are said to be* mutually unbiased *if* $|\langle a^{(1)}|b^{(2)}\rangle| = 1/\sqrt{d}$*, for all* $a, b \in \{0, \ldots, d-1\}$*. A set* $\{\mathcal{B}_0, \ldots, \mathcal{B}_{L-1}\}$ *of orthonormal bases in* $\mathbb{C}^d$ *is called a* set of mutually unbiased bases *if each pair of bases is mutually unbiased.*

For example, the eigenbases of the Pauli X and Z matrices in dimension $d = 2$ are mutually unbiased. Similarly, in dimension $d = 2^n$, the well-known computational and Hadamard bases are mutually unbiased. These are simply the eigenbases of $\mathbb{I}^{\otimes n}$ and $H^{\otimes n}$, where $\mathbb{I}$ and H are respectively the identity and Hadamard operations[3] on $\mathbb{C}^2$.

Let $N(d)$ denote the maximal number of bases possible in a set of MUBs in dimension d. For any dimension d, it is known that $N(d) \leq d + 1$ [7]. If $d = p^k$ is a prime power, it has been shown that $N(d) = d + 1$ and explicit constructions are known [7, 134]. In square dimensions $d = s^2$, it has been shown [131] using a construction based on Latin squares that $N(d) > \text{MOLS}(s)$, where MOLS(s) is the number of mutually orthogonal $s \times s$ Latin squares. In general, it is known that $N(mn) \geq \min[N(n), N(m)]$, for all $n, m \in \mathbb{N}$ [68, 137]. Finally, an explicit construction is known for 3 MUBs in any dimension $d \geq 2$ [52]. However, there is not much else that is known. For example, it is still an open problem as to whether there exists a set of 7 (or even 4!) MUBs in dimension $d = 6$. We refer the reader to a recent review by Durt et al. [40] for a comprehensive survey of the existing constructions of MUBs, their properties and applications.

To gain some insight into the construction of these bases, we briefly describe here the procedure due to Bandyopadhyay et al. [7], of constructing $d + 1$ MUBs when $d = p^k$, a prime power. Define the generalized Pauli operators X_p, Z_p, which act on the computational basis $\{|0\rangle, |1\rangle, \ldots, |p-1\rangle\}$ as follows:

$$X_d|j\rangle = |(j+1) \text{mod} p\rangle; \;\; Z_d|j\rangle = \omega^j |j\rangle, \; \forall j = 0, \ldots, p-1,$$

where $\omega = \exp(2\pi i/p)$. Consider *strings of Pauli operators* of the form $(X_p)^{a_1}(Z_p)^{b_1} \otimes \ldots \otimes (X_p)^{a_k}(Z_p)^{b_k}$, where $a_i, b_i \in \{0, \ldots, p-1\}$. Then, the set of all $d^2 - 1$ Pauli strings excluding the identity, can be grouped into $d + 1$ classes $\mathcal{C}_0, \ldots, \mathcal{C}_d$ such that $|\mathcal{C}_i| = d - 1, \; \forall i = 0, \ldots, d$, the elements of each $\mathcal{C}_i$ commute and $\mathcal{C}_i \cap \mathcal{C}_j = \{\phi\}, \; \forall i \neq j$. If $\mathcal{B}_i$ denotes the common eigenbasis of the operators in the set $\mathcal{C}_i$, it can be shown that the $d + 1$ bases $\mathcal{B}_0, \mathcal{B}_1, \ldots, \mathcal{B}_d$ are mutually unbiased.

The proof of this final statement follows from a general theorem proved in [7] for a basis of unitary operators in the space of $d \times d$ complex matrices, denoted as $\mathbb{M}_d(\mathbb{C})$. We state the theorem and its proof here, since it forms the basis for our construction of MUBs. First, note that there exist

[3]The Hadamard transformation is the unitary operator given by $H = \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$ in the computational basis $\{|0\rangle, |1\rangle\}$.

at most d pairwise orthogonal commuting unitary matrices in $\mathbb{M}_d(\mathbb{C})$. Let $\mathbb{U} = \{\mathcal{U}_0, \mathcal{U}_1, \ldots, \mathcal{U}_{d^2-1}\}$ be a basis of unitary matrices for $\mathbb{M}_d(\mathbb{C})$. Then, without loss of generality we can assume $\mathcal{U}_0 = \mathbb{I}$, the identity operator on $\mathbb{M}_d(\mathbb{C})$. The basis $\mathbb{U}$ is called a **maximally commuting** basis if there exists partitioning of $\mathbb{U}$ into classes $\mathcal{C}_0, \mathcal{C}_1, \ldots, \mathcal{C}_d$ such that

$$\mathbb{U} = \{\mathbb{I}\} \bigcup \mathcal{C}_0 \bigcup \mathcal{C}_1 \ldots \bigcup \mathcal{C}_d \tag{3.13}$$

where each class $\mathcal{C}_i$ contains exactly $d-1$ commuting matrices from $\mathbb{U}$. Note that $\{\mathbb{I}_d\} \bigcup \mathcal{C}_i$ is a set of d commuting orthogonal unitary matrices, which as we know is maximal.

Theorem 3.2.3 (**Unitary basis and MUBs [7]**). *If there exists a maximal commuting basis of orthogonal unitary matrices in $\mathbb{M}_d(\mathbb{C})$, then there exist a set of $d+1$ mutually unbiased bases in dimension d.*

Proof. Let $\mathbb{U}$ be a maximally commuting basis of unitary matrices as defined above, with a partitioning as described by (3.13). For any $0 < j < d$, let

$$\mathcal{C}_j = \{\mathcal{U}_{(j,0)}, \mathcal{U}_{(j,1)}, \ldots, \mathcal{U}_{(j,d-1)}\}$$

be a maximal set of orthogonal, commuting matrices, where $\mathcal{U}_{(j,0)} = \mathbb{I}$. Thus for each $0 < j < d$, there exists an orthonormal basis,

$$\mathcal{B}_j = \{|\psi_j^1\rangle, |\psi_j^2\rangle, \ldots, |\psi_j^d\rangle\},$$

such that every $\mathcal{U}_{(j,t)},\ 0 < t < d-1$ is diagonal in the basis $\mathcal{B}_j$. That is,

$$\mathcal{U}_{(j,t)} = \sum_{m=1}^{d} \lambda_{j,t,m} |\psi_j^m\rangle\langle\psi_j^m|.$$

Consider any two classes $\mathcal{C}_j$ and $\mathcal{C}_k$. The orthogonality condition on the unitaries implies that for

$0 \leq s,t \leq d-1$, $\mathrm{tr}[\mathcal{U}_{(j,s)}^{\dagger}\mathcal{U}_{(k,t)}] = d\delta_{s,0}\delta_{t,0}$. This in turn implies,

$$\begin{aligned}
\mathrm{tr}[\mathcal{U}_{(j,s)}^{\dagger}\mathcal{U}_{(k,t)}] &= \sum_{m,n=1}^{d} \lambda_{j,s,m}^{*}\lambda_{k,t,n}|\psi_j^m\rangle\langle\psi_j^m|\psi_k^n\rangle\langle\psi_k^n| \\
&= \sum_{m,n=1}^{d} \lambda_{j,s,m}^{*}\lambda_{k,t,n}|\langle\psi_j^m|\psi_k^n\rangle|^2 \\
&= d\delta_{s,0}\delta_{t,0}. \qquad (3.14)
\end{aligned}$$

Since we have set $\mathcal{U}_{(j,0)} = \mathbb{I},\ \forall 0 \leq j \leq d$, we have, $\lambda_{j,0,n} = 1,\ \forall 0 \leq j \leq d,\ 1 \leq n \leq d$. This along with (3.14) immediately gives us, that for any pair of classes $\mathcal{C}_j, \mathcal{C}_k$, the corresponding eigenbases $\mathcal{B}_j = \{|\psi_j^m\rangle\}_m$ and $\mathcal{B}_k = \{|\psi_k^m\rangle_m\}$ satisfy

$$|\langle\psi_j^m|\psi_k^n\rangle|^2 = \frac{1}{d},\ \forall 1 \leq m,n \leq d. \qquad (3.15)$$

■

A special case of the construction described above are the three mutually unbiased bases in dimension $d = 2^k$ given by the unitaries $\mathbb{I}^{\otimes k}, \mathbb{H}^{\otimes k}$ and $\mathbb{K}^{\otimes k}$ applied to the computational basis, where $\mathbb{H}$ is the Hadamard transform, and $\mathbb{K} = (\mathbb{I} + i\sigma_x)/\sqrt{2}$. In particular, three mutually unbiased bases in dimension $d = 2$ are given by the eigenvectors of the Pauli matrices X, Z, and Y . A very interesting aspect of such MUBs is that there exists an ordering $\mathcal{B}_1, \ldots, \mathcal{B}_{d+1}$ and a unitary U that cyclically permutes all bases, that is, $U\mathcal{B}_j = U\mathcal{B}_{j+1}$ for all j, where $U\mathcal{B}_{d+1} = \mathcal{B}_1$ [132]. As we see in the following section, this *symmetry* property plays an important role in obtaining tight uncertainty relations.

3.2.3 MUBs and Strong Uncertainty Relations

We have already noted that mutually unbiased bases lead to maximally strong uncertainty relations (see (3.9) and (3.11)), when we consider only 2 measurement settings. It was further shown [61,112] that strong uncertainty relations are obtained when the complete set of $d+1$ bases, $\{\mathcal{B}_0, \mathcal{B}_1, \ldots, \mathcal{B}_d\}$, exists:

$$\frac{1}{d+1}\sum_{j=0}^{d} H(\mathcal{B}_j|\rho) \geq \log(d+1) - 1. \qquad (3.16)$$

It is interesting to note that this inequality in fact follows from a bound on the collision entropy:

$$\frac{1}{d+1}\sum_{j=0}^{d} H_2(\mathcal{B}_j|\rho) \geq \log(d+1) - 1. \tag{3.17}$$

For more than 2 measurements in different bases, being mutually unbiased is indeed a *necessary* condition to obtain strong uncertainty relations. To see this, consider two bases $\mathcal{B}_1$ and $\mathcal{B}_2$ which are not mutually unbiased, so that there exist basis vectors $|x\rangle \in \mathcal{B}_1$ and $|y\rangle \in \mathcal{B}_2$ that have a higher overlap $|\langle x|y\rangle|^2 > 1/d$. Then, choosing $\rho = |x\rangle\langle x|$ yields zero entropy when measured in basis $\mathcal{B}_1$ and less than full entropy [4] when measured in the basis $\mathcal{B}_2$.

However, whereas being mutually unbiased is necessary, it was demonstrated recently that it is not a *sufficient* condition to obtain maximally strong uncertainty relations for the Shannon entropy. In particular, there do exist large sets of up to $\sqrt{d}$ mutually unbiased bases in square dimensions for which we do obtain very weak uncertainty relations [6]. Recently, Ambainis [4] has shown that for any three bases from the "standard" mutually unbiased bases construction [7, 134] in prime dimension, the lower bound cannot exceed $\left(\frac{1}{2} + o(1)\right)\log d$, for large dimensions. For dimensions of the form $4k+3$ and $8k+5$ no further assumption is needed, but the proof assumes the Generalized Riemann Hypothesis for dimensions of the form $8k+1$. Furthermore, for any $0 \leq \epsilon \leq 1/2$, there always exist $k = d^{\epsilon}$ of these bases such that the lower bound cannot be larger than $\left(\frac{1}{2} + \epsilon + o(1)\right)\log d$. Only if we use the maximal set of $d+1$ mutually unbiased bases that can be found for any given prime power dimension do we obtain strong uncertainty relations [61, 112].

At present, we also know that there do exist arbitrarily large sets of two outcome measurements that give us maximally strong uncertainty relations [129], and that in larger dimensions selecting a sufficiently large number of bases (of order $(\log d)^4$) at random does lead to strong relations [56]. However, it remains an intriguing open question as to whether there even exist three measurements with three outcomes in dimension $d > 2$ that are maximally incompatible with respect to the Shannon entropy. We refer the reader to a recent survey by Wehner and Winter [130] for a more detailed review of known results and open questions in the study of EURs.

Wootters and Sussman [132] made the interesting observation that for the maximal set of $d+1$ mutually unbiased bases in dimension $d = 2^n$, the lower bound of the entropic uncertainty relation

[4] Note that the entropy of performing a measurement corresponding to an orthonormal basis in dimension d can never exceed $\log d$, where the maximum is attained when the distribution over the outcomes is uniform $(1/d)$.

in terms of the collision entropy given in (3.17) is tight, and the minimum is attained by a state that is invariant under a unitary that cyclically permutes the set of all $d+1$ MUBs. A similar unitary was noted to exist by Chau [24]. Wootters and Sussman derive their transformation from phase space arguments. Their unitary can in fact easily be generalized to cyclically permute L bases, whenever L divides $d+1$ (see Section 3.3.1). The results in [132] have recently been generalized by Appleby [5], who shows that in prime power dimensions of the form $d = 1$ or $3 \mod 4$, there exists a unitary operation that cyclically permutes the first and second halves of the full set of MUBs. This raises the question of whether smaller sets of MUBs also exhibit such symmetries. And can we exploit such symmetries to obtain tight uncertainty relations? In particular, is the minimizing state always an invariant of such a transformation as observed for two bases in (3.8)?

3.3 Min-entropic Uncertainty Relations

Several entropic measures could be considered when it comes to quantifying uncertainty, as described in Section 3.1.1, and each has its merits. Recall that the min-entropy $H_\infty(X)$ is determined by the highest peak in the distribution $P_X(x)$ of the random variable $\mathcal{X}$, so that $2^{-H_\infty(X)} = \max_{x\in\mathcal{X}} P_X(x)$. The min-entropy is thus related to the probability of "guessing" the value of the random variable x and is of particular interest in cryptography. Since $H(\cdot) \geq H_\infty(\cdot)$, min-entropic uncertainty relations also provide us with bounds on uncertainty relations in terms of the Shannon entropy.

To gain more intuition on why the min-entropy might be a more useful quantity in cryptography than the Shannon entropy, we may consider the following example: Let $\mathcal{X} = \{0,1\}^n$ and let $x_0 = 0,\ldots,0$ be the all-zero string. Suppose that the distribution P_X is such that $P_X(x_0) = 1/2 + 1/(2^{n+1})$ and $P_X(x) = 1/(2^{n+1})$ for $x \neq x_0$, that is, the string x_0 is chosen with probability $1/2$, and with probability $1/(2^{n+1})$ the other strings are chosen uniformly at random. Then, for large n, $H(X) \approx n/2$, whereas $H_\infty(X) \approx 1$! If x were to correspond to an encryption key used to encrypt an n bit message, we would certainly not talk about security if we can guess the key with probability at least $1/2$! Yet, the Shannon entropy is quite high, and is clearly not a suitable measure of how secure an encryption key x would be. We refer the reader to [74] for a more detailed discussion on the operational meaning of the min-entropy and its usefulness in cryptography.

3.3.1 Symmetries

Apart from cryptographic applications, min-entropic uncertainty relations are also appealing since the problem of determining tight uncertainty relations can be simplified considerably in the presence of symmetries. First, note that Jensen's inequality [28] implies that

$$\frac{1}{L}\sum_{j=0}^{L-1} H_\infty(\mathcal{B}_j|\rho) \tag{3.18}$$

$$\geq -\log \frac{1}{L}\sum_{j=0}^{L-1} \max_{b^{(j)}} \mathrm{tr}(\rho|b^{(j)}\rangle\langle b^{(j)}|), \tag{3.19}$$

where the inequality becomes equality if all terms $\mathrm{tr}(\rho|b^{(j)}\rangle\langle b^{(j)}|)$ are the same. For $\vec{b} = (b^{(0)}, \ldots, b^{(L-1)}) \in \{0, \ldots, d-1\}^{\times L}$, define

$$P_{\vec{b}} := \sum_{b^{(j)}} |b^{(j)}\rangle\langle b^{(j)}| \ . \tag{3.20}$$

Determining a tight lower bound in (3.19) is thus equivalent to determining

$$\max_{\vec{b}} \max_{\rho} \mathrm{tr}(\rho P_{\vec{b}}) \ . \tag{3.21}$$

Clearly, any ζ such that

$$P_{\vec{b}} \leq \zeta \mathbb{I} \text{ for all } \vec{b} \tag{3.22}$$

gives us a lower bound for (3.18). For any set of bases, this makes the problem of finding a bound more approachable as it reduces the problem to finding the largest eigenvalue for the operator $P_{\vec{b}}$. In particular, it can be phrased as a semidefinite program to minimize ζ such that (3.22) holds for all $\vec{b}$.

It is now easy to see why symmetries simplify our goal of determining tight uncertainty relations for the min-entropy. The following Lemma makes use of the above simplification to throw light on the structure of the maximally certain states, for some sets of mutually unbiased bases. In particular, we note that for L MUBs in dimension $d = 2^n$, the state that minimizes the min-entropic uncertainty relations is an invariant of a certain unitary, whenever L divides $d+1$.

Lemma 3.3.1. *Suppose that for every* $\vec{b} \in \{0, \ldots, d-1\}$ *there exists a unitary* $U_{\vec{b}}$ *such that* $U_{\vec{b}}|b^{(j)}\rangle = |b^{(j+1 \mod L)}\rangle$. *Then there exists a* $\vec{b'}$ *such that the minimum in* (3.18) *is attained for a state* ρ *that is invariant under* $U_{\vec{b'}}$.

Proof. First of all, note that

$$\frac{1}{L}\sum_{j=0}^{L-1}(U_{\vec{b}}^{j})P_{\vec{b}}(U_{\vec{b}}^{j})^{\dagger} = P_{\vec{b}} \ , \tag{3.23}$$

and hence for $\rho_{\mathrm{sym}} = (1/L)\sum_j (U_{\vec{b}}^{j})^{\dagger}\rho(U_{\vec{b}}^{j})$

$$\mathrm{tr}(\rho_{\mathrm{sym}}P_{\vec{b}}) = \mathrm{tr}(\rho P_{\vec{b}}) \ . \tag{3.24}$$

In particular, this holds for the state $\rho = |\psi\rangle\langle\psi|$ corresponding to the eigenvector $|\psi\rangle$ with the largest eigenvalue of $P_{\vec{b'}}$. When looking for the minimizing state on the right hand side of (3.18) we can thus restrict ourselves to states which are invariant under $U_{\vec{b'}}^{j}$. Note that in this case, we further have that

$$\mathrm{tr}(\rho_{\mathrm{sym}}|b^{(j)}\rangle\langle b^{(j)}|) = \frac{1}{L}\,\mathrm{tr}(\rho P_{\vec{b}}) \ , \tag{3.25}$$

show that the inequality (3.18) is tight whenever we have such a symmetry. ■

The question of course remains, as to whether such unitaries do exist in general. Wootters and Sussman [134] have shown that there exists a unitary U that cyclically permutes the set of all $d+1$ MUBs for $d = 2^n$ by constructing a unitary that corresponds to a rotation around the origin in phase space. Clearly, by considering the unitary U^k one can trivially adapt their construction to obtain a unitary that cyclically permutes L MUBs whenever $L \cdot k = d+1$. By first translating any point in the phase space to the origin, then applying the transformation U^k and finally translating the origin back to the original point, one can obtain the desired unitaries $U_{\vec{b}}$ that enable us to find tight bounds for the min-entropic uncertainty relations. This is indeed the first time we have some insight into the structure of the states that minimize (3.18).

3.3.2 Discrete Wigner Function

The min-entropy is also related to the well studied extrema of the discrete Wigner function. To see how finding a lower bound for min-entropic uncertainty relations for $d+1$ MUBs relates to finding the extrema of the discrete Wigner function, let us first recall the properties of the discrete Wigner function. The discrete phase space is a two-dimensional vector space over a finite field $\mathbb{F}_d$, where here we focus on the case of $d = 2^n$. For every state ρ, we can associate a function W_α with every point α in the discrete phase space, known as the discrete Wigner function. For completeness, we provide a short summary on how to determine W_α; a detailed account can be found in [50].

First of all, note that the d^2 points of the discrete phase space can be parititioned into d parallel lines each of which contains d points. Any such partition is called a *striation*, and it is known that $d+1$ such striations can be found [50]. One may define the discrete Wigner function by relating each striation to one of the $d+1$ possible mutually unbiased bases [50]: Let $\lambda_{b,j}$ denote the b-th line in the striation j. With each such line, we associate a projector

$$Q(\lambda_{b,j}) = |b^{(j)}\rangle\langle b^{(j)}| \ ,$$

onto the bth element of the basis $\mathcal{B}_j$, in a specific order so as to satisfy certain symmetry constraints [50]. Defining the phase-space point operator

$$A_\alpha := \sum_{\substack{\lambda_{b,j} \\ \alpha \subset \lambda_{b,j}}} Q(\lambda_{b,j}) - \mathbb{I} \ ,$$

one can now define the discrete Wigner function as

$$W_\alpha := \frac{1}{d} \operatorname{tr}(A_\alpha \rho) \ . \tag{3.26}$$

The *extrema of the discrete Wigner function* at each phase-space point α are defined as the minimum and maximum of (3.26) over quantum states ρ.

Note that when considering $L = d+1$ mutually unbiased bases, each point α in the discrete phase space can be contained in exactly one line from each basis, as all lines in a striation, i.e., one basis are parallel. Hence, there is a one-to-one correspondence between points α in discrete phase space and vectors $\vec{b} \in \{0, \ldots, d-1\}^{\times d+1}$. In terms of the phase space operator this means

that $A_\alpha + \mathbb{I} = P_{\vec{b}}$, where $P_{\vec{b}}$ is defined as in (3.20). Note that the maximum of the discrete Wigner function,

$$W_\alpha^{\max} = \max_\rho \frac{1}{d} \operatorname{tr}(A_\alpha \rho) \ , \tag{3.27}$$

is simply the largest eigenvalue of A_α (or $P_{\vec{b}} - \mathbb{I}$) up to a factor of $1/d$. We thus have that

$$\zeta := d \cdot \left[\max_\alpha W_\alpha^{\max} + 1 \right]$$

satisfies $P_{\vec{b}} \leq \zeta \mathbb{I}$ and the maximum of the discrete Wigner function provides a lower bound to the min-entropic uncertainty relations as follows:

$$\frac{1}{d+1} \sum_{j=0}^{d} H_\infty(\mathcal{B}_j || \psi\rangle) \geq -\log \left[d \cdot \left(\max_\alpha W_\alpha^{\max} + 1 \right) \right] \ . \tag{3.28}$$

The extrema $W_\alpha^{\max}$ were evaluated numerically in [23] for small d. However, as noted in Section 3.3.1, one may use symmetries to solve the problem of determining $W_\alpha^{\max}$ directly.

3.4 New Lower Bounds on the Average Min-entropy

In this section we state and prove our min-entropic uncertainty relation for an arbitrary set of L mutually unbiased bases. As mentioned in Section 3.3, the problem of finding a lower bound for the average min-entropy reduces to the problem of finding the maximum eigenvalue of the operator $P_{\vec{b}}$ defined in (3.20). In Section 3.4.1, we use a result due to Schaffner [114] obtained using the techniques of Kittaneh [67], to show that for any set of L mutually unbiased bases in dimension d, the maximum eigenvalue of $P_{\vec{b}}$ is bounded by

$$P_{\vec{b}} \leq \frac{1}{L} \left(1 + \frac{L-1}{\sqrt{d}} \right) \mathbb{I}, \text{ for all } \vec{b}. \tag{3.29}$$

Using this, we obtain the following simple bound for the average min-entropy.

Theorem 3.4.1. *Let $\mathcal{B}_0, \ldots, \mathcal{B}_{L-1}$ be a set of mutually unbiased bases in dimension d. Then,*

$$\frac{1}{L}\sum_{j=0}^{L-1}\mathcal{H}_\infty(\mathcal{B}_j||\psi\rangle) \geq -\log\left[\frac{1}{L}\left(1+\frac{L-1}{\sqrt{d}}\right)\right] \ . \tag{3.30}$$

For the case of $L = 2$ MUBs in dimension d, our bound exactly matches the wellknown result of Deutsch (see (3.8)). For $L > 2$, the only other known lower bound for the average min-entropy is the one obtained in [114], where it is shown that for a set of $L < \sqrt{d}$ MUBs in dimension $d = 2^n$, the average min-entropy satisfies

$$\frac{1}{L}\sum_{j=0}^{L-1}\mathcal{H}_\infty(\mathcal{B}_j||\psi\rangle) \geq -\log\left[\frac{1}{L}\left(1+\frac{L-1}{\sqrt{d}}\max_{0\leq i<j\leq L-1}\sqrt{|X^i||X^j|}\right)\right], \tag{3.31}$$

where $X^i, X^j \subset \{0,1\}^n$ are subsets of n-bit strings. In the case of min-entropic uncertainty relations, these subsets contain only a single string, which corresponds to the peak of the probability distribution induced on the state $|\psi\rangle$ by the corresponding bases $\mathcal{B}^i$ and $\mathcal{B}^j$, so that

$$\max_{0\leq i<j\leq L-1}\sqrt{|X^i||Y^j|} = 1.$$

Thus, in dimension $d = 2^n$, our bound in (3.30) is clearly the same as (3.31). The reason we obtain a more general bound for $L \leq d+1$ MUBs in any dimension d is that we reduce the problem directly to an eigenvalue problem without going through the representation in terms of bit strings as in [114].

Using an alternate approach involving a Bloch sphere like representation of the basis vectors $|b^{(j)}\rangle$, we show that the maximum eigenvalue of $P_{\vec{b}}$ can be bound differently, as follows:

$$P_{\vec{b}} \leq \frac{1}{d}\left(1+\frac{d-1}{\sqrt{L}}\right)\mathbb{I}, \text{ for all } \vec{b}. \tag{3.32}$$

As we show in Section 3.4.2, this implies the following.

Theorem 3.4.2. *Let $\mathcal{B}_0, \ldots, \mathcal{B}_{L-1}$ be a set of mutually unbiased bases in dimension d. Then,*

$$\frac{1}{L}\sum_{j=0}^{L-1}\mathcal{H}_\infty(\mathcal{B}_j||\psi\rangle) \geq -\log\left[\frac{1}{d}\left(1+\frac{d-1}{\sqrt{L}}\right)\right] \ . \tag{3.33}$$

Notice that this alternate bound on the min-entropy is stronger than (3.30) when $L > d$. In particular, for the complete set of $d+1$ MUBs in dimension d, this alternate bound in (3.33) is stronger than any of the previously known bounds. When $L = d$, the two bounds in (3.30) and (3.33) that we derive are indeed equivalent.

3.4.1 Proof of Theorem 3.4.1

Recall that in Section 3.3.1, we had reduced the problem of lower bounding the average min-entropy to an eigenvalue problem (3.21). It is easy to see that this maximum is always attained at a pure state, so we can restrict the problem to an optimization over pure states. Thus, solving for ζ in

$$\max_{|\psi\rangle} \mathrm{tr}[P_{\vec{b}}|\psi\rangle\langle\psi|] \leq \zeta \ , \tag{3.34}$$

immediately leads to a min-entropic uncertainty relation of the form

$$\frac{1}{L}\sum_{j=0}^{L-1} \mathcal{H}_\infty(\mathcal{B}_j||\psi\rangle\langle\psi|) \geq -\log\zeta \ . \tag{3.35}$$

Proof. To solve the eigenvalue problem in (3.34), we use the following result of Schaffner [114], which was obtained using the methods of Kittaneh [67]. For a set of L orthogonal projectors $A_0, A_1, \ldots, A_{L-1}$, the norm of the sum satisfies,

$$\| \sum_{j=0}^{L-1} A_j \| \leq 1 + (L-1)\left(\max_{0\leq j<k\leq L-1} \| A_j A_k \| \right), \tag{3.36}$$

where $\| (.) \|$ denotes the operator norm, which here is simply the maximum eigenvalue for Hermitian operators. Applying this result to sums of basis vectors $|b^{(j)}\rangle$, we have

$$\| \sum_{j=0}^{L-1} |b^{(j)}\rangle\langle b^{(j)}| \| \leq \ 1 + \ (L-1)\left(\max_{0\leq j<k\leq L-1} \| (|b^{(j)}\rangle\langle b^{(j)}|)(|b^{(k)}\rangle\langle b^{(k)}|) \| \right),$$

which implies,

$$\| P_{\vec{b}} \| \leq \ \frac{1}{L} + \left(\frac{L-1}{L}\right)\left(\max_{0\leq j<k\leq L-1} \| \, |b^{(j)}\rangle(\langle b^{(j)}||b^{(k)}\rangle)\langle b^{(k)}| \, \| \right).$$

Recall, that for all mutually unbiased basis vectors $|b^{(j)}\rangle, |b^{(k)}\rangle$, where $b^{(j)}, b^{(k)} \in \{0, \ldots, d-1\}$,

$$\langle b^{(j)}|b^{(k)}\rangle = e^{i\phi}\frac{1}{\sqrt{d}}, \text{ for any } j \neq k,$$

where ϕ is some phase factor. Further, since the vectors $|b^{(j)}\rangle$ are normalized, the Cauchy-Schwarz inequality gives

$$\| \, |b^{(j)}\rangle\langle b^{(k)}| \, \| \leq 1, \, \forall \, b^{(j)}, b^{(k)} \in \{0, \ldots, d-1\}.$$

Combining these with (3.37) gives the following bound on the maximum eigenvalue of the operator $P_{\vec{b}}$:

$$\zeta = \frac{1}{L}\left(1 + \frac{L-1}{\sqrt{d}}\right) \, . \tag{3.37}$$

By (3.35), this immediately proves our claim. ■

3.4.2 Proof of Theorem 3.4.2

Here, we present an alternate approach to bound the maximum eigenvalue of $P_{\vec{b}}$, using a Bloch-vector-like representation of the MUB basis states. The bound that we obtain here, stated in Theorem 3.4.2, is stronger than the last one when $L > d$. In particular, when we consider the complete set ($L = d+1$) of MUBs in any dimension d, this approach yields the best known bound.

Proof. First, we switch to a basis of Hermitian operators, so that every state in $\mathcal{H}$ has a parametrization in terms of vectors in a real vector space. Any state $\rho \in \mathcal{H}$ can be written as

$$\rho = \frac{1}{d}\mathbb{I} + \frac{1}{2}\sum_{i=1}^{d^2-1} \alpha^{(i)}\hat{A}_i \, , \tag{3.38}$$

where $\{\hat{A}_i\}$ are Hermitian, traceless operators that are orthogonal with respect to the Hilbert-Schmidt norm. That is, $\mathrm{tr}[\hat{A}_i^\dagger \hat{A}_j] = 2\,\delta_{ij}$, and the scalars $\{\alpha^{(i)}\} \in \mathbb{R}$. We can thus parameterize any state in our d-dimensional Hilbert space with a vector $\vec{\alpha} = (\alpha^{(1)}, ..., \alpha^{(d^2-1)}) \in \mathbb{R}^{d^2-1}$. When ρ is a pure state ($\mathrm{tr}[\rho^2] = 1$), the vector $\vec{\alpha}$ corresponding to this pure state satisfies the following

normalization condition:

$$\begin{aligned} \mathrm{tr}\left[\left(\frac{1}{d}\mathbb{I}+\frac{1}{2}\sum_{i=1}^{d^2-1}\alpha^{(i)}\hat{A}_i\right)^2\right] &= 1 \\ \Rightarrow \frac{1}{d}+\frac{1}{2}\sum_{i=1}^{d^2-1}|\alpha^{(i)}|^2 &= 1 \\ \Rightarrow |\vec{\alpha}| = \sqrt{\sum_{i=1}^{d^2-1}|\alpha^{(i)}|^2} &= \sqrt{\frac{2(d-1)}{d}}. \end{aligned} \tag{3.39}$$

Furthermore, in this representation, the vectors $\{\vec{\alpha}_{(b,j)}\}$ corresponding to the MUB states $\{|b^{(j)}\rangle\}$ satisfy the following special properties:

- (M1) *Normalization*: $\mathrm{Tr}[|b^{(j)}\rangle\langle b^{(j)}||b^{(j)}\rangle\langle b^{(j)}|] = 1$ implies that $|\vec{\alpha}_{(b,j)}| = \sqrt{\frac{2(d-1)}{d}}$, $\forall\ b \in \{0,...,d-1\}$, $j \in \{0,...,L-1\}$. (By an argument similar to the one that leads to (3.39).)

- (M2) *Constant inner-product*: $|\langle b^{(j)}|\hat{b}^{(k)}\rangle|^2 = \frac{1}{d}$ implies that $\vec{\alpha}_{(b,j)}.\vec{\alpha}_{(\hat{b},k)} = 0$, $\forall\ j \neq k$, $\forall\ b, \hat{b} \in \{0,...,d-1\}$. This is easily seen, as follows:

$$\begin{aligned} \mathrm{tr}[|b^{(j)}\rangle\langle b^{(j)}||b^{(k)}\rangle\langle b^{(k)}|] &= \frac{1}{d}+\frac{1}{2}\sum_i \alpha^{(i)}_{(b,j)}\alpha^{(i)}_{(\hat{b},k)} = \frac{1}{d} \\ \Rightarrow \vec{\alpha}_{(b,j)}.\vec{\alpha}_{(\hat{b},k)} &= 0. \end{aligned} \tag{3.40}$$

Now, using this representation of MUB states and density operators, we can rewrite the maximization problem of (3.34) as

$$\begin{aligned} \max_{|\psi\rangle}\mathrm{tr}[P_{\vec{b}}|\psi\rangle\langle\psi|] &= \max_{|\psi\rangle}\mathrm{tr}\left[\frac{1}{L}\sum_j |b^{(j)}\rangle\langle b^{(j)}||\psi\rangle\langle\psi|\right] \\ &\leq \max_{\vec{\alpha}}\frac{1}{L}\sum_j \mathrm{Tr}\left[\left(\frac{\mathbb{I}}{d}+\frac{\sum_j \alpha^j_{(b^{(j)},j)}\hat{A}_j}{2}\right)\left(\frac{\mathbb{I}}{d}+\frac{\sum_i \alpha^{(i)}\hat{A}_i}{2}\right)\right] \\ &= \frac{1}{d}+\max_{\vec{\alpha}}\frac{1}{2L}\sum_j \vec{\alpha}_{(b^{(j)},j)}.\vec{\alpha}. \end{aligned}$$

Now we only need to find the real (d^2-1)-dimensional vector $\vec{\alpha}$, that maximizes the sum $\sum_j \vec{\alpha}_{(b^{(j)},j)}.\vec{\alpha}$.

We now define an "average" vector corresponding to each string $\vec{b}$, as follows:

$$\frac{1}{L}\sum_j \vec{\alpha}_{(b^{(j)},j)} = \vec{\alpha}_{(\text{avg})}.$$

Then, it becomes obvious that the maximum is attained when $\vec{\alpha}$ is parallel to $\vec{\alpha}^{(\text{avg})}$. Since it is a vector corresponding to a pure state, its norm is given by (3.39), so that

$$\vec{\alpha}_{(\text{max})} = \sqrt{\frac{2(d-1)}{d}}\frac{\vec{\alpha}_{(\text{avg})}}{|\vec{\alpha}_{(\text{avg})}|}.$$

Note that this maximizing vector has a constant overlap with all vectors $\vec{\alpha}^{(b^{(j)},j)}$, for a given string $\vec{b}$. In other words, for each string $\vec{b}$, the maximum is attained by the vector that makes equal angles with all the vectors that constitute the "average" vector ($\vec{\alpha}_{(\text{avg})}$) corresponding to that string. Note however that this vector may not always correspond to a valid state.

Now that we know the maximizing vector, we can go ahead and compute the value of ζ in (3.34).

$$\begin{aligned}
\max_{|\psi\rangle} \text{tr}[P_{\vec{b}}|\psi\rangle\langle\psi|] &\leq \frac{1}{d} + \max_{\vec{\alpha}} \frac{1}{2L}\sum_j \vec{\alpha}_{(b^{(j)},j)}.\vec{\alpha} \\
&= \frac{1}{d} + \frac{1}{2}\max_{\vec{\alpha}} \vec{\alpha}_{(\text{avg})}.\vec{\alpha} \\
&= \frac{1}{d} + \frac{1}{2}\frac{\vec{\alpha}_{(\text{avg})}.\vec{\alpha}_{(\text{avg})}}{|\vec{\alpha}_{(\text{avg})}|}\sqrt{\frac{2(d-1)}{d}} \\
&= \frac{1}{d}\left(1 + \frac{d-1}{\sqrt{L}}\right) ,
\end{aligned}$$

where we have used the fact that the vector $\vec{\alpha}_{(\text{avg})}$ has a constant norm which can be computed as follows:

$$\begin{aligned}
\vec{\alpha}_{(\text{avg})}.\vec{\alpha}_{(\text{avg})} &= \frac{1}{L^2}\sum_{j,k} \vec{\alpha}_{(b^{(k)},k)}.\vec{\alpha}_{(b^{(j)},j)} \\
&= \frac{1}{L^2}\sum_j \vec{\alpha}_{(b^{(j)},j)}.\vec{\alpha}_{(b^{(j)},j)} \qquad (3.41) \\
&= \frac{1}{L^2}(L)\left[\frac{2(d-1)}{d}\right] \\
\Rightarrow |\vec{\alpha}_{(\text{avg})}| &= \frac{1}{\sqrt{L}}\sqrt{\frac{2(d-1)}{d}} , \qquad (3.42)
\end{aligned}$$

thus proving our claim. Equation (3.41) follows from the fact that vectors corresponding to different MUB states have zero inner product (see property $(M2)$ above). ■

The fact that the bases are mutually unbiased was crucial in giving rise to properties $(M1)$ and $(M2)$, which in turn enabled us to identify the maximizing vector $\alpha_{\max}$. Indeed the maximizing vector corresponding to a given string $\vec{b}$ might not always correspond to a valid state, in which case the bound derived above cannot be achieved. However, there exist strings of basis elements $\vec{b}$, for which we can explicitly construct a state that has equal trace overlap with the states that constitute the corresponding operator $P_{\vec{b}}$. These are in fact states of the form

$$P_{\vec{b}} = \frac{1}{L}\sum_j |b^{(j)}\rangle\langle b^{(j)}|, \text{ where } \vec{b} = \{c, ..., c\}, \tag{3.43}$$

for any $c \in \{0, \ldots, d-1\}$. For the symmetric MUBs that we will construct in the following section (see (3.45)), an eigenstate of the unitary U that cycles between the different MUBs has the same trace overlap with each of the states $\{|b^{(j)}\rangle, j = 0, \ldots, L-1\}$, for a fixed value of b. To see this, suppose $|\phi\rangle$ is an eigenvector of U with eigenvalue λ, then for all $0 \leq j \leq L-1$ and a given value of b,

$$\begin{aligned} \mathrm{tr}[|b^{(j)}\rangle\langle b^{(j)}||\phi\rangle\langle\phi|] &= |\langle b^{(j)}|\phi\rangle|^2 = |\langle b^{(1)}|(U^\dagger)^{j-1}|\phi\rangle|^2 \\ &= (|\lambda|^2)|\langle b^{(1)}|\phi\rangle|^2 \\ &= |\langle b^{(1)}|\phi\rangle|^2. \end{aligned}$$

This is indeed the case for $L = 4$ MUBs in $d = 4$ (3.47), where the lower bound derived here is achieved by eigenstates of U.

3.5 Construction of Symmetric MUBs

We now state our main result on constructing *symmetric* sets of MUBs using the generators of the Clifford algebra in dimension $d = 2^n$.

Theorem 3.5.1. *Suppose that $2 \leq L \leq 2n+1$ is prime, and either L divides n or $L = 2n+1$. Then in dimension $d = 2^n$, there exist L mutually unbiased bases $\mathcal{B}_0, \ldots, \mathcal{B}_{L-1}$ for which there*

exists a unitary U that cyclically permutes them,

$$U\mathcal{B}_j = \mathcal{B}_{j+1 \mod L} \ . \tag{3.44}$$

In other words, we provide an explicit construction of MUBs $\mathcal{B}_0, \ldots, \mathcal{B}_{L-1}$ with $\mathcal{B}_j = \{|b^{(j)}\rangle\}$ and a unitary U such that

$$U|b^{(j)}\rangle\langle b^{(j)}|U^\dagger = |b^{(j+1 \mod L)}\rangle\langle b^{(j+1 \mod L)}| \tag{3.45}$$
$$\text{for all } |b^{(j)}\rangle \in \mathcal{B}_j \ .$$

Furthermore, in dimension $d = 4$, we actually find such a unitary for any set of L MUBs, where $2 \leq L \leq 5$. Our approach exploits properties of the Clifford algebra, and this might yield new insights into the structure of these MUBs. It is entirely distinct from the phase space approach which was used in [134] to construct such a unitary for the full set of $d+1$ MUBs. Note that our construction gives at most $O(\log d)$ bases, but shows that there is indeed an additional symmetry which has previously gone unnoticed. For $L = 2$ bases, U is simply the Fourier transform, and it would be interesting to investigate general properties of our transformation and whether it has applications in other areas.

Our construction of mutually unbiased bases makes essential use of the techniques developed in [7], together with properties of the Clifford algebra. We follow the procedure outlined in [7], described above in Section 3.2.2, but now applied to a subset of the operators in $\mathcal{S} \setminus \{\mathbb{I}\}$, where $\mathcal{S}$ is the set of all the Clifford generators and their higher products:

$$\mathcal{S} = \{\mathbb{I}, \Gamma_j, i\Gamma_i\Gamma_j, \Gamma_i\Gamma_j\Gamma_k, \ldots, i\Gamma_1 \ldots \Gamma_{2n}\}$$

We seek to group these operators into classes of commuting operators, i.e., sets $\{\mathcal{C}_0, \mathcal{C}_1, \ldots, \mathcal{C}_{L-1} \mid \mathcal{C}_j \subset \mathcal{S} \setminus \{\mathbb{I}\}\}$ of size $|\mathcal{C}_j| = d - 1$ such that

(i) the elements of $\mathcal{C}_j$ commute amongst themselves, for all $0 \leq j \leq L-1$,

(ii) $\mathcal{C}_j \cap \mathcal{C}_k = \emptyset$ for all $j \neq k$.

As shown in Theorem 3.2.3, the common eigenbases of L such classes form a set of L mutually unbiased bases.

Note that no class can contain two generators Γ_j and Γ_k since they do not commute. When forming the classes we hence ensure that each one contains exactly one generator Γ_j, which clearly limits us to constructing at most $2n+1$ such classes. The difficulty in obtaining a partitioning that is suitable for our purpose is to ensure that the unitary U that cyclically permutes the generators $\Gamma_0, \ldots, \Gamma_{L-1}$ also permutes the corresponding bases by transforming the products of the operators appropriately. Further details of our general construction, including a formal proof of Theorem 3.5.1 can be found in Appendix B.

3.5.1 Examples

Let us consider two simple examples of such classes in dimension $d = 4$. These are not obtained from our general construction, but nevertheless provide us with the necessary intuition. For $L = 3$ MUBs the classes are given by

$$\begin{aligned} \mathcal{C}_0 &= \{\Gamma_0, i\Gamma_1\Gamma_4, i\Gamma_3\Gamma_2\}, \\ \mathcal{C}_1 &= \{\Gamma_1, i\Gamma_2\Gamma_4, i\Gamma_3\Gamma_0\}, \\ \mathcal{C}_2 &= \{\Gamma_2, i\Gamma_0\Gamma_4, i\Gamma_3\Gamma_1\}. \end{aligned} \tag{3.46}$$

It is easy to see that the unitary U that achieves the transformation $\Gamma_0 \to \Gamma_1 \to \Gamma_2 \to \Gamma_0$, leaving Γ_3 and Γ_4 invariant, cyclically permutes the bases given above. We can show that an eigenstate of the *commuting* operators Γ_0, $i\Gamma_2\Gamma_4$, and $i\Gamma_3\Gamma_1$ minimizes the average collision entropy H_2. In fact this minimizing state achieves the lower bound stated earlier (3.17), showing that the uncertainty relation is indeed *tight* for 3 MUBs in $d = 4$.

For $L = 4$ we obtain the classes

$$\begin{aligned} \mathcal{C}_0 &= \{\Gamma_0, i\Gamma_1\Gamma_4, i\Gamma_2\Gamma_3\}, \\ \mathcal{C}_1 &= \{\Gamma_1, i\Gamma_2\Gamma_4, i\Gamma_3\Gamma_0\}, \\ \mathcal{C}_2 &= \{\Gamma_2, i\Gamma_3\Gamma_4, i\Gamma_0\Gamma_1\}, \\ \mathcal{C}_3 &= \{\Gamma_3, i\Gamma_0\Gamma_4, i\Gamma_1\Gamma_2\}. \end{aligned} \tag{3.47}$$

It is easy to see that the unitary U that achieves the transformation $\Gamma_0 \to \Gamma_1 \to \Gamma_2 \to \Gamma_3 \to \Gamma_0$,

leaving Γ_4 invariant, cyclically permutes the bases given above. For $L = 4$ classes the minimum in the entropic uncertainty relation for H_∞ (3.30) is attained for a state that is invariant under the transformation U. However, we also know that for $L = 4$ or $L = 8$ classes in dimension $d = 8$, no partitioning of operators is possible that satisfies our requirements. The values of L and d for which such a unitary can be found, indeed remains an interesting open question.

3.6 Tight Lower Bounds for Symmetric MUBs

Based on our construction, we now show that (3.30) is in fact tight for 4 MUBs in dimension $d = 4$, where the minimum is attained for an invariant state of the transformation U that cyclically permutes all 4 bases. Even though this is a somewhat restricted statement, it is the first time that a tight entropic uncertainty relation has been obtained for this case. The minimizing state here has an appealing symmetry property, just as for the case of 2 bases in (3.8), where the minimum is attained by a state that is invariant under the Fourier transform.

Note that our construction only gives unitaries $U_{\vec{b}}$ for $\vec{b} = (c, \ldots, c)$ for any $c \in \{0, \ldots, d-1\}$. This means that our complementarity transform U leads to tight bounds only if the largest eigenvalue of any $P_{\vec{b}}$ occurs for a $\vec{b}$ of this form. This is for indeed the case for $L = 4$ in $d = 4$, where we cannot obtain a unitary from the phase space approach of [134]. Here, the largest eigenvalue of $P_{\vec{b}}$ occurs for a $\vec{b}$ of the form $\vec{b} = (c, \ldots, c)$ for any $c \in \{0, \ldots, 3\}$. The states that achieve the lower bound are in fact eigenvectors of U, which can be expressed in terms of the MUB basis vectors as follows,

$$|\psi_b\rangle = \frac{1}{2} \sum_{j=0}^{3} \exp(i\pi j/4) |b^{(j)}\rangle, \; b \in \{0, \ldots, 3\}. \tag{3.48}$$

For the set of 4 MUBs in dimension $d = 4$ constructed from the classes given in (3.47), it is easy to check that our bound

$$\frac{1}{4} \sum_{j=0}^{3} \mathcal{H}_\infty(\mathcal{B}_j || \psi\rangle) \geq -\log \left[\frac{1}{4} \left(1 + \frac{3}{2} \right) \right] \approx 0.678$$

is tight, and the minimum is indeed achieved by an invariant state, as defined in (3.48).

For the collision entropy H_2, Wootters [132] has shown that the lower bound in (3.17) is attained by an invariant state, while considering the full set of $d + 1$ MUBs. Here, however, we are able to

show that uncertainty relation for H_2 is tight for 3 bases in $d = 4$, where our bases have an entirely different structure and the minimum is not attained by an invariant state of our transformation. Nevertheless, we have for the first time a *tight* entropic uncertainty relation for *all* possible MUBs in a dimension larger than $d = 2$, where the Bloch sphere representation makes the problem easily accessible. In $d = 4$, we have a tight relation for H_∞ for $L = 2, 4$, and tight relations for H_2 for $L = 3, 5$.

Our results indicate that due to the different properties of the minimizing state for different numbers of bases, the problem may be even more daunting than previously imagined. Yet, our work shows that in each case the minimizing state is by no means arbitrary. It has a well defined, albeit different structure in each of the cases.

In Figs. 3.1 and 3.2 we plot the bounds in (3.30) and (3.33) for the MUBs that we have constructed in dimensions $d = 4$ and $d = 8$ respectively. We also compare our lower bounds to the actual numerical minimum and the value attained by an invariant state, in each case. In both figures, the crosses denote numerically computed minima of the average min-entropy for MUBs obtained using our construction, and the circles denotes the average min-entropy for invariant states constructed as in (3.48). The bound in (3.30) is clearly tight for both $L = 3$ and $L = 4$ in dimension $d = 4$. For 4 MUBs in $d = 4$ the minimum of the average min-entropy is indeed attained by states invariant under U. Similarly, in dimension $d = 8$, the bound in (3.30) is close to tight for $L = 3$, and for $L = 6$ in $d = 8$, the minimum of the average min-entropy is nearly attained by states invariant under U.

3.7 Conclusions and Open Questions

We have shown that there exist up to $2 \leq L \leq 2n+1$ mutually unbiased bases in dimension $d = 2^n$ for which we can find a unitary that cyclically permutes these bases, whenever L is prime and L divides n or $L = 2n + 1$. This unitary is found by exploiting symmetry properties of the Clifford algebra. Our approach is quite distinct from the phase space approaches that were previously used to show that there exists such a unitary for the set of all $d+1$ MUBs [134], or for two halves of the full sets of MUBs when $d = 1$ or $3(\mod 4)$ [5]. Our unitary can be understood as a generalization of the Fourier transform, and it would be interesting to see whether it has other applications in quantum information.

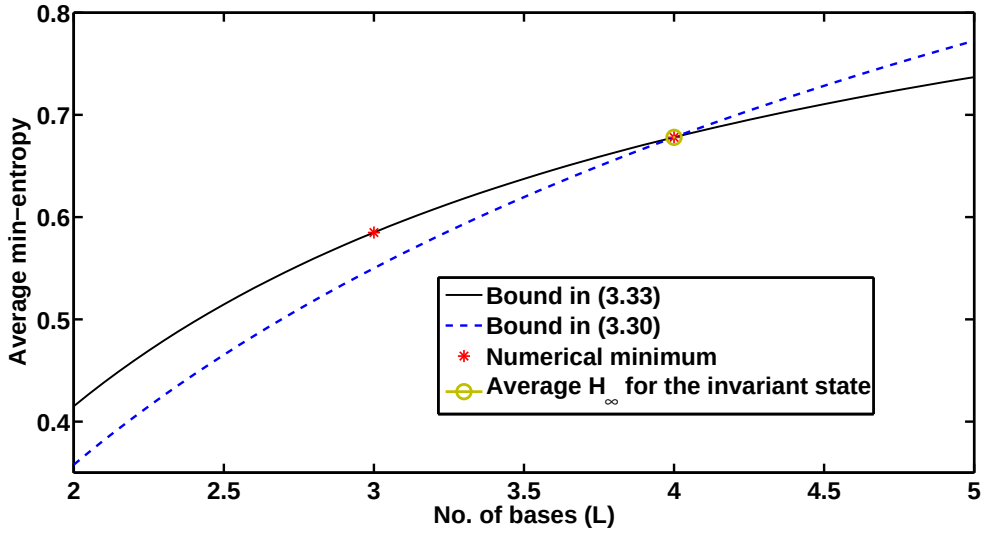


Figure 3.1: Average min-entropy for different sets of MUBs in dimension $d = 4$. The dashed line represents the bound in (3.30), and the solid line represents the bound in (3.33). The circle denotes the average min-entropy for the invariant states given in (3.48).

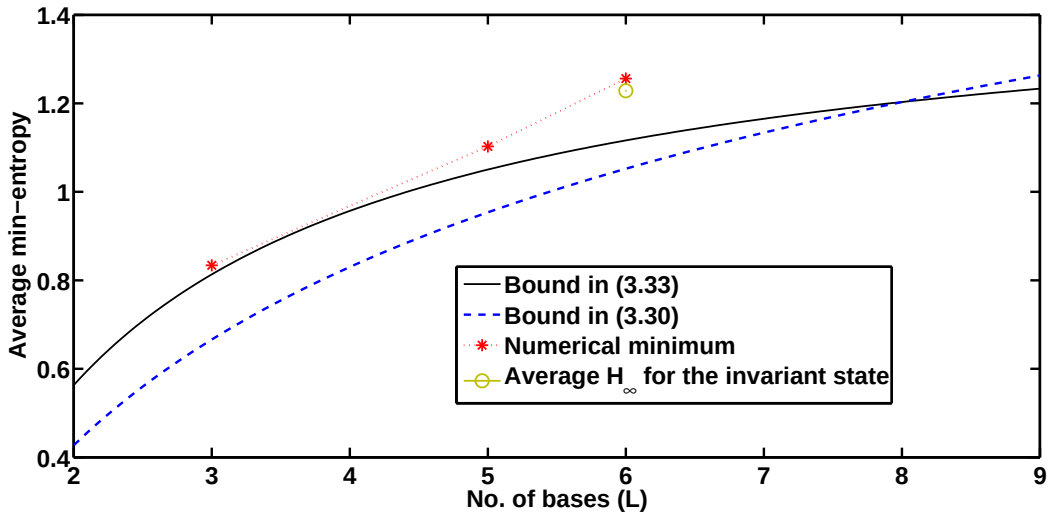


Figure 3.2: Average min-entropy for different sets of MUBs in dimension $d = 8$. The circle denotes the average min-entropy for invariant states constructed in dimension $d = 8$, similar to the states described in (3.48).

It is an interesting open question to generalize our result to other dimensions, or to a different number of bases. In prime dimension, one could consider generalized Clifford algebras [27]. Even though it does not have the full $\mathrm{SO}(2n+1)$ symmetry, it nevertheless exhibits enough symmetries to allow an exchange of generators. This stems from the way the (generalized) Clifford algebra is obtained [27, 83], which permits any transformation that preserves the p-norm for $p \geq 2$ in dimension p. Yet, this is only the first step of our construction. As for generalizing our result to any L bases in dimension $d = 2^n$, we note that it *is* indeed possible to find such classes even when L is not prime, as our example for $L = 4$ in dimension $d = 4$ shows. However, we also know that for $L = 8$ classes in dimension $d = 16$, no partitioning of operators can be found satisfying our requirements. It is an interesting open question as to when such a partitioning can be found in general.

We have also used our complementarity transform to obtain a tight uncertainty relation for the min-entropy for $L = 4$ bases in dimension $d = 4$. No tight relations are known for this case before. We also used a slight generalization of the unitary from [134] to show that when $d = 2^n$ and L divides $d + 1$, the minimizing state is an invariant of a certain unitary. This is the first time that significant insight has been obtained about the structure of the minimizing states for min-entropic uncertainty relations for mutually unbiased bases. It is an exciting open question to obtain tight relations in general, and understand the structure of the minimizing states.

Chapter 4

Achieving the Physical Limits of the Bounded-Storage Model

Two-party cryptography enables Alice and Bob to solve problems in cooperation even if they do not trust each other. Important examples of such tasks include auctions and secure identification. In the latter, Alice wants to identify herself to Bob (possibly a fraudulent ATM machine) without revealing her password, as depicted in Fig. 4. More generally, Alice and Bob wish to solve problems where Alice holds an input x (e.g.her password) and Bob holds an input y (e.g.the password an honest Alice should possess), and they want to obtain the value of some function $f(x, y)$ (e.g."yes" if $x = y$, and "no" otherwise). Known as *Secure Function Evaluation*, security in this case implies that Alice should not learn anything about y and Bob should not learn anything about x, apart from what can be inferred from $f(x, y)$ [136].

Contrary to quantum key distribution where honest Alice and Bob can work together to detect the presence of an outside eavesdropper [11, 42], two-party cryptography is made difficult by the

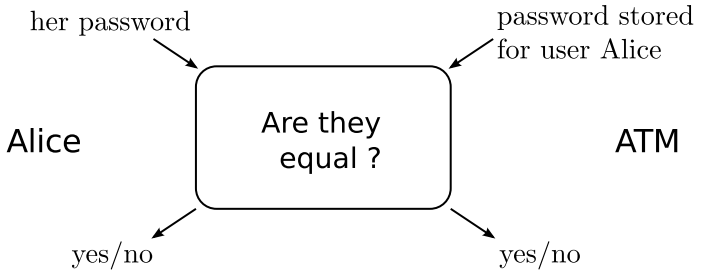


Figure 4.1: Secure Identification.

fact that Alice and Bob do not trust each other and have to fend for themselves. Indeed, two-party cryptography is impossible without making assumptions about the adversary, even when we allow quantum communication [86]. The security of most cryptographic systems in use today is based on the premise that certain computational problems are hard to solve for the adversary. Concretely, the security relies on the assumption that the adversary's computational resources are limited, and the underlying problem is hard in some precise complexity-theoretic sense. While the former assumption may be justified in practice, the latter statement is usually an unproven mathematical conjecture.

It is thus a natural question as to whether other, more physical assumptions regarding the two parties' resources allow us to obtain security without relying on any additional unproven hardness results. This is indeed known to be possible if we assume that the adversary's classical [21,93,94] or quantum storage is limited [30–32] or more generally if his memory is simply imperfect [76,115,128]. In the context of quantum cryptography, it has been shown that security can be achieved if the adversary can store strictly less then half of the qubits transmitted during the protocol. This special case is known as the *bounded-storage model*, and it has long been an open question as to whether security can still be achieved if the adversary's storage were any larger. Here,[1] we answer this question positively and demonstrate a two-party protocol which is secure as long as the adversary cannot store even a small fraction of the transmitted pulses. We also show that in the more general setting of the *noisy-storage model*, security can be extended to a larger class of noisy quantum memories.

The rest of this chapter is organized as follows. In Section 4.1 we formally introduce the noisy-storage model along with other concepts that are used in our security analysis. Next we describe the primitive of *weak string erasure* in Section 4.2 along with our protocol to realize this primitive in the noisy-storage model and the corresponding security proofs. Section 4.3 contains a brief description of some of the standard cryptographic tools that we use in our protocol to realize oblivious transfer from weak string erasure. Finally, Section 4.4 contains our protocol for oblivious transfer along with the relevant security proofs.

[1]The work described in this chapter was done in collaboration with Stephanie Wehner and the results have been published in [91]. We thank Robert König and Jürg Wullschleger for many interesting and useful discussions.

4.1 Preliminaries

We begin by briefly introducing some of the important concepts which will be used in this chapter. First, in Section 4.1.1 we define some of the entropic measures used to quantify the information gained by a dishonest party during a protocol. Next we describe the noisy-storage model in Section 4.1.2, and introduce some of the parameters that describe how well information is transmitted through the adversary's noisy quantum memory, in Section 4.1.3. We conclude with a review of earlier results that relate security to the storage rate and channel capacity in Section 4.1.4 and briefly state our contributions in this context.

4.1.1 Quantifying Adversarial Information

Consider a classical random variable X distributed according to the distribution P_X with elements $x \in \mathcal{X}$ drawn from the set $\mathcal{X}$. The probability distribution P_X over $\mathcal{X}$ can be encoded into a state $\rho_x \in \mathcal{H}_\mathcal{X}$, where $\mathcal{H}_\mathcal{X}$ is the Hilbert space $\mathcal{H}_\mathcal{X} \cong \mathbb{C}^{|\mathcal{X}|}$ with an orthonormal basis $\{|x\rangle,\ x \in \mathcal{X}\}$, as follows:

$$\rho_X = \sum_{x\in\mathcal{X}} P_X(x)|x\rangle\langle x| \ . \tag{4.1}$$

Of particular interest is the uniform distribution over X, which gives rise to the completely mixed state on $\mathcal{H}_\mathcal{X}$ denoted as

$$\tau_\mathcal{X} := \frac{1}{|\mathcal{X}|}\sum_{x\in\mathcal{X}} |x\rangle\langle x| \ . \tag{4.2}$$

Information about the classical random variable X can also be encoded into a *classical-quantum* state $\rho_{XQ} \in \mathcal{H}_\mathcal{X} \otimes \mathcal{H}_Q$ of the form

$$\rho_{XQ} = \sum_{x\in\mathcal{X}} P_X(x) \underbrace{|x\rangle\langle x|}_{X} \otimes \underbrace{\rho_x}_{Q}, \tag{4.3}$$

where $\mathcal{H}_Q$ is an additional Hilbert space. In other words, the state ρ_{XQ} encodes the ensemble of states $\{P_X(x), \rho_x\}_{x\in\mathcal{X}}$ on the Hilbert space $\mathcal{H}_Q$, where ρ_x is the conditional state on Q given $X = x$. In analyzing the security of our protocols, we will often be interested in quantifying the adversary's quantum information Q about a classical bit-string X. A natural measure of this is the average success probability that a party holding Q has in guessing the value of X. For a *cq*-state

of the form ρ_{XQ} defined above, this *guessing probability* is defined as

$$P_{guess}(X|Q)_\rho := \max_{\{M_x\}} \sum_x P_X(x) \operatorname{tr}[M_x \rho_x] \ , \tag{4.4}$$

where the maximization is over all POVMs $\{M_x\}_{x\in\mathcal{X}}$ on $\mathcal{H}_Q$. As mentioned in Section 3.3 this directly relates to the **conditional min-entropy**, as follows:

$$\mathrm{H}_\infty(X|Q)_\rho := -\log P_{guess}(X|Q)_\rho \ . \tag{4.5}$$

We will also make use of the following general definition of the min-entropy [107] for any arbitrary bipartite density operator ρ_{AB},

$$\mathrm{H}_\infty(A|B)_\rho = -\log \inf\{\operatorname{tr}[\sigma_B] | \sigma_B \geq 0 \text{ and } \rho_{AB} \leq \mathbb{I}_A \otimes \sigma_B\} \ . \tag{4.6}$$

The advantage of this definition which is equivalent to (4.5) is that it allows us to maximize over a neighborhood of our *cq*-state ρ_{XQ}, leading to the notion of **smooth min-entropy**, which is defined as follows:

$$\mathrm{H}^\varepsilon_\infty(X|Q)_\rho := \sup_{\substack{\bar{\rho}_{XQ}\geq 0 \\ \frac{1}{2}\|\bar{\rho}_{XQ}-\rho_{XQ}\|_1 \leq \operatorname{tr}(\rho_{XQ})\cdot\varepsilon \\ \operatorname{tr}(\bar{\rho}_{XQ})\leq \operatorname{tr}(\rho_{XQ})}} \mathrm{H}_\infty(X|Q)_{\bar{\rho}} \ . \tag{4.7}$$

In other words, a state ρ_{XQ} which is ϵ-close to a state $(\rho_{\bar{X}Q})$ with high min-entropy, will have high *smooth* min-entropy.

4.1.2 Noisy-Storage Model

The quantum noisy-storage model that we describe here is a generalization of the quantum bounded-storage model [31] and an earlier version of the noisy-storage model [128]. Whereas the former assumes that the adversary's quantum storage is noiseless but bounded; the latter deals with the case where the adversary's quantum storage is noisy, but allows for a larger amount of storage. The general model that we describe here due to König et al. [76], incorporates both the amount of storage and noise. As explained below, the earlier settings are special cases of this model.

Formally, the adversary's noisy quantum memory is here modeled as a device whose input states

are bounded linear operators in some Hilbert space $\mathcal{H}_{\text{in}}$. A state ρ in the device decoheres over time. That is, the state in the memory after a time t is given by $\mathcal{F}_t(\rho)$, where $\mathcal{F}_t : \mathcal{B}(\mathcal{H}_{\text{in}}) \to \mathcal{B}(\mathcal{H}_{\text{out}})$ is a completely positive trace preserving (CPTP) map (see Section 2.1.1) corresponding to the noise in the memory. Since the amount of noise might depend on the storage time, the noisy memory is in fact described by the family of maps $\{\mathcal{F}_t\}_{t>0}$. We assume that the noise is Markovian, so that $\mathcal{F}_0 = \mathbb{I}$ and $\mathcal{F}_{t_1+t_2} = \mathcal{F}_{t_1} \circ \mathcal{F}_{t_2}$. In other words, the noise in the storage only increases with time and the adversary cannot gain any information by delaying the readout. This is the only restriction imposed on the adversary, who may otherwise be all-powerful. All his other actions including computation, communication, measurement, and state preparation are allowed to be instantaneous. Further, the adversary has unlimited classical storage and computational resources, possibly quantum. In our protocol, we will introduce certain time delays Δt which force any adversary to use his storage device for a waiting time of at least Δt. The assumption of noisy-storage model entails that during such waiting times Δt in a protocol, the adversary has to measure/discard all his quantum information except what he can encode (arbitrarily) into his quantum memory. The adversary's noisy quantum storage can thus be simply modeled as the CPTP map $\mathcal{F} \equiv \mathcal{F}_{\Delta t}$, when analyzing security, rather than the family $\{\mathcal{F}_t\}_{t>0}$.

In our work, we focus on the case where the input space is an n-fold tensor product $\mathcal{H}_{\text{in}} \cong (\mathbb{C}^d)^{\otimes \nu n}$, the protocols involve n-qu*d*its of communication and the noise is of the form $\mathcal{F} \equiv \mathcal{N}^{\otimes \nu n}$ with $\mathcal{N} : \mathcal{B}(\mathbb{C}^d) \to \mathcal{B}(\mathbb{C}^d)$. The constant $\nu > 0$ is referred to as the *storage rate* as it captures the fraction of the transmitted qu*d*its that could could potentially be stored by the adversary.

To see how previously analyzed cases fit into this model, note that the bounded-storage model [31] corresponds to the case where $\mathcal{H}_{\text{in}}$ is of limited input dimension, and $\mathcal{F} = \mathbb{I}$ is the identity operator on $\mathcal{H}_{\text{in}}$. Concretely, protocols with n qubits of communication have been constructed for storage devices described by $\mathcal{H}_{\text{in}} \cong (\mathbb{C}^2)^{\otimes \nu n}$, and security established [30, 31] for storage rates $\nu < 1/4$. In the context of the noisy-storage model, protocols with n qubits of communication have been analyzed [128], where the noise $\mathcal{F} \equiv \mathcal{N}^{\otimes n}$ is an n-fold tensor product of a noisy single-qubit channel $\mathcal{N} : \mathcal{B}(\mathbb{C}^2) \to \mathcal{B}(\mathbb{C}^2)$, that is, $\mathcal{H}_{\text{in}} \cong (\mathbb{C}^2)^n$ and $\nu = 1$. The adversary was further restricted to performing product measurements on the qubits received in the protocol. The more recent noisy-storage model due to König *et al.* [76] deals with the case where $\mathcal{H}_{\text{in}} \cong (\mathbb{C}^2)^{\otimes \nu n}$, and the noise is of the form $\mathcal{F} \equiv \mathcal{N}^{\otimes \nu n}$, without any restrictions on the kinds of measurements the adversary can

perform. It is easy to see that the model we analyze here, involving qu*d*its instead of qubits, is a direct generalization of this noisy-storage model.

4.1.3 Characterizing the Noisy Quantum Storage

While the security of the protocols constructed in the earlier bounded-storage model was shown to depend only on the storage rate ν, in the noisy-storage model the security of protocols is related to the problem of sending information through the noisy-storage channel. Specifically, it was shown that the number of classical bits that can be sent through the noisy-storage channel being limited is a sufficient condition for security. We will make this statement precise in this section.

For a fixed n, we denote by $P^{\mathcal{F}}_{succ}(nR)$, the success probability of correctly transmitting a randomly chosen nR-bit string $x \in \{0,1\}^{nR}$ through the (storage) channel $\mathcal{F}$. Then,

$$P^{\mathcal{F}}_{\text{succ}}(nR) = \max_{\{M_x\},\{\rho_x\}} \frac{1}{2^{nR}} \sum_{x\in\{0,1\}^n R} \text{tr}\left(M_x \mathcal{F}(\rho_x)\right) \ , \tag{4.8}$$

where the maximization is taken over encodings $\{\rho_x\}$ on $\mathcal{H}_{\text{in}}$ and decoding POVMs $\{M_x\}$ on $\mathcal{H}_{\text{out}}$. As in [76], we show that security can be obtained for channels with the property that this decoding probability decays exponentially above a certain threshold, that is, there exist constants $n_0 > 0$ and $\gamma > 0$ such that the decoding probability satisfies $P^{\mathcal{F}}_{\text{succ}}(nR) \leq 2^{-\gamma n}$, for all $n > n_0$ and $0 < R < 1/2$. Recall that we are dealing with tensor product channels of the form $\mathcal{F} = \mathcal{N}^{\otimes \nu n}$, where n is the number of qu*d*its sent in the protocol and $\nu > 0$ is the storage rate. Our proof thus relates the security of protocols for such channels to the *classical capacity* $C_{\mathcal{N}}$ of $\mathcal{N}$. This provides a quantitative expression to our intuition that noisy channels which are of little use for classical information transmission in fact give rise to security in the noisy-storage model.

Clearly a necessary condition for an exponential decay as described above is that the classical capacity $C_{\mathcal{N}}$ of the channel be strictly smaller than the rate R at which we send information through the channel. This however is not sufficient, since $R > C_{\mathcal{N}}$ is not generally known to imply such an exponential decay for $\mathcal{F} = \mathcal{N}^{\otimes n}$. We are therefore interested in channels N which satisfy the following **strong-converse property**. The success probability (4.8) decays exponentially for rates R above the capacity, that is,

$$P^{\mathcal{N}^{\otimes n}}_{\text{succ}}(nR) \leq 2^{-n\gamma^{\mathcal{N}}(R)}, \ \text{ where } \ \gamma^{\mathcal{N}}(R) > 0, \ \text{forall } R > C_{\mathcal{N}} \ . \tag{4.9}$$

It has been shown by König and Wehner [75] that property (4.9) does indeed hold for a large class of channels including the *depolarizing channel*.

4.1.4 Security, Storage Rate, and Channel Capacity

Clearly, the storage rate ν plays a crucial role in deciding whether security can be obtained from a particular storage device. For example, in the case of bounded storage where we have no noise ($\mathcal{N} = \mathbb{I}$), we can never hope to obtain security if the adversary can store all quantum information made available to him during the protocol, that is, if $\nu = 1$ and the input space is $\mathcal{H}_{\text{in}} = (\mathbb{C}^d)^{\otimes n}$. Apart from this trivial condition, however, no bounds were known that restrict our ability to obtain security. In [30] it was shown that security can be achieved in a protocol based on qubits ($d = 2$) as long as $\nu < 1/4$. This was improved to $\nu < 1/2$ in [76]. More generally, it was shown that security in the noisy-storage model can be obtained [76] if

$$C_{\mathcal{N}} \cdot \nu < \frac{1}{2} \,, \tag{4.10}$$

where $C_{\mathcal{N}}$ is the classical capacity of the quantum channel $\mathcal{N}$.

Here, we show that for the case of bounded storage, security can be obtained if the cheating party can store all but a constant fraction of the transmitted pulses. That is, the trivial condition $\nu < 1$ stated above is in fact optimal! The honest players thereby need no quantum storage at all in order to execute the protocol. This not only settles the question, but also highlights the sharp contrast to the case of classical bounded storage, where it was shown that security can only be obtained if the adversary's classical storage is at most quadratic in the storage required by the honest players [41]. Unlike the protocols in [30,31,76,128] which use BB84 encoded qubits, we make use of states encoded in higher-dimensional mutually unbiased bases.[2] Of course, we also scale the storage size accordingly to $\mathcal{H}_{\text{in}} = (\mathbb{C}^d)^{\otimes \nu n}$ when sending d dimensional states. More specifically, we show that security in the setting of bounded storage is possible as long as

$$\nu < \frac{\log(d+1) - 1}{\log d} \,, \tag{4.11}$$

where the right hand side approaches 1 for large d. We stress that for large values of d, the resulting

[2] Mutually unbiased bases are defined in Section 3.2.2.

protocols will be much harder to implement experimentally, and even though the errors decrease exponentially with n, they converge very slowly for large d. Note, however, that here we are merely interested in exploring the fundamental physical limitations of this model.

For the general setting of noisy quantum storage we further show that security is possible for devices $\mathcal{F} = \mathcal{N}^{\otimes \nu n}$, where the channel $\mathcal{N} : \mathcal{B}(\mathbb{C}^d) \to \mathcal{B}(\mathbb{C}^d)$ satisfies the strong converse property [75], whenever

$$C_{\mathcal{N}} \cdot \nu < \log(d+1) - 1 \ , \tag{4.12}$$

thus extending the range of storage devices for which security can be achieved [76]. Our proof relies on an entropic uncertainty relation[3] for mutually unbiased bases, but is completely general in the sense that any other set of encodings satisfying such a relation could be used in our protocol instead.

We would like to emphasize that that the setting considered here differs greatly from that of quantum key distribution (QKD) [11, 42], where again, higher-dimensional states have been used to some advantage. In QKD, Alice and Bob *trust each other*, but are trying to protect themselves from an outside eavesdropper. An important advantage gained by Alice and Bob in this setting is that they can work together to try and *detect* interference by such an eavesdropper. In contrast, in the scenario we are considering there is no analogous way for Alice to check on any of Bob's actions, and vice versa. Hence, we require an entirely different proof of security from that used in quantum key distribution, and whereas results from QKD may provide some clues, they merely indicate that higher-dimensional states could be useful for our problem.

4.1.5 Techniques

We conclude this section with a brief overview of the steps involved in obtaining our result. The constant 1/2 in the bound in (4.10) is a result of using BB84-states [11] in the protocol, and stems from an uncertainty relation for measurements in these two bases [89]. It is thus natural to consider a protocol that uses more than two mutually unbiased bases (MUBs) for which uncertainty relations are known to exist [112]. Our first step is to obtain a modified protocol for the simple two-party primitive weak string erasure described in Section 4.2. This modified protocol, which we

[3] See Sections 3.2.1 and 3.2.3 for an overview of entropic uncertainty relations.

call *n*onuniform weak string erasure, is obtained using the full set of $d+1$ MUBs that are known to exist in prime power dimensions [7, 134]. Next, we show that there is still a secure protocol for the cryptographic primitive of oblivious transfer using this variant of weak string erasure. This is done by purely classical postprocessing of the output of the quantum primitive weak string erasure. Since it is known that any two-party cryptographic problem can be solved using oblivious transfer [64], our protocols can in principle be used to realize any two-party cryptographic task securely.

4.2 Weak String Erasure

The quantum primitive weak string erasure (WSE) was originally introduced in [76], where it was used as a first step to realizing other primitives including oblivious transfer. Weak string erasure provides Alice with a random bit-string $X^n \in \{0,1\}^n$, while Bob receives a randomly chosen substring $X_{\mathcal{I}} = (X_{i_1}, \ldots, X_{i_r})$, together with the index set $\mathcal{I} = \{i_1, \ldots, i_r\}$ specifying the location of the bits that he has information about. Security of weak string erasure roughly means that even a dishonest Bob cannot gain much information about Alice's entire string X^n, while security against a dishonest Alice means that she does not learn anything about the index set $\mathcal{I}$. A simple quantum protocol that securely realizes WSE in the noisy-storage model was constructed in [76], in which the honest parties do not require any quantum memory at all to execute the protocol.

4.2.1 Nonuniform Weak String Erasure

We now describe a variant of weak string erasure, which we may term *nonuniform* weak string erasure. Intuitively, this primitive provides Alice with a string $X^n = (X_1, ..., X_n) \in \{0, 1, \ldots, d-1\}^n$, where each entry X_i takes on one of d possible values. Bob obtains a set of index locations $\mathcal{I} = \{i_1, ..., i_{|\mathcal{I}|} \mid i_j \in [n]\}$, where any index $i \in \{1, \ldots, n\} =: [n]$ is chosen to be in $\mathcal{I}$ with some probability p. In addition, Bob receives the entries of the string X^n corresponding to the indices $\mathcal{I}$, which we denote by the substring $X_{\mathcal{I}} = (X_{i_1}, X_{i_2}, \ldots, X_{i_{|\mathcal{I}|}})$. Security here means that even if Alice is dishonest, she cannot learn which entries are known to Bob, i.e., she cannot learn anything about the index set $\mathcal{I}$. Conversely, if Bob is dishonest, then his information about the entire string X^n should still be limited in the sense that the probability that he can guess all of X^n given his

information B' is small. That is,

$$P_{\text{guess}}(X|B') \leq 2^{-\lambda n} \ , \tag{4.13}$$

for some $\lambda > 0$. As explained in Section 4.1.1, this is equivalent to demanding that his min-entropy is bounded by

$$\mathrm{H}_\infty(X^n|B')_\rho = -\log P_{guess}(X^n|B') \geq \lambda n \ . \tag{4.14}$$

In practice, we allow this condition to fail with error parameter ε, which is equivalent to demanding that the smooth min-entropy (4.7) satisfies,

$$\mathrm{H}^\varepsilon_\infty(X^n|B') \geq \lambda n \ . \tag{4.15}$$

Our definition of nonuniform WSE closely follows that of WSE in [76], except that the string X^n is now chosen from a larger alphabet and the indices in $\mathcal{I} \subseteq [n]$ are not chosen uniformly at random. Instead, the probability p that honest Bob learns the value of X_i for $i \in [n]$ is equal to the probability that he chooses the same basis as Alice, that is, $p = 1/(d+1)$. Clearly, the probability that Bob learns a particular subset $\mathcal{I}$ satisfies

$$\Pr(\mathcal{I}) = p^{|\mathcal{I}|}(1-p)^{n-|\mathcal{I}|}. \tag{4.16}$$

Note that we can write the subset $\mathcal{I}$ as a string $(y_1, \ldots, y_n) \in \{0,1\}^n$ where $y_i = 1$ if and only if $i \in \mathcal{I}$, allowing us to identify $|\mathcal{I}\rangle := |y_1\rangle \otimes \ldots \otimes |y_n\rangle$. As in (4.1), the probability distribution over subsets $\mathcal{I} \subseteq [n]$ can then be encoded into the state

$$\Psi(p) := \sum_{\mathcal{I} \subseteq 2^{[n]}} p^{|\mathcal{I}|}(1-p)^{n-|\mathcal{I}|} |\mathcal{I}\rangle\langle\mathcal{I}| \ . \tag{4.17}$$

Furthermore, we will follow the notation in (4.2) and denote the uniform distribution over a set $\mathcal{S}$ as

$$\tau_{\mathcal{S}} := \frac{1}{|\mathcal{S}|} \sum_{s \in \mathcal{S}} |s\rangle\langle s| \ . \tag{4.18}$$

Definition 4.2.1 (Nonuniform WSE). *An $(n, \lambda, \varepsilon, p, d)$–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 $\sigma_{X^n \mathcal{I} X_{\mathcal{I}}}$ is defined such that*

1. *The joint distribution of the n-dit string X^n and subset $\mathcal{I}$ is given by*

$$\sigma_{X^n \mathcal{I}} = \tau_{\{0,1,\ldots,d-1\}^n} \otimes \Psi(p) \ , \tag{4.19}$$

2. *The joint state ρ_{AB} created by the real protocol is equal to the ideal state: $\rho_{AB} = \sigma_{X^n \mathcal{I} X_{\mathcal{I}}}$ where we identify (A, B) with $(X^n, \mathcal{I} X_{\mathcal{I}})$.*

Security for Alice: *If A is honest, then there exists an ideal state $\sigma_{X^n B'}$ such that*

1. *The amount of information B' gives Bob about X^n is limited:*

$$\frac{1}{n} H_\infty(X^n | B')_\sigma \geq \lambda. \tag{4.20}$$

2. *The joint state $\rho_{AB'}$ created by the real protocol is ε-close to the ideal state, i.e., $\sigma_{X^n B'} \approx_\varepsilon \rho_{AB'}$ where we identify (X^n, B') with (A, B').*

Security for Bob: *If B is honest, then there exists ideal state $\sigma_{A' \hat{X}^n \mathcal{I}}$ where $\hat{X}^n \in \{0, 1, ..., d-1\}^n$ and $\mathcal{I} \subseteq [n]$ such that*

1. *The random variable $\mathcal{I}$ is independent of $A' \hat{X}^n$ and distributed over $2^{[n]}$ according to the probability distribution given by (4.16):*

$$\sigma_{A' \hat{X}^n \mathcal{I}} = \sigma_{A' \hat{X}^n} \otimes \Psi(p) \ . \tag{4.21}$$

2. *The joint state $\rho_{A'B}$ created by the real protocol is equal to the ideal state: $\rho_{A'B} = \sigma_{A'(\mathcal{I} \hat{X}_{\mathcal{I}})}$, where we identify (A', B) with $(A', \mathcal{I} \hat{X}_{\mathcal{I}})$.*

Next we outline a simple protocol that achieves the functionality described above. It is a straightforward generalization of the original protocol in [76] to multiple encodings, the main difference being that the indices in $\mathcal{I} \subseteq [n]$ are no longer chosen uniformly at random. Instead, the

probability p that honest Bob learns the value of X_i for $i \in [n]$ is equal to the probability that he chooses the same basis as Alice, that is, $p = 1/(d+1)$. Let $2^{[n]}$ denote the set of all subsets of $[n]$.

Protocol 1: Nonuniform WSE

Outputs: $x^n \in \{0,1,\ldots,d-1\}^n$ to Alice, $(\mathcal{I}, z^{|\mathcal{I}|}) \in 2^{[n]} \times \{0,1,\ldots,d-1\}^{|\mathcal{I}|}$ to Bob.

1: Alice: Picks an n-$d$it string uniformly at random, $x^n \in \{0,1,...,d-1\}^n$. She encodes each $d$it into one of the $d+1$ MUBs, $\mathcal{B}_{\theta_1},\ldots,\mathcal{B}_{\theta_n}$, that is, she chooses a basis string $\theta^n = (\theta_1,\ldots,\theta_n) \in \{0,...,d\}^n$ uniformly at random, so that the $d$it x_j is encoded in basis $\mathcal{B}_{\theta_j}$, and sends it to Bob.

2: Bob: Chooses a basis string $\tilde{\theta}^n \in \{0,1,...,d\}^n$ uniformly at random. When receiving the ith qu$d$it, he measures it in the basis $\mathcal{B}_{\tilde{\theta}_i}$, to obtain outcome $\tilde{x}_i$.

Both parties wait for a timeperiod Δt.

3: Alice: Sends the basis information θ^n to Bob, and outputs x^n.

4: Bob: Computes $\mathcal{I} := \{i \in [n] | \theta_i = \tilde{\theta}_i\}$, and outputs $(\mathcal{I}, \tilde{x}_{\mathcal{I}})$.

We are now ready to state our result on the security of nonuniform weak string erasure. We first state the general result for quantum memories, and then focus on the tensor-product channels of the type $\mathcal{F} = \mathcal{N}^{\otimes \nu n}$.

Theorem 4.2.2. *(i) Let $\delta \in]0, \frac{1}{2}[$ and let Bob's storage be given by $\mathcal{F} : \mathcal{B}(\mathcal{H}_{\text{in}}) \to \mathcal{B}(\mathcal{H}_{\text{out}})$. Then Protocol 1 is an $(n, \lambda(\delta,d), \varepsilon(\delta,d), 1/(d+1), d)$-weak string erasure protocol with min-entropy rate*

$$\lambda(\delta,d) = -\lim_{n\to\infty} \frac{1}{n} P_{succ}^{\mathcal{F}}\left((\log(d+1) - 1 - \delta)\cdot n\right) \ ,$$

and error $\varepsilon(\delta,d) = 2\exp(-f(\delta,d)n)$ with

$$f(\delta,d) := \frac{(\delta/4)^2}{32\left(\log((d+1)\cdot d) + \log\frac{4}{\delta}\right)^2} > 0. \tag{4.22}$$

(ii) Suppose $\mathcal{F} = \mathcal{N}^{\otimes \nu n}$ for a storage rate $\nu > 0$, where $\mathcal{N}$ satisfies the strong-converse property (4.9) and has capacity $C_{\mathcal{N}}$ bounded by

$$C_{\mathcal{N}} \cdot \nu < \log(d+1) - 1 \ . \tag{4.23}$$

Let $\delta \in]0, \frac{1}{2} - C_{\mathcal{N}} \cdot \nu[$. Then Protocol 1 is an $(n, \tilde{\lambda}(\delta, d), \varepsilon(\delta, d), 1/(d+1), d)$-weak string erasure protocol for sufficiently large n, where

$$\tilde{\lambda}(\delta, d) = \nu \cdot \gamma^{\mathcal{N}} \left(\frac{\log(d+1) - 1 - \delta}{\nu} \right) \ . \tag{4.24}$$

Note that for the special case of bounded-storage, where $\mathcal{N} = \mathbb{I}_d$, the classical capacity $C_{\mathcal{N}} = \log d$, so that the bound in (4.24) holds for a storage rate of

$$\nu < \frac{\log(d+1) - 1}{\log d} \approx 1, \text{ for large } d. \tag{4.25}$$

Thus for the case of bounded storage, security can in principle be obtained for any storage rate $\nu < 1$, provided we choose a large enough system size d.

It is easy to see that the protocol is correct if both parties are honest: if Alice is honest, her string $X^n = x^n$ is chosen uniformly at random from $\{0, 1, \ldots, d-1\}^n$ as desired, and if Bob is honest, he clearly obtains $\tilde{x}_i = x_i$ whenever $i \in \mathcal{I}$ for a random subset $\mathcal{I} \subseteq [n]$. In the remainder of this section, we demonstrate security when either party is dishonest.

4.2.2 Security against Dishonest Bob

We begin by modeling Bob's attack as a CPTP map $\mathcal{E} : \mathcal{B}((\mathbb{C}^d)^{\otimes n}) \to \mathcal{B}(\mathcal{H}_{\text{in}} \otimes \mathcal{H}_K)$ so that for any input state $\rho \in (\mathbb{C}^d)^{\otimes n}$, provided by Alice before the waiting time, he obtains an output state $\zeta_{Q_{\text{in}}K} = \mathcal{E}(\rho)$. Here Q_{in} is the quantum information he puts into his quantum storage, and K is any additional classical information he retains. Hence, the joint state of Alice and Bob before his storage noise is applied is of the form

$$\rho_{X^n \Theta^n K Q_{in}} = \frac{1}{d^n (d+1)^n} \sum_{x^n, \theta^n, k} P_{K|X^n = x^n, \Theta^n = \theta^n}(k) \underbrace{|x^n\rangle\langle x^n| \otimes |\theta^n\rangle\langle\theta^n|}_{\text{Alice}} \otimes \underbrace{|k\rangle\langle k| \otimes \zeta_{x^n \theta^n k}}_{\text{Bob}}, \tag{4.26}$$

where $\zeta_{x^n\theta^n k}$ is the state on $\mathcal{H}_{\text{in}}$ depending on Alice's choice of string x^n, bases θ^n and Bob's classical information k. Bob's storage then undergoes noise described by $\mathcal{F} : \mathcal{B}(\mathcal{H}_{\text{in}}) \to \mathcal{B}(\mathcal{H}_{\text{out}})$, and the state evolves to $\rho_{X^n\Theta^n K\mathcal{F}(Q_{\text{in}})}$. After time Δt, Bob also receives the basis info $\Theta^n = \theta^n$. Then their joint state is

$$\rho_{X^n\Theta^n K\mathcal{F}(Q_{\text{in}})} = \frac{1}{d^n(d+1)^n} \sum_{x^n,\theta^n,k} P_{K|X^n=x^n,\Theta^n=\theta^n}(k) \underbrace{|x^n\rangle\langle x^n|}_{\text{Alice}} \otimes \underbrace{|\theta^n\rangle\langle\theta^n| \otimes \mathcal{F}(\zeta_{x^n\theta^n k})}_{\text{Bob B'}}, \tag{4.27}$$

where Bob now holds $B' = \Theta^n K\mathcal{F}(Q_{\text{in}})$.

We next need to show that even when Bob is dishonest, he cannot learn much about the entire string X^n. In other words, our goal is to show that there exists some $\lambda > 0$, such that (4.15) is satisfied. Our proof now proceeds in three steps. First, we consider Bob's information about the string X^n given only his classical information K, and the basis information Θ^n that he receives. This can be quantified using entropic uncertainty relations in terms of the Shannon entropy for $(d+1)$ MUBs in $\mathbb{C}_d$ (see Section 3.2.1). Recall that the set of $(d+1)$ MUBs in $\mathbb{C}_d$ satisfies [112]

$$\frac{1}{d+1}\sum_{i=1}^{d+1} H(\mathcal{B}_i|\rho) \geq \log(d+1) - 1, \forall\, \rho \in \mathbb{C}_d, \tag{4.28}$$

where $H(\mathcal{B}_i|\rho)$ is the Shannon entropy of the probability distribution induced by measuring the state ρ in the basis $\mathcal{B}_i$. Using [114, Theorem 4.22] this uncertainty relation implies a bound on Bob's information in terms of the smooth min-entropy,

$$H_\infty^{\varepsilon/2}(X^n|K\Theta^n)_\rho \geq \left(\log(d+1) - 1 - \frac{\delta}{2}\right) n\ , \tag{4.29}$$

for any $0 < \delta < \frac{1}{2}$ with

$$\varepsilon = 2\exp\left(-\frac{(\delta/4)^2 n}{32\left(\log((d+1)\cdot d) + \log\frac{4}{\delta}\right)^2}\right)\ . \tag{4.30}$$

That is, the error decreases exponentially with n, as desired. Note that instead of mutually unbiased bases, we could have used any other form of encoding, which obeys a strong uncertainty relation.

Next we consider Bob's information when he is also given the output of his storage device $\mathcal{F}(Q)$. We know from [76] that the uncertainty relation (4.29) determines the rate at which Bob needs

to send information through his storage device. Using [76, Lemma 2.2] together with (4.29), we obtain

$$\begin{aligned} H_\infty^\varepsilon(X^n|\Theta^n K\mathcal{F}(Q_{in}))_\rho &\geq -\log P_{succ}^{\mathcal{F}}\left[n\left(\log(d+1)-1-\tfrac{\delta}{2}\right)-\log\tfrac{2}{\varepsilon}\right] \\ &\geq -\log P_{succ}^{\mathcal{F}}\left[n(\log(d+1)-1)-n\tfrac{\delta}{2}\right] \ , \end{aligned} \tag{4.31}$$

where, $P_{\text{succ}}^{\mathcal{F}}(nR)$ is the average probability of sending a randomly chosen string $x \in \{0,1\}^{nR}$ through the storage $\mathcal{F}$, as defined in (4.8). The second inequality above follows from the monotonicity of $P_{\text{succ}}^{\mathcal{F}}$ and the fact that $\log\frac{2}{\varepsilon} < \frac{\delta}{2}n$ for $0 < \delta < \frac{1}{2}$. By definition of the smooth min-entropy, this implies that there exists an ideal state $\sigma_{X^nB'}$ such that

1. $\sigma_{X^nB'} \approx_\varepsilon \rho_{X^nB'}$,
2. $\frac{1}{n}H_\infty(X^n|B')_\sigma \geq -\frac{1}{n}\log P_{succ}^{\mathcal{F}}\left[n\log(d+1)-n-\frac{\delta}{2}n\right]$,

which proves part (i) of Theorem 4.2.2.

In the special case that $\mathcal{F}$ is the tensor product channel $\mathcal{F} = \mathcal{N}^{\otimes\nu n}$, the right hand side of (4.31) is the success probability of sending νn bits at a rate $R = (\log(d+1)-1-\delta/2)/\nu$. The final step is to note that for channels $\mathcal{N}$ satisfying the strong converse property (4.9), this success probability drops off exponentially with n according to the parameter $\gamma^{\mathcal{N}}(R/\nu)$, whenever $R > C_{\mathcal{N}}$. Thus we obtain that there exists an ideal state $\sigma_{X^nB'}$ that is ε-close to $\rho_{X^nB'}$ and has a min-entropy

$$\frac{1}{n}H_\infty(X^n|B')_\sigma \geq \nu.\gamma^{\mathcal{N}}\left(\frac{\log(d+1)-1-\frac{\delta}{2}}{\nu}\right) > 0 \ , \tag{4.32}$$

whenever

$$C_{\mathcal{N}}\cdot\nu < \log(d+1)-1-\frac{\delta}{2} \ . \tag{4.33}$$

This proves part (ii) of Theorem 4.2.2. As before, the error rate in (4.30) shows that exponential security (in n) is possible for any constant $\delta > 0$.

4.2.3 Security against Dishonest Alice

When Alice is dishonest, it is intuitively obvious that she is unable to gain any information about the index set $\mathcal{I}$, since she never receives any information from Bob during our protocol. However, a more careful security analysis is required if we want to use weak string erasure to build more complicated primitives like oblivious transfer. The proof of security when Alice is dishonest is essentially analogous to [76] (see Section 3.4 and Figures 7 and 8), where an imaginary "simulator" with perfect quantum memory is introduced, to define the desired ideal state. Here, we merely state how to adapt the proof of [76]: here we naturally obtain $\Psi(p)$ in place of the uniform distribution $\tau_{2^{[n]}}$ in our simulation. Similarly, the subset $\mathcal{I}$ is not chosen uniformly at random, but with probability

$$\Pr(\mathcal{I}) := \left(\frac{1}{d+1}\right)^{|\mathcal{I}|} \left(\frac{d}{d+1}\right)^{n-|\mathcal{I}|} . \tag{4.34}$$

We have already discussed in Section 4.1.3 how our bounds in (4.25) and (4.33) are an improvement over pervious security bounds. As a concrete example, we consider the case when the noisy storage is in fact the depolarizing channel, that is, for any $\rho \in \mathcal{B}(\mathcal{H}_{\text{in}})$, we have $\mathcal{N}(\rho) = r\rho + (1-r)\mathbb{I}/d$. We compare the security regions obtained from previous bounds and those obtained from our new bound, for WSE with depolarizing noise in Fig. 4.2.

4.3 Cryptographic Tools

In order to construct a protocol for oblivious transfer from weak string erasure, we will need a few additional cryptographic tools, which we briefly describe here. For an integer n, we denote $[n] := \{1, \ldots, n\}$. We use $2^{[n]} := \{\mathcal{S} | \mathcal{S} \subseteq [n]\}$ to refer to the set of all possible subsets of $[n]$, including the empty set ϕ. For an n-bit string $X^n = (X_1, \ldots, X_n)$, we denote by $X_{\mathcal{S}} = (X_{i_1}, \ldots, X_{i_\ell})$, the substring corresponding to the subset $\mathcal{S} = \{i_1, \ldots, i_\ell\} \in 2^{[n]}$.

4.3.1 Privacy Amplification

Intuitively, privacy amplification [107, 108] allows us to turn a long string X, about which the adversary holds some quantum information Q, into a shorter string $Z = \mathsf{Ext}(X, R)$ about which he is almost entirely ignorant, using a 2-*universal hash function* $\mathsf{Ext}(X, R)$. A function Ext :

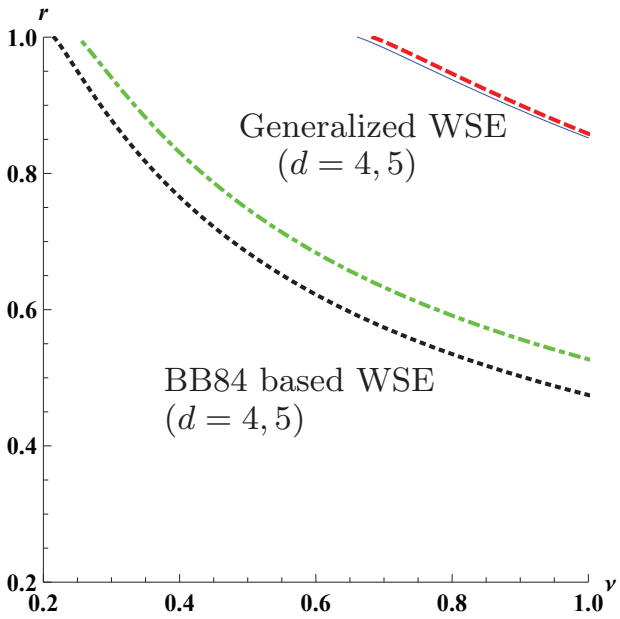


Figure 4.2: Security regions (r, ν) for weak string erasure (WSE) with depolarizing noise $\mathcal{N}(\rho) = r\rho + (1-r)\mathbb{I}/d$, in dimensions $d = 4, 5$. Previously [76], security was shown in the regions below the dotted black curve for $d = 4$ and the dot-dashed green curve for $d = 5$. Our analysis extends the security region to the solid blue curve ($d = 4$) and the dashed red curve ($d = 5$) respectively.

$\{0,1\}^n \otimes \mathcal{R} \to \{0,1\}^\ell$ is called 2-universal if for all $x \neq x' \in \{0,1\}^n$ and uniformly chosen $r \in_R \mathcal{R}$, we have, $\Pr[\mathsf{Ext}(x,r) = \mathsf{Ext}(x',r)] \leq 2^{-\ell}$. The following theorem quantifies how the length ℓ of this new string is directly related to the min-entropy $\mathrm{H}_\infty(X|Q)$ defined in (4.5).

Theorem 4.3.1 (Privacy amplification [107,108]). *Consider a set of 2-universal hash functions* $\mathsf{Ext} : \{0,1\}^n \otimes \mathcal{R} \to \{0,1\}^\ell$*, and a cq-state* ρ_{X^nQ}*, where* X^n *is an* n*-bit string. Define* $\rho_{X^nQR} = \rho_{X^nQ} \otimes \tau_{\mathcal{R}}$*, that is,* R *is a random variable uniformly distributed on* $\mathcal{R}$ *as defined in* (4.2)*, and independent of* X^nQ*. Then*

$$\rho_{\mathsf{Ext}(X^n,R)RQ} \approx_{\varepsilon'} \tau_{\{0,1\}^\ell} \otimes \rho_{RQ} \qquad \text{for } \varepsilon' := 2^{-\frac{1}{2}(\mathrm{H}_\infty^\varepsilon(X^n|Q)-\ell)-1} + 2\varepsilon \qquad \textit{for all } \varepsilon > 0 \ .$$

It should be stressed here that the extracted string $\mathsf{Ext}(X^n, R)$ is secure even if the adversary is given R in addition to Q.

4.3.2 Min-entropy Sampling

The sampling property of min-entropy was first established by Vadhan [125] in the classical case, and in [73] for the classical-quantum case. Given a cq-state ρ_{X^nQ}, where X^n is an n-bit string, an important property of smooth min-entropy is that the *min-entropy rate* $\mathrm{H}_\infty^\varepsilon(X^n|Q)/n$ is approximately preserved when considering a randomly chosen substring $X_\mathcal{S}$ of X^n. The min-entropy rate can thus be thought of as the (average) min-entropy of an individual bit X_i given Q. The analogous property in the quantum setting holds when each X_i is a block, that is, a β-bit string instead of a single bit. The following lemma proved in [76] as a special case of earlier results [73], makes this statements more precise.

Lemma 4.3.2 (Min-entropy sampling [73, 76]). *Let $\rho_{\mathbf{Z}Q}$ be a cq-state, where $\mathbf{Z} = (Z_{i,\alpha})_{(i,\alpha)\in[m]\times[\beta]} \in \mathbb{M}_{m\times\beta}(\{0,1\})$ is an $m \times \beta$-matrix with entries in $\{0,1\}$. Let $Z_i := (Z_{i,1}, \dots Z_{i,\beta}) \in \{0,1\}^\beta$ be the i-th row of $\mathbf{Z}$, such that $Z^m = (Z_1, \dots, Z_m) \equiv \mathbf{Z}$. Let*

$$\frac{\mathrm{H}_\infty^\varepsilon(\mathbf{Z}|Q)}{m\beta} \geq \lambda$$

be a lower bound on the smooth min-entropy rate of $\mathbf{Z}$ given Q. Let $\omega \geq 2$ be a constant, and assume $s, \beta \in \mathbb{N}$ are such that

$$s \geq m/4, \qquad \text{and} \qquad \beta \geq \max\left\{67, \frac{256\omega^2}{\lambda^2}\right\}, \tag{4.35}$$

and let $P_\mathcal{S}$ be the uniform distributions over subsets of $[m]$ of size s. Then,

$$\Pr_\mathcal{S}\left[\frac{\mathrm{H}_\infty^{\varepsilon+4\delta}(Z_\mathcal{S}|Q)}{s\beta} \geq \left(\frac{\omega-1}{\omega}\right)\lambda\right] \geq 1-\delta^2, \qquad \text{where } \delta = 2^{-m\lambda^2/(512\omega^2)}.$$

Adversarial partitions: In Lemma 4.3.2, βm bits were partitioned into m blocks $Z_1, \dots, Z_m$ of β bits each, by arranging the bits in the matrix $\mathbf{Z}$. It was shown in [76] that Lemma 4.3.2 can in fact be generalized to deal with arbitrarily chosen partitions, even in an adversarial manner. To formally state the corresponding generalization of Lemma 4.3.2, first observe that any partition of βm bits into m blocks is described by a permutation $\pi : [m] \times [\beta] \to [m] \times [\beta]$ where $\pi \in S_{m\beta}$. We are interested in the min-entropy of the $s\beta$-bit substring $\pi(\mathbf{Z})_S = (\pi(\mathbf{Z})_{i_1}, \dots, \pi(\mathbf{Z})_{i_s})$, where $\mathcal{S} = i_1, \dots, i_s \subset [m]$ and $\pi(\mathbf{Z})_i$ denotes the ith row of the matrix $\pi(\mathbf{Z})$ obtained by the action of π

on $\mathbf{Z}$. Note that the permutation $\Pi = \pi$ is a random variable which may in general depend on the adversary's quantum information Q. More precisely, we will assume that Π is the result of a CPTP map applied to Q. Such a CPTP map takes the form $\mathcal{E} : \mathcal{B}(\mathcal{H}_Q) \to \mathcal{B}(\mathcal{H}_{Q'} \otimes \mathcal{H}_\Pi)$, and has the property that Π is classical and a permutation in $S_{m\cdot\beta}$ for any input state. The following lemma essentially follows from the easily verified fact that the min-entropy is invariant under reordering, that is,

$$\mathrm{H}_\infty^\varepsilon(\mathbf{Z}|Q) = \mathrm{H}_\infty^\varepsilon(\pi(\mathbf{Z})|Q) \qquad \text{for all permutations} \qquad \pi \in S_{m\cdot\beta} \ . \tag{4.36}$$

Lemma 4.3.3 (**[76]**). *Let $\rho_{\mathbf{Z}Q}$ be a cq-state, where $\mathbf{Z} = (Z_{i,\alpha})_{(i,\alpha)\in[m]\times[\beta]} \in \mathbb{M}_{m\times\beta}(\{0,1\})$ is a $m \times \beta$-matrix with entries in $\{0,1\}$. Assume that*

$$\frac{\mathrm{H}_\infty^\varepsilon(\mathbf{Z}|Q)}{m\beta} \geq \lambda \ ,$$

and that λ and $s, \beta \in \mathbb{N}$ satisfy condition (4.35) of Lemma 4.3.2. Let $\mathcal{E} : \mathcal{B}(\mathcal{H}_Q) \to \mathcal{B}(\mathcal{H}_{Q'} \otimes \mathcal{H}_\Pi)$ be a permutation-computing CPTP map, as explained above, and let

$$\rho_{\mathbf{Z}Q'\Pi} = (\mathbb{I}_\mathbf{Z} \otimes \mathcal{E})\rho_{\mathbf{Z}Q} \ .$$

Finally, let $P_\mathcal{S}$ be the uniform distribution over subsets of $[m]$ of size s. Then for any constant $\omega \geq 2$

$$\Pr_\mathcal{S}\left[\frac{\mathrm{H}_\infty^{\varepsilon+4\delta}(\Pi(\mathbf{Z})_\mathcal{S}|Q'\Pi)}{s\beta} \geq \left(\frac{\omega-1}{\omega}\right)\lambda\right] \geq 1-\delta^2 \qquad \textit{where } \delta = 2^{-m\lambda^2/(512\omega^2)} \ .$$

4.3.3 Interactive Hashing

A final tool we need is *interactive hashing* [37, 113] first introduced in [100]. This is a two-party primitive where Bob inputs some string W^t, and Alice has no input. The primitive then generates two strings W_0^t, W_1^t, with the property that one of the two equals W^t. For a protocol implementing this primitive, security is intuitively specified by the following conditions: Alice does not learn which of the two strings is equal to W^t, and Bob has very little control over the other string created by the protocol. This is stated formally in the following Lemma, proved in [37, 113].

Lemma 4.3.4 (Interactive Hashing [37,113]). *There exists a protocol called* interactive hashing *between two players, Alice and Bob, such that Alice has no input, Bob has input $W^t \in \{0,1\}^t$ and both players output $(W_0^t, W_1^t) \in \{0,1\}^t \times \{0,1\}^t$, satisfying the following:*

Correctness: *If both players are honest, then $W_0^t \neq W_1^t$ and there exists a $D \in \{0,1\}$ such that (a) $W_D^t = W^t$, and (b) the distribution of W_{1-D}^t is uniform on $\{0,1\}^t \backslash \{W^t\}$.*

Security for Bob: *If Bob is honest, then $W_0^t \neq W_1^t$ and there exists a $D \in \{0,1\}$ such that $W_D^t = W^t$. If Bob chooses W^t uniformly at random, then D is uniform and independent of Alice's view.*

Security for Alice: *If Alice is honest, then for every subset $\mathcal{S} \subseteq \{0,1\}^t$,*

$$\Pr[W_0^t \in \mathcal{S} \text{ and } W_1^t \in \mathcal{S}] \leq 16 \cdot \frac{|\mathcal{S}|}{2^t} \tag{4.37}$$

4.4 Oblivious Transfer

Oblivious transfer (OT), which was first introduced by Rabin [104], is a special case of the problem of secure function evaluation described earlier. We will describe here a variant of this, known as *Fully Randomized Oblivious Transfer* [30, 44]. This primitive outputs two strings $S_0^\ell, S_1^\ell \in \{0,1\}^\ell$ to Alice, and a choice bit $C \in \{0,1\}$ and S_C^ℓ to Bob. Security means that if Alice is dishonest, she should not learn anything about C. If Bob is dishonest, we demand that there exists some random variable C such that Bob is entirely ignorant about S_{1-C}^ℓ. That is, he may learn at most one of the two strings which are generated. It has been shown that fully randomized OT can easily be converted into standard 1–2 oblivious transfer [9, 14]. Furthermore, since any two-party cryptographic problem can be solved using OT [64], our final goal will be to demonstrate a secure protocol for fully randomized OT in the noisy-storage model.

We begin, as before, with a formal definition of the primitive.

Definition 4.4.1. *An (ℓ, ε)–*fully randomized oblivious transfer (FROT) scheme *is a protocol between Alice and Bob satisfying the following:*

Correctness: *If both parties are honest, then the ideal state $\sigma_{S_0^\ell S_1^\ell C S_C^\ell}$ is defined such that*

1. *The distribution over S_0^ℓ, S_1^ℓ and C is uniform:*

$$\sigma_{S_0^\ell S_1^\ell C} = \tau_{\{0,1\}^\ell} \otimes \tau_{\{0,1\}^\ell} \otimes \tau_{\{0,1\}} \ .$$

2. *The real state $\rho_{S_0^\ell S_1^\ell C Y^\ell}$ created during the protocol is ε-close to the ideal state:*

$$\rho_{S_0^\ell S_1^\ell C Y^\ell} \approx_\varepsilon \sigma_{S_0^\ell S_1^\ell C S_C^\ell} \ , \tag{4.38}$$

where we identify $A = (S_0^\ell, S_1^\ell)$ and $B = (C, Y^\ell)$.

Security for Alice: *If Alice is honest, then there exists an ideal state $\sigma_{S_0^\ell S_1^\ell B' C}$, where C is a random variable on $\{0,1\}$, such that*

1. *Bob is ignorant about S_{1-C}^ℓ:*

$$\sigma_{S_{1-C}^\ell S_C^\ell B' C} \approx_\varepsilon \tau_{\{0,1\}^\ell} \otimes \sigma_{S_C^\ell B' C} \ .$$

2. *The real state $\rho_{S_0^\ell S_1^\ell B'}$ created during the protocol is ε-close to the ideal state:*

$$\rho_{S_0^\ell S_1^\ell B'} \approx_\varepsilon \sigma_{S_0^\ell S_1^\ell B'} \ .$$

Security for Bob: *If Bob is honest, then there exists an ideal state $\sigma_{A' S_0^\ell S_1^\ell C}$ such that*

1. *Alice is ignorant about C:*

$$\sigma_{A' S_0^\ell S_1^\ell C} = \sigma_{A' S_0^\ell S_1^\ell} \otimes \tau_{\{0,1\}} \ .$$

2. *The real state $\rho_{A' C Y^\ell}$ created during the protocol is ε-close to the ideal state:*

$$\rho_{A' C Y^\ell} \approx_\varepsilon \sigma_{A' C S_C^\ell} \ ,$$

where we identify $B = (C, Y^\ell)$.

We first state a simplified version of the actual protocol which executes fully randomized oblivious transfer from WSE, which contains all the essential ingredients for understanding the main

steps of our security proof. This is a purely classical protocol, using the quantum primitive WSE. The full protocol follows later. The main difference from the protocol presented in [76] is the fact that $\mathcal{I}$ is no longer uniform, and honest Bob only learns about pn entries x_j, whereas in the case of uniform WSE he could learn roughly $n/2$. Here, we introduce a new parameter $\eta = 2(d+1)$ in the protocol, such that with high probability Bob learns at least n/η of the indices.

Our protocol uses two ingredients, privacy amplification and the primitive interactive hashing, both of which are described in Section 4.3 above. In our context, the interactive hashing protocol is useful in the following way. It takes as inputs a subset $\mathcal{I}_{\text{tr}}$ (encoded as a string w) from Bob, and outputs two subsets $\mathcal{I}_0, \mathcal{I}_1 \in [n]$ (encoded as strings w_0, w_1) to both Alice and Bob. The protocol ensures that there exists a $c \in \{0,1\}$, such that $\mathcal{I}_c = \mathcal{I}_{\text{tr}}$, that is, one of the two subsets it outputs is equal to Bob's original input. Note that since Bob knows his input, he can of course compute c. Nevertheless, interactive hashing ensures that Alice cannot learn which subset is the same as Bob's input, that is, Alice cannot learn c. And while Bob can choose one of these subsets (namely $\mathcal{I}_c$), the choice of the other subset is not under his control (4.37). In fact, $\mathcal{I}_{1-c}$ is essentially chosen at random.

Protocol 2: Oblivious Transfer: Naive Protocol

Outputs: $(s_0^\ell, s_1^\ell) \in \{0,1\}^\ell \times \{0,1\}^\ell$ to Alice, and $(c, y^\ell) \in \{0,1\} \times \{0,1\}^\ell$ to Bob

1: Alice and Bob: Execute WSE. Alice gets a string $x^n \in \{0,1,\ldots,d-1\}^n$, Bob a set $\mathcal{I} \subset [n]$ and a string $s = x_{\mathcal{I}}$. If $|\mathcal{I}| < n/\eta$, Bob chooses uniformly at random a set $\mathcal{I}_{\text{tr}}$ of size $|\mathcal{I}_{\text{tr}}| = n/\eta$. Otherwise, he randomly truncates $\mathcal{I}$ to $|\mathcal{I}_{\text{tr}}|$ of size n/η, and deletes the corresponding values in s.

2: Alice and Bob: Execute interactive hashing with Bob's input w equal to a description of $\mathcal{I}_{\text{tr}} = \mathsf{Enc}(w)$. Interpret the outputs w_0 and w_1 as descriptions of subsets $\mathcal{I}_0$ and $\mathcal{I}_1$ of $[n]$.

3: Alice: Chooses $r_0, r_1 \in_R \mathcal{R}$ and sends them to Bob.

4: Alice: Outputs $(s_0^\ell, s_1^\ell) := (\mathsf{Ext}(x_{\mathcal{I}_0}, r_0), \mathsf{Ext}(x_{\mathcal{I}_1}, r_1))$ using the 2-universal hash function known from quantum key distribution (see Section 4.3.1) $\mathsf{Ext} : \{0,\ldots,d-1\}^{n/\eta} \times \mathcal{R} \to \{0,1\}^\ell$.

5: Bob: Computes $c \in \{0,1\}$ with $\mathcal{I} = \mathcal{I}_c$, and $x_{\mathcal{I}}$ from s. He outputs $(c, y^\ell) := (c, \mathsf{Ext}(s, r_c))$.

Let us assume that the subset $\mathcal{I}_{1-c}$ generated by the interactive hashing protocol is uniformly distributed over subsets of size n/η not equal to $\mathcal{I}$. The string $x_{\mathcal{I}_{1-c}}$ is then obtained by sampling from the string x^n which, by the definition of nonuniform WSE, has high min-entropy. We therefore expect the value s^ℓ_{1-c} to be uniform and independent of Bob's view. This should imply security for Alice, whereas security for Bob immediately follows from the properties of interactive hashing.

In this intuitive argument, we have ignored the fact that the sampling result only applies to blocks (Lemma 4.3.3) and not individual bits. To make use of the sampling results, we hence need to make slight modification to the simple protocol given above. We partition x^n (where $n = \beta m$) into m blocks of β *d*its each. To use interactive hashing in conjunction with subsets, the actual protocol requires an encoding of subsets as strings. Since our subsets will now be smaller than in [76], we choose t such that $2^t \leq \binom{m}{m/\eta} \leq 2 \cdot 2^t$, and an injective encoding $\mathsf{Enc} : \{0,1\}^t \to \mathcal{T}$, where $\mathcal{T}$ is the set of possible subsets of size m/η. Note that this again means that not all subsets can be encoded but at least half of them will. We are now ready to state our complete protocol to realize fully randomized oblivious transfer (FROT).

Protocol 2: WSE-to-FROT

Parameters: Set $\eta := 2(d+1)$. Integers n, β such that $m := n/\beta$ is a multiple of η. Outputs: $(s^\ell_0, s^\ell_1) \in \{0,1\}^\ell \times \{0,1\}^\ell$ to Alice, and $(c, y^\ell) \in \{0,1\} \times \{0,1\}^\ell$ to Bob.

1: Alice and Bob: Execute $(n, \lambda, \varepsilon, 1/(d+1), d)$–WSE.

Alice gets a string $x^n \in \{0,1,\ldots,d-1\}^n$, Bob a set $\mathcal{I} \subset [n]$ and a string $s = x_{\mathcal{I}}$. If $|\mathcal{I}| < n/\eta$, then Bob simply chooses $\mathcal{I}_{\text{tr}}$ from all subsets of size $|\mathcal{I}| = n/\eta$ uniformly at random. Otherwise, he randomly truncates $\mathcal{I}$ to $\mathcal{I}_{\text{tr}}$ of size n/η, and deletes the corresponding values in s.

We arrange x^n into a matrix $\mathbf{z} \in \mathbb{M}_{m\times\beta}(\{0,1,\ldots,d-1\})$, by $\mathbf{z}_{j,\alpha} := x_{(j-1)\cdot\beta+\alpha}$ for $(j,\alpha) \in [m] \times [\beta]$.

2: Bob:

1. Randomly chooses a string $w^t \in_R \{0,1\}^t$ corresponding to an encoding of a subset $\mathsf{Enc}(w^t)$ of $[m]$ with m/η elements.

2. Randomly partitions the n *d*its of x^n into m blocks of β *d*its each: He randomly chooses a permutation $\pi : [m] \times [\beta] \to [m] \times [\beta]$ of the entries of $\mathbf{z}$ such that he knows $\pi(\mathbf{z})_{\mathsf{Enc}(w^t)}$ (that is, these *d*its are permutation of the *d*its of s). Formally, π is uniform over permutations satisfying the following condition: for all $(j, \alpha) \in [m] \times [\beta]$ and $(j', \alpha') := \pi(j, \alpha)$, we have $(j-1) \cdot \beta + \alpha \in \mathcal{I} \Leftrightarrow j' \in \mathsf{Enc}(w^t)$.

3. Bob sends π to Alice.

3: Alice and Bob: Execute interactive hashing with Bob's input equal to w^t. They obtain $w_0^t, w_1^t \in \{0,1\}^t$ with $w^t \in \{w_0^t, w_1^t\}$.

4: Alice: Chooses $r_0, r_1 \in_R \mathcal{R}$ and sends them to Bob.

5: Alice: Outputs $(s_0^\ell, s_1^\ell) := (\mathsf{Ext}(\pi(\mathbf{z})_{\mathsf{Enc}(w_0^t)}, r_0), \mathsf{Ext}(\pi(\mathbf{z})_{\mathsf{Enc}(w_1^t)}, r_1))$.

6: Bob: Computes c, where $w^t = w_c^t$, and $\pi(\mathbf{z})_{\mathsf{Enc}(w^t)}$ from s. He outputs $(c, y^\ell) := (c, \mathsf{Ext}(\pi(\mathbf{z})_{\mathsf{Enc}(w^t)}, r_c))$.

Theorem 4.4.2 (Oblivious Transfer). *For any $\omega \geq (d+1)$ and $\beta \geq \max\{67, 256\omega^2/\lambda^2\}$, the protocol WSE-to-FROT implements an $(\ell, 43 \cdot 2^{-\frac{\lambda^2}{512\omega^2\beta}n} + 2\varepsilon)$-FROT from one instance of of $(n, \lambda, \varepsilon, p, d)$-nonuniform WSE, where $\ell := \left\lfloor \left(\left(\frac{\omega-1}{\omega}\right) \frac{\lambda}{4(d+1)} - \frac{\lambda^2}{512\omega^2\beta} \right) n - \frac{1}{2} \right\rfloor$.*

4.4.1 Security for Bob

To show that the protocol is secure against a cheating Alice, we have to show that there is no way for her to learn which of the two strings is known to honest Bob. Formally, let $\tilde{\rho}_{A''CY^\ell}$ denote the joint state at the end of the protocol, where A'' is the quantum output of a malicious Alice and (C, Y^ℓ) is the classical output of an honest Bob. In what follows, we show that we can construct an ideal state $\tilde{\sigma}_{A''S_0^\ell S_1^\ell C} = \tilde{\sigma}_{A''S_0^\ell S_1^\ell} \otimes \tau_{\{0,1\}}$ that satisfies $\tilde{\rho}_{A''CY^\ell} = \tilde{\sigma}_{A''CS_C^\ell}$.

We analyze the actions of a malicious Alice in two parts. First, she executes the WSE protocol with Bob, after which they share the state $\rho_{A'X_\mathcal{I}\mathcal{I}}$. Recall that the properties of weak string erasure ensure that a dishonest Alice does not know which *d*its $x_\mathcal{I}$ of x^n are known to Bob, that is, she is ignorant about the index set $\mathcal{I}$. In other words, there exists an ideal state $\sigma_{A'\hat{X}^n\mathcal{I}\hat{X}_\mathcal{I}}$ such that the reduced states satisfy $\rho_{A'X_\mathcal{I}\mathcal{I}} = \sigma_{A'\hat{X}_\mathcal{I}\mathcal{I}}$. Next Alice takes A' as input and interacts with Bob in the

rest of the protocol. To analyze the resulting joint output state $\tilde{\rho}_{A''CY^\ell}$, we can use the properties of weak string erasure, starting from the state $\sigma_{A'\hat{X}^n\mathcal{I}}$. Following the arguments in [76], it can be easily shown that $\tilde{\rho}_{A''CY^\ell} = \tilde{\sigma}_{A''CS_C^\ell}$, where $\tilde{\sigma}$ denotes the ideal state at the end of the protocol.

It only remains to be shown that Alice does not learn anything about C, that is, $\tilde{\sigma}_{A''S_0^\ell S_1^\ell C} = \tilde{\sigma}_{A''S_0^\ell S_1^\ell} \otimes \tau_{\{0,1\}}$. From the properties of nonuniform WSE it follows that $\sigma_{A'\hat{X}^n\mathcal{I}} = \sigma_{A'\hat{X}^n} \otimes \Psi(1/(d+1))$. Since Bob randomly truncates $\mathcal{I}$ to $\mathcal{I}_{\text{tr}}$ such that $|\mathcal{I}_{\text{tr}}| = n/\eta$, the truncated set is independent of A'. Furthermore, although $\mathcal{I}$ is not distributed uniformly over $2^{[n]}$, we can show that the truncated set $\mathcal{I}_{\text{tr}}$ is indeed distributed uniformly over all subsets of size n/η. Intuitively this follows from the fact that the distribution of the set $\mathcal{I}$ depends only on $|\mathcal{I}|$, the number of elements in $\mathcal{I}$.

Formally, let $p(\bar{A})$ denote the probability that $|\mathcal{I}| \geq n/\eta$. Then, the probability of a given truncated set $\mathcal{I}_{\text{tr}}$ can be written in terms of the probability $p(\bar{A})$ as follows,

$$p(\mathcal{I}_{\text{tr}}|\bar{A}) = \sum_{\substack{\mathcal{I}\subseteq[n] \\ |\mathcal{I}|\geq n/\eta}} \frac{p(\mathcal{I}|\bar{A})}{\binom{|\mathcal{I}|}{n/\eta}} p(|\mathcal{I}| \geq n/\eta) = \frac{1}{p(\bar{A})} \sum_{\mathcal{I}} \frac{p(\mathcal{I})}{\binom{|\mathcal{I}|}{n/\eta}},$$

independent of the choice of truncation. Here $1/\binom{|\mathcal{I}|}{n/\eta}$ is the probability that we pick a particular $\mathcal{I}_{\text{tr}}$ from the original $\mathcal{I}$ and $p(\mathcal{I}|\bar{A})$ is the conditional probability of a set $\mathcal{I}$, given that Bob obtains a sufficient number of indices. The last step is simply an application of the Bayes' rule, $p(\bar{A})p(\mathcal{I}|\bar{A}) = p(\bar{A}|\mathcal{I})p(\mathcal{I})$, where $p(\bar{A}|\mathcal{I}) = 1$ for the subsets $\mathcal{I}$ in the sum. Note that if $|\mathcal{I}| < n/\eta$ then Bob chooses a subset of the desired size uniformly at random from all subsets of size $|\mathcal{I}| = n/\eta$ and hence $\Pr(\mathcal{I}_{\text{tr}})$ is always uniform. Hence, conditioned on any fixed $S^t = s^t$, the permutation Π is uniform and independent of A'. It follows that the string S^t is also uniform and independent of A' and Π. From the properties of interactive hashing we are guaranteed that C is uniform and independent of Alice's view afterwards, and hence,

$$\tilde{\sigma}_{A''S_0^\ell S_1^\ell C} = \tilde{\sigma}_{A''S_0^\ell S_1^\ell} \otimes \tau_{\{0,1\}} \ .$$

Finally, the fact that $\mathcal{I}_{\text{tr}}$ is uniform, together with the properties of interactive hashing (Lemma 4.3.4), ensure that Alice cannot gain any information as to which of the two subsets $\mathcal{I}_0$ and $\mathcal{I}_1$ of bits are known to Bob. Hence, she cannot learn C, as desired.

4.4.2 Security for Alice

Again, it follows from weak string erasure that a dishonest Bob gains only a limited amount of information about the string X^n. The properties of interactive hashing ensure that Bob has very little control over the subset $\mathcal{I}_{1-c}$ chosen by the interactive hashing. Therefore, by the results on min-entropy sampling, Bob has only limited information about the $d$its in this subset. Privacy amplification can then be used to turn this into almost complete ignorance. The security proof for the case where Bob is dishonest is analogous to [76], and employs Lemma 4.3.3 with a subset size of $|\mathcal{S}| = m/\eta$.

4.4.3 Correctness

It remains to prove that if both parties are honest, then honest Bob can indeed learn the desired S_C. This requires us to show that for our choice of η, Bob can learn sufficiently many indices $i \in [n]$.

Lemma 4.4.3 (Correctness). *Protocol WSE-to-FROT satisfies correctness with an error of*

$$43 \cdot 2^{-\frac{\lambda^2}{512\omega^2\beta}n} \,. \tag{4.39}$$

First, using the Hoeffding bound [60], we show that the probability that a subset of $[n]$—each of whose entries is chosen with probability $p = 1/(d+1)$—has less than n/η elements is at most $\exp(-2n/\eta^2)$. Consider a sequence of independent random variables $\{X_1, \ldots, X_n\}$, which are bounded as follows: $\Pr(X_i - \mathbf{E}(X_i) \in [a_i, b_i]) = 1,\ \forall 1 < i < n$. Then, Hoeffding's inequality states that the sum $S = X_1 + \ldots + X_n$ satisfies,

$$\Pr(\mathbf{E}(S) - S \geq t) \leq \exp\left(-\frac{2t^2}{\sum_{i=1}^{n}(b_i - a_i)^2}\right). \tag{4.40}$$

In our context, X_i is the binary variable which takes on the value 1 if the index $i \in \mathcal{I}$, and 0 otherwise. The sum S is thus simply equal to $|\mathcal{I}|$, the number of elements in the index set $\mathcal{I}$, which is a random subset of $[n]$. For the case of $d+1$ encodings, $\Pr(X_i = 1) = 1/(d+1)$ and

$\Pr(X_i = 0) = d/(d+1)$, so that the expectation value satisfies

$$\mathbf{E}(S) = \mathbf{E}(|\mathcal{I}|) = \frac{n}{d+1} \ . \tag{4.41}$$

Furthermore, we can take $a_i = 0$ and $b_i = 1$ for all i. Applying Hoeffding's inequality to the sum $S = |\mathcal{I}|$ gives

$$\Pr(\frac{n}{d+1} - |\mathcal{I}| \geq \frac{n}{d+1} - \frac{n}{\eta}) \leq \exp\left(-2n\left[\frac{1}{d+1} - \frac{1}{\eta}\right]^2\right). \tag{4.42}$$

Rearranging terms, we obtain the probability that a random set $\mathcal{I}$ has less than n/η elements:

$$\Pr(|\mathcal{I}| \leq \frac{n}{\eta}) \leq \exp\left(-2n\left[\frac{1}{d+1} - \frac{1}{\eta}\right]^2\right). \tag{4.43}$$

Since our work is mainly a proof of achievability, we do not at this stage care about optimality or efficiency. We simply pick a choice of η that will satisfy this condition, and set $\eta = 2(d+1)$. Thus, the probability that a random subset of $[n]$ has less than n/η elements is at most $\exp(-2n/\eta^2)$.

Let $\xi := 2^{-n/\eta^2}$. We have to show that the state $\tilde{\rho}_{S_0^\ell S_1^\ell C Y^\ell}$ at the end of the protocol is close to the given ideal state $\tilde{\sigma}_{S_0^\ell S_1^\ell C S_C^\ell}$. As shown above, the probability that a subset of $[n]$ has less than n/η elements is at most

$$\exp(-2n/\eta^2) \leq \xi \ . \tag{4.44}$$

Hence, the probability that Bob does not learn sufficiently many indices when both parties are honest is at most ξ. Let $\mathcal{A}$ be the event that $|\mathcal{I}| \geq n/\eta$. It remains to show that the state $\tilde{\rho}_{S_0^\ell S_1^\ell C Y^\ell|\mathcal{A}}$ is close to the given ideal state $\sigma_{S_0^\ell S_1^\ell C S_C^\ell}$.

Note that the correctness condition of WSE ensures that the state created by WSE is equal to $\rho_{X^n \mathcal{I} X_\mathcal{I}} = \sigma_{X^n \mathcal{I} X_\mathcal{I}}$, where $\sigma_{X^n \mathcal{I}} = \tau_{\{0,1,\ldots,d-1\}^n} \otimes \Psi(1/(d+1))$. Since $\mathcal{I}_0$ and $\mathcal{I}_1$ are chosen independently of X^n, $X_{\mathcal{I}_0}$ and $X_{\mathcal{I}_1}$ have a min-entropy of n/η each. Since $\ell \leq n/2\eta \leq n/\eta - 2\log 1/\xi$, it follows from privacy amplification that S_C^ℓ is independent and ξ-close to uniform. Since dishonest Bob is only more powerful than honest Bob, we can carry over from the proof against dishonest Bob, that S_{1-C}^ℓ is independent and uniform except with an error of at most

$\hat{\varepsilon} = 41 \cdot 2^{-\frac{\lambda^2}{512\omega^2\beta}n}$, where we used the fact that Bob is also honest during weak string erasure ($\varepsilon = 0$). Finally, by the same arguments showing security for Bob, we have that C is uniform and independent of S_0^ℓ and S_1^ℓ. Hence,

$$\rho_{S_0^\ell S_1^\ell C|\mathcal{A}} \approx_{\xi+\hat{\varepsilon}} \sigma_{S_0^\ell S_1^\ell C} \ .$$

Since the extra condition on the permutation Π implies that Bob can indeed calculate $\Pi(\mathbf{Z})_{\mathsf{Enc}(W)}$ from $X_{\mathcal{I}}$, we have that $Y^\ell = S_C^\ell$. Using $\Pr[\mathcal{A}] \geq 1 - \xi$, we get

$$\rho_{S_0^\ell S_1^\ell C Y^\ell} \approx_{2\xi+\hat{\varepsilon}} \sigma_{S_0^\ell S_1^\ell C S_C^\ell} \ .$$

Finally, since $\lambda \leq 1$, $\beta > 1$ and $\omega \geq (d+1)$, we have $1/\eta^2 = 1/(4(d+1)^2) > \lambda^2/(512\omega^2\beta)$. Adding up all errors and noting that

$$2 \cdot 2^{-\frac{1}{\eta^2}n} \leq 2 \cdot 2^{-\frac{\lambda^2}{512\omega^2\beta}n} \ ,$$

we obtain the error bound in (4.39).

4.5 Conclusion

We have shown that any two-party cryptographic primitive can be implemented securely in the setting of bounded quantum storage, even if the adversary can store all but a fraction of the transmitted pulses. This is optimal, since we can never hope to achieve security if the cheating party could store *all* quantum communication made available to him. Our result demonstrates that there is no physical principle that prevents us from achieving security even with a very high storage rate $\nu < 1$. We have also shown in the noisy-storage setting that security is possible for a much larger range of noisy quantum memories.

To achieve our result we use higher-dimensional states which are difficult to create in practice. It is therefore an interesting open question whether the same result could be obtained using merely BB84 encoded qubits. We may also note, that our approach merely relies on the existence of entropic uncertainty relations for multiple encodings, and our protocols and proofs will carry over if we were to use any other encodings for which strong uncertainty relations are known to exist. For example, it is conceivable that uncertainty relations for multiple encodings can be based on top of

BB84 encoded qubits [54], which would lead to a protocol that is easy to implement experimentally.

Appendix A

The AQEC Algorithm for Qubit Codes

For the most practically relevant case of codes encoding a single qubit, that is, when $\mathcal{C}$ is of dimension $d = 2$, the algorithm given in Section 2.5 has considerable simplifications. In this appendix, we show that the maximum eigenvalue $\lambda_{\max}$ of Δ_{sum} required in Step 2 can be easily computed without requiring any diagonalization of Δ_{sum}. In addition, we show that the fidelity loss η_P needed in Step 3 is also simple to compute. In general, obtaining the value for η_P requires an exhaustive optimization over all pure states in the code space. For a qubit code, however, we will see that computing η_P requires no such exhaustive optimization, and can be done using only eigenanalysis.

A.1 Computing the Maximum Eigenvalue of Δ_{sum}

Given an orthonormal basis $\{|v_1\rangle, |v_2\rangle\}$ for the qubit codespace, we can construct the Pauli basis $\{\sigma_0 \equiv \mathbb{I}_2, \sigma_x, \sigma_y, \sigma_z\}$:

$$\begin{aligned} \sigma_0 &= |v_1\rangle\langle v_1| + |v_2\rangle\langle v_2| \equiv \mathbb{I}_2, \\ \sigma_x &= |v_1\rangle\langle v_2| + |v_2\rangle\langle v_1|, \\ \sigma_y &= -i(|v_1\rangle\langle v_2| - |v_2\rangle\langle v_1|), \\ \text{and} \quad \sigma_z &= |v_1\rangle\langle v_1| - |v_2\rangle\langle v_2|. \end{aligned} \tag{A.1}$$

The Pauli basis forms a basis for the qubit operator algebra. The AQEC conditions (2.22) can then be written as

$$PE_i^\dagger \mathcal{E}(P)^{-1/2} E_j P = \beta_{ij} \mathbb{I}_2 + \sum_a \gamma_{ij}^a \sigma_a, \tag{A.2}$$

where $a = x, y, z$, and γ_{ij}^a are some coefficients so that $\Delta_{ij} \equiv \sum_a \gamma_{ij}^a \sigma_a$. The right-hand side of (A.2) can be viewed as an expansion of the left-hand side in the Pauli basis.

The first simplification of the algorithm given in Section 2.5, for a qubit code, comes from the following lemma.

Lemma A.1.1. *For $\mathcal{C}$ encoding a single qubit,*

$$\Delta_{sum} = \Big(1 - \sum_{ij} |\beta_{ij}|^2\Big) P. \tag{A.3}$$

Proof. Using $\Delta_{ij} = \sum_a \gamma_{ij}^a \ \sigma_a$, we can write Δ_{sum} as

$$\begin{aligned} \Delta_{\text{sum}} &= \sum_{ij} \sum_{ab} \gamma_{ij}^{a*} \gamma_{ij}^b \sigma_a \sigma_b \\ &= P \sum_{ij} \sum_a \left|\gamma_{ij}^a\right|^2 + \sum_{ij} \sum_{a \neq b} \gamma_{ij}^{a*} \gamma_{ij}^b \sigma_a \sigma_b. \end{aligned} \tag{A.4}$$

Since $\gamma_{ij}^{a*} = \gamma_{ji}^a$, the second term on the right-hand side of (A.4) can be written as

$$\begin{aligned} &\frac{1}{2}\Big(\sum_{ij} \sum_{a \neq b} \gamma_{ji}^a \gamma_{ij}^b \sigma_a \sigma_b + \sum_{ij} \sum_{a \neq b} \gamma_{ji}^a \gamma_{ij}^b \sigma_a \sigma_b \Big) \\ &= \frac{1}{2}\Big(\sum_{ij} \sum_{a \neq b} \gamma_{ji}^a \gamma_{ij}^b \sigma_a \sigma_b + \sum_{ij} \sum_{a \neq b} \gamma_{ij}^b \gamma_{ji}^a \sigma_b \sigma_a \Big) \\ &= \frac{1}{2} \sum_{ij} \sum_{a \neq b} \gamma_{ji}^a \gamma_{ij}^b \left(\sigma_a \sigma_b + \sigma_b \sigma_a\right) = 0. \end{aligned}$$

In the first equality, we have interchanged the indices $a \leftrightarrow b$ and $i \leftrightarrow j$ in the second term. The last equality comes from the fact that the Pauli matrices anticommute. We are thus left with only the first term in (A.4). Now, the TP condition for $\mathcal{R}_P \circ \mathcal{E}$ acting on $\mathcal{C}$ gives $\sum_{ij} |\beta_{ij}|^2 + \sum_{ij} \sum_a |\gamma_{ij}^a|^2 = 1$. This means that we have $\Delta_{\text{sum}} = P \sum_{ij} \sum_a |\gamma_{ij}^a|^2 = (1 - \sum_{ij} |\beta_{ij}|^2) P$, thus proving the lemma. ■

Lemma A.1.1 tells us that Δ_{sum} has a flat spectrum. Its maximum eigenvalue is thus simply

given by

$$\lambda_{\max} = \|\Delta_{\text{sum}}\| = 1 - \sum_{ij} |\beta_{ij}|^2, \tag{A.5}$$

with β_{ij} as in (2.27b). Observe that the proof of Lemma A.1.1 required detailed properties of the Pauli matrices which, together with the identity operator $\mathbb{I}_2$, form a basis for the qubit operator space. In general, this lemma does not hold for higher-dimensional codes.

A.2 Computing the Fidelity Loss for the Transpose Channel

In Step 3 of our algorithm in Section 2.5, we have to compute the fidelity loss η_P for the recovery $\mathcal{R}_P$, or equivalently, the worst-case fidelity for the map $\mathcal{R}_P \circ \mathcal{E}$. For a qubit code, $(\mathcal{R}_P \circ \mathcal{E}) : \mathcal{B}(\mathcal{C}) \to \mathcal{B}(\mathcal{C})$ is just a qubit map. Observe that $\mathcal{R}_P \circ \mathcal{E}$ is not only CPTP but is also unital (i.e., $(\mathcal{R}_P \circ \mathcal{E})(P) = P$). Hence, we only need to consider a qubit map that is CPTP and unital. Here, we show that the worst-case fidelity for a unital, CPTP qubit map is very easy to compute.

Even though our context here only requires a unital, CPTP qubit map, we begin with a general CP map $\Phi \sim \{K_i\}$ on a d-dimensional Hilbert subspace $\mathcal{C}$. This will highlight why the qubit case is particularly simple. It is convenient to use a matrix description for Φ by going to the Hilbert-Schmidt space in which operators on $\mathcal{C}$ are represented as vectors and linear maps on operators are represented as matrices. The Hilbert-Schmidt vector space is endowed with the inner product $\langle A|B\rangle \equiv \text{tr}(A^\dagger B)$ where $|A\rangle$ and $|B\rangle$ are the vectors corresponding to the operators A and B. To go from the operator description to the Hilbert-Schmidt space, one picks any orthonormal basis $\{O_i\}$ for $\mathcal{B}(\mathcal{H})$. Then, the vector corresponding to any operator $A \in \mathcal{B}(\mathcal{H})$ has entries given by $\text{tr}\{O_i^\dagger A\}$; the matrix corresponding to a linear map $\mathcal{E}$ on operators in $\mathcal{B}(\mathcal{H})$ has matrix elements given by $\text{tr}\{O_i^\dagger \mathcal{E}(O_j)\}$.

Let us make use of a Hermitian basis $\{O_0, O_1, \ldots, O_{d^2-1}\}$ for $\mathcal{B}(\mathcal{C})$ where

$$O_0 \equiv \mathbb{I},\ O_\alpha^\dagger = O_\alpha\ \forall \alpha,\ \ \text{tr}\{O_\alpha^\dagger O_\beta\} = \delta_{\alpha\beta} d\ \ \forall\ \alpha, \beta.$$

The operators O_α for $\alpha = 1, \ldots, d^2 - 1$ are clearly traceless. Such a basis exists for any d—for example, one can use the standard generators of the $SU(d)$ group, augmented with the identity operator, as the basis elements. Then, the action of Φ can be represented as a matrix $\mathcal{M}$ acting on

the Hilbert-Schmidt space with matrix elements

$$\mathcal{M}_{\alpha\beta} \equiv \frac{1}{d} \operatorname{tr}\{O_\alpha \Phi(O_\beta)\}. \tag{A.6}$$

Since Φ is CP and O_i's are Hermitian, we have that $\mathcal{M}^*_{\alpha\beta} = \mathcal{M}_{\alpha\beta}$, so $\mathcal{M}$ is a real matrix.

Now, the density operator corresponding to any pure state $|\psi\rangle$ in $\mathcal{C}$ can be expanded in terms of the Hermitian basis as

$$|\psi\rangle\langle\psi| = \frac{1}{d}\left(\mathbb{I} + \mathbf{s}\cdot\mathbf{O}\right) = \frac{1}{d}\ \vec{s}\cdot\vec{O}, \tag{A.7}$$

where $\mathbf{s}$ is a real (d^2-1)-element vector, $\vec{s} \equiv (1, \mathbf{s})$, $\mathbf{O} \equiv (O_1, O_2, \ldots, O_{d^2-1})$, and $\vec{O} \equiv (\mathbb{I}, \mathbf{O})$. $\mathbf{s}$ is not an arbitrary vector, but has to obey some constraints in order for it to correspond to a pure state.

Using (A.6) and (A.7), we can compute the fidelity for a state $|\psi\rangle \in \mathcal{C}$ under the map Φ as

$$\begin{aligned} F^2\left[|\psi\rangle, \Phi(|\psi\rangle\langle\psi|)\right] &= \operatorname{tr}\{|\psi\rangle\langle\psi|\Phi(|\psi\rangle\langle\psi|)\} \\ &= \frac{1}{d^2}\sum_{\alpha,\beta=0}^{d^2-1} s_\alpha s_\beta \operatorname{tr}\{O_\alpha \Phi(O_\beta)\} \\ &= \frac{1}{d}\sum_{\alpha,\beta=0}^{d^2-1} s_\alpha\ \mathcal{M}_{\alpha\beta}\ s_\beta \\ &= \frac{1}{d}\ s^T \mathcal{M}\ s, \end{aligned} \tag{A.8}$$

where s is just $\vec{s}$ viewed as a column vector, and the superscript T denotes the transpose. A simple way to understand this expression is to observe that the right-hand side of the first line of (A.8) is the inner product between the vector in the Hilbert-Schmidt space corresponding to the operator $|\psi\rangle\langle\psi|$ (which is just s up to some normalization factor), and the vector corresponding to $\Phi(|\psi\rangle\langle\psi|)$ (which is just $\mathcal{M}s$ up to some normalization factor). The final expression in (A.8) is then just this Hilbert-Schmidt inner product, with the factor of $\frac{1}{d}$ to take care of the normalization of the operator basis.

We can rewrite the expression in (A.8) for the fidelity using $\mathcal{M}_{\text{sym}} \equiv \frac{1}{2}(\mathcal{M} + \mathcal{M}^T)$, the symmetrized version of $\mathcal{M}$. Observe that $s^T \mathcal{M}_{\text{sym}}\ s = \frac{1}{2}(s^T \mathcal{M}\ s + (s^T \mathcal{M}\ s)^T) = s^T \mathcal{M}\ s$. Equation (A.8)

can hence be rewritten as

$$F^2\big(|\psi\rangle, \Phi(|\psi\rangle\langle\psi|\big) = s^T \mathcal{M}_{\text{sym}}\ s. \tag{A.9}$$

From this, we see that finding the worst-case fidelity is equivalent to the following minimization problem for the real, symmetric matrix $\mathcal{M}_{\text{sym}}$:

$$\text{minimize:} \quad s^T \mathcal{M}_{\text{sym}}\ s, \tag{A.10a}$$

$$\text{constraint:} \quad s \text{ corresponds to a pure state.} \tag{A.10b}$$

For $d > 2$, the constraint (A.10b) is difficult to write down. This constrained minimization problem is hence not simple for a general d.

For qubits ($d = 2$) however, the constraint equation *is* simple to write down. In this case, the operator basis can be chosen to be the Pauli basis $\{\sigma_0, \sigma_x, \sigma_y, \sigma_z\}$. Then, Eq. (A.7) corresponds to the Bloch sphere representation of a pure state, with the Bloch vector $\mathbf{s} \equiv (s_x, s_y, s_z)$ satisfying $\|\mathbf{s}\| = (s_x^2 + s_y^2 + s_z^2)^{1/2} = 1$. The constraint (A.10b) becomes

$$\text{constraint:} \quad s = (1, \mathbf{s}), \quad \text{with } \|\mathbf{s}\| = 1. \tag{A.10b$'$}$$

The constrained minimization problem can then be solved using the Lagrange multiplier method.

For the case of a CPTP qubit map that is also unital, the minimization problem can be further simplified. For any CPTP, unital Φ (arbitrary d), $\mathcal{M}$ takes the form

$$\mathcal{M} = \left(\begin{array}{c|ccc} 1 & 0 & \dots & 0 \\ \hline 0 & & & \\ \vdots & & \mathcal{T} & \\ 0 & & & \end{array}\right). \tag{A.11}$$

The first row comes from the fact that Φ is TP, since we have set $O_0 = \mathbb{I}$, and all O_α's for $\alpha > 0$ are traceless. The first column comes from the fact that Φ is unital. $\mathcal{T}$ is a $(d-1) \times (d-1)$ real matrix. Defining $\mathcal{T}_{\text{sym}} \equiv \frac{1}{2}(\mathcal{T} + \mathcal{T}^T)$, (A.9) can be written as

$$F^2\big(|\psi\rangle, \Phi(|\psi\rangle\langle\psi|\big) = \frac{1}{d}(1 + \mathbf{s}^T \mathcal{T}_{\text{sym}}\ \mathbf{s}).$$

This means that we can equivalently minimize $\mathbf{s}^T \mathcal{T}_{\text{sym}}\, \mathbf{s}$ instead of the original $s^T \mathcal{M}_{\text{sym}}\, s$ in (A.10a). Note that, for Φ with a Hermitian-closed Kraus set,[1] $\mathcal{T}$ is symmetric so that $\mathcal{T}_{\text{sym}} = \mathcal{T}$. This is indeed the case for $\Phi \equiv \mathcal{R}_P \circ \mathcal{E} \circ \mathcal{P} \sim \{PE_i^\dagger \mathcal{E}(P)^{-1/2} E_j^\dagger P\}$. For a qubit CPTP, unital Φ then, the constrained minimization problem, with the operator basis $\{O_\alpha\}$ chosen as the Pauli basis, becomes

$$\text{minimize:} \quad \mathbf{s}^T \mathcal{T}_{\text{sym}}\, \mathbf{s}, \tag{A.12a}$$

$$\text{constraint:} \quad \|\mathbf{s}\| = \sqrt{s_x^2 + s_y^2 + s_z^2} = 1. \tag{A.12b}$$

The constraint simply tells us to minimize the expectation value of $\mathcal{T}_{\text{sym}}$ with respect to all real *unit* vectors $\mathbf{s}$.

Now, since $\mathcal{T}_{\text{sym}}$ is real and symmetric, it can be diagonalized with an orthogonal matrix Q so that $\mathcal{T}_{\text{sym}} = Q^T \mathcal{T}_D Q$, where $\mathcal{T}_D$ is a real, diagonal matrix of eigenvalues of $\mathcal{T}_{\text{sym}}$. Then $s^T \mathcal{T}_{\text{sym}}\, s = (Qs)^T \mathcal{T}_D (Qs)$. Q, being orthogonal, preserves the length of the vector it acts on. The minimization problem (A.12) simply corresponds to minimizing the expectation value of $\mathcal{T}_D$ over all real unit vectors. As $\mathcal{T}_D$ is real and diagonal, this minimum expectation value is exactly the smallest eigenvalue of $\mathcal{T}_D$ (and hence of $\mathcal{T}_{\text{sym}}$), attained by the corresponding eigenvector normalized to unit length. Therefore, we see that the fidelity loss for a CPTP, unital qubit map Φ is given by

$$\eta_\Phi = 1 - \min_{|\psi\rangle \in \mathcal{C}} F^2(|\psi\rangle, \Phi(|\psi\rangle\langle\psi|) = \frac{1}{2}(1 - t_{\min}),$$

where $t_{\min}$ is the smallest eigenvalue of $\mathcal{T}_{\text{sym}}$ corresponding to the map Φ. Setting $\Phi = \mathcal{R}_P \circ \mathcal{E} \circ \mathcal{P}$ gives η_P.

[1] A set $\mathcal{K} \equiv \{K_i\}$ is Hermitian-closed if $K_i \in \mathcal{K}$ if and only if $K_i^\dagger \in \mathcal{K}$.

Appendix B

Constructing Maximally Commuting Classes of Clifford Generators

In (3.46) and (3.47) we gave examples of constructing $L = 3$ and $L = 4$ MUBs in dimension $d = 4$, such that they are cyclically permuted under the action of a unitary U that permutes the Clifford generators in $d = 4$. Here, we show by a general construction that it is always possible to construct L such classes in dimension $d = 2^n$, whenever $L|n$ and L is prime. We also outline a construction for $L = 2n + 1$ classes, given a unitary U that cycles through *all* $2n + 1$ Clifford generators, when $2n + 1$ is prime.

Given the $2n$ generators of the Clifford algebra in dimension $d = 2^n$, we consider the set

$$\mathcal{S} = \{\mathbb{I}, \Gamma_j, i\Gamma_j\Gamma_k, \Gamma_j\Gamma_k\Gamma_l, ..., i\Gamma_0\Gamma_1..\Gamma_{2n-1} \equiv \Gamma_{2n}\}. \tag{B.1}$$

To generate a set of $L \leq 2n + 1$ MUBs, we seek to group the elements of $\mathcal{S}$ into L classes of commuting operators, i.e., sets $\{\mathcal{C}_0, \mathcal{C}_1, \ldots, \mathcal{C}_{L-1} \mid \mathcal{C}_j \subset \mathcal{S} \setminus \{\mathbb{I}\}\}$ of cardinality $|\mathcal{C}_j| = d - 1$, such that,

(**P1**) The elements of $\mathcal{C}_j$ commute for all $0 \leq j \leq L - 1$,

(**P2**) The classs are all *mutually disjoint*, that is,

$$\mathcal{C}_j \cap \mathcal{C}_k = \emptyset \ \text{ for all } \ j \neq k, \tag{B.2}$$

(**P3**) The unitary U that cyclically permutes the generators $\Gamma_0, \ldots, \Gamma_{L-1}$, also permutes the corresponding classes by permuting products of operators appropriately.

To obtain such a set of classes, we first pick $d-1$ elements for the class $\mathcal{C}_0$ and then generate the rest of the classes by repeated application of U to the elements of $\mathcal{C}_0$. This automatically ensures property (**P3**). To ensure (**P1**) and (**P2**), the $d-1$ operators $\mathcal{C}_0 \equiv \{\mathcal{O}_1, \mathcal{O}_2, \ldots \mathcal{O}_{d-1}\}$ must satisfy the following:

(i) For any pair $\mathcal{O}_i, \mathcal{O}_j \in \mathcal{C}_0$, $[\mathcal{O}_i, \mathcal{O}_j] = 0$, and

(ii) The operators in $\mathcal{C}_0$ cycle through *mutually disjoint* sets of operators under the action of U.

To understand condition (ii) better, consider an operator $\mathcal{O}_i$ in $\mathcal{C}_0$. Then, by construction, $U^k(\mathcal{O}_i) \in \mathcal{C}_k$ for $0 \leq k \leq L-1$, assuming that we construct a total of L classes. In addition, property (ii) implies $U^k(\mathcal{O}_i) \notin \mathcal{C}_j$, for any $j \neq k$. In other words, given any two operators $\mathcal{O}_i, \mathcal{O}_j \in \mathcal{C}_0$ that cycle through the sets

$$S_i = \{U^k(\mathcal{O}_i) | 0 \leq k \leq L-1\}, \text{ and} \tag{B.3}$$

$$S_j = \{U^k(\mathcal{O}_j) | 0 \leq k \leq L-1\}\ , \tag{B.4}$$

respectively, under the action of U, property (ii) demands that $S_i \cap S_j = \emptyset$, for all $i \neq j$ and $i, j = 1, 2, \ldots, d-1$.

Finally, we note that no class can contain two generators Γ_j and Γ_k, since they do not commute. When forming the classes we hence ensure that each one contains exactly one generator Γ_j, which we refer to as the *singleton* Γ-operator of the class, as opposed to the rest of the elements which will be *products* of Γ-operators. The fact that each class can contain at most one singleton operator limits us to constructing a maximum of $2n+1$ such classes.

B.1 Mathematical Tools

Before proceeding to outline our construction, we establish some useful mathematical facts which will help motivate our algorithm for the construction of mutually disjoint classes. For the rest of the section, we will work with a set of p Γ-operators $\{\Gamma_0, \Gamma_2, \ldots, \Gamma_{p-1}\}$ that are cycled under the action of U, as follows,

$$U : \Gamma_0 \to \Gamma_1 \to \ldots \Gamma_{p-1} \to \Gamma_0, \tag{B.5}$$

In other words, we are given a set of Γ-operators whose *cycle-length* is p.

B.1.1 Length-2 Operators

First, we consider products of two Γ-operators of the form $\Gamma_i\Gamma_j$, which we call *length-2* operators. It is convenient to characterize such pairs in terms of the *spacing*—(S)—between the operators that constitute them. The spacing function S, for a given set of p operators, is simply defined as: $S(\Gamma_i\Gamma_j) = (j-i) \bmod p$. Then, the following holds:

Lemma B.1.1 (**Unique spacings imply nonintersecting cycles**). *The action of U on any length-2 operator $\Gamma_i\Gamma_j$ leaves its spacing function $S(.)$ invariant. Thus, length-2 operators that have unique spacings cycle through mutually disjoint sets of operators under the action of U.*

Proof. Recall, $U : \Gamma_i \rightarrow \Gamma_{(i+1) \bmod p}$. It clearly follows that

$$\begin{aligned} U : S(\Gamma_i\Gamma_j) &\rightarrow S\left(\Gamma_{(i+1) \bmod p}\Gamma_{(j+1) \bmod p}\right) \\ &= (j-i) \bmod p = S(\Gamma_i\Gamma_j). \end{aligned}$$

■

B.1.2 Higher-Length Operators

Similar to the length-2 operators, we refer to any product of ℓ Γ-operators as a *length-ℓ* operator. For operators of length higher than 2, it becomes convenient to refer to them using their corresponding index sets. For example, the operator $\Gamma_{i_1}\Gamma_{i_2}\ldots\Gamma_{i_\ell}$ will be simply denoted by the index set $(i_1, i_2, \ldots i_\ell)$. In the following lemma, we obtain a condition for any set of length-ℓ operators to cycle through mutually disjoint sets under the action of U.

Lemma B.1.2 (**Mutually disjoint cycles for length ℓ**). *Suppose the length-ℓ operators (for $3 \leq \ell \leq p-1$) that belong to the class $\mathcal{C}_0$ are such that they correspond to index sets $(i_1, i_2, \ldots, i_\ell)$ which sum to the same value*

$$i_1 + i_2 + \ldots + i_\ell = c_\ell \, mod \, p, \; \forall \; (i_1, i_2, \ldots, i_\ell) \in \mathcal{C}_0. \tag{B.6}$$

Then, no given index set of length ℓ can belong to more than one class, for prime values of p.

Proof. Given the operators $\{\Gamma_{i_1}\Gamma_{i_2}\ldots\Gamma_{i_\ell}\} \in \mathcal{C}_0$, such that the corresponding index sets $(i_1, i_2, \ldots, i_\ell)$

sum to

$$i_1 + i_2 + \ldots + i_\ell = c_\ell \bmod p,\ \forall (i_1, i_2, \ldots, i_\ell) \in \mathcal{C}_0. \tag{B.7}$$

Under the action of U, these index sets change to

$$\begin{aligned} (i_1, i_2, \ldots, i_\ell) &\rightarrow (i_1^{(1)}, i_2^{(1)}, \ldots, i_\ell^{(1)}) \\ &= (i_1 + 1, i_2 + 1, \ldots, i_\ell + 1) \bmod p\ . \end{aligned}$$

For any index set $(i_1^{(1)}, i_2^{(1)}, \ldots, i_\ell^{(1)}) \in \mathcal{C}_1$, the sum of the indices corresponding to the new operators $\{\Gamma_{i_1^{(1)}} \Gamma_{i_2^{(1)}} \ldots \Gamma_{i_\ell^{(1)}}\} \in \mathcal{C}_1$ becomes

$$i_1^{(1)} + i_2^{(1)} + \ldots + i_\ell^{(1)} = (c_\ell + \ell) \bmod p.$$

Proceeding similarly, the corresponding operators in the class $\mathcal{C}_k$ have index sets $(i_1^{(k)}, i_2^{(k)}, \ldots, i_\ell^{(k)})$ that sum to

$$i_1^{(k)} + i_2^{(k)} + \ldots + i_\ell^{(k)} = (c_\ell + k\ \ell) \bmod p, \tag{B.8}$$

for all $(i_1^{(k)}, i_2^{(k)}, \ldots, i_\ell^{(k)}) \in \mathcal{C}_k$. Thus, starting with a constraint on the length-ℓ operators in $\mathcal{C}_0$, we have obtained a constraint on the corresponding operators in a generic class $\mathcal{C}_k$.

Now, to arrive at a contradiction, suppose that an index set $(j_1, j_2, \ldots, j_\ell)$ whose indices $\{j_m\}_m$ take values from the set $\{0, 1, \ldots, p-1\}$, belongs to two different classes, $\mathcal{C}_k$ and $\mathcal{C}_k'$ (with $k \neq k'$). The constraint imposed by (B.8) implies

$$\begin{aligned} (c_\ell + k\ \ell) \bmod p &= (c_\ell + k'\ \ell) \bmod p \\ \Rightarrow (k - k')\ell \bmod p &= 0. \end{aligned} \tag{B.9}$$

Without loss of generality, let $k > k'$. Since we can form at most p classes, the difference $(k - k')$ can be at most $(p-1)$. Finally, since $\ell \leq p-1$, condition (B.9) cannot be satisfied for prime values of p. ■

Recall that to construct any p classes, we first construct the class $\mathcal{C}_0$, and then obtain the rest

by successive application of U. Therefore, the fact that any index set of a certain length ℓ cannot belong to more than one class implies that each length-ℓ operator in $\mathcal{C}_0$ cycles through a unique set of length-ℓ operators under U. In other words, the length-ℓ operators cycle through mutually disjoint sets, as desired.

Lemma B.1.2 thus provides us with a sufficient condition for the set of length-ℓ operators in $\mathcal{C}_0$ to cycle through mutually disjoint sets under U, given a set of Γ-operators whose cyclelength is primevalued. We only need to ensure that the length-ℓ operators in the first class that we construct, $\mathcal{C}_0$, correspond to index sets that *all* sum to the same value. This condition is of course subject to the constraint that the maximum allowed length for the operators in $\mathcal{C}_0$ (and by extension, in any class) is $p-1$.

B.2 Constructing 2n + 1 Prime Classes

To start with, we construct $L = 2n+1$ classes in dimension $d = 2^n$, when $2n+1$ is prime. This case is particulary easy, and illustrates how the results of the previous sections are used in our construction.

Theorem B.2.1 ($2n+1$ prime classes). *Let $\mathcal{G}^{(full)} = \{\Gamma_0, \ldots, \Gamma_{2n}\}$ denote the complete set of $(2n+1)$ Γ-operators, and let U be the unitary that cycles through all of them, that is,*

$$U \quad : \quad \Gamma_0 \to \Gamma_1 \ldots \Gamma_{2n-1} \to \Gamma_{2n} \to \Gamma_0 \ . \tag{B.10}$$

*If $2n+1$ is prime, then there exist $2n+1$ classes $\mathcal{C}_0, \mathcal{C}_1, \ldots, \mathcal{C}_{2n}$ satisfying properties **(P1)** through **(P3)**.*

Proof. We prove the existence of $2n+1$ classes by construction. We first outline an algorithm to pick $d-1$ operators that constitute the class $\mathcal{C}_0$. The remaining classes are easily obtained by the application of U to the elements of $\mathcal{C}_0$. Then, we make use of Lemmas B.1.1 and B.1.2 to prove that the classes obtained through our construction do satisfy the desired properties.

Algorithm

1. Pick one of the elements of $\mathcal{G}^{(\text{full})}$, Γ_0, as the singleton operator.

2. Pair up the remaining operators in $\mathcal{G}^{(\text{full})}$ to form $(n-1)$ length-2 operators which commute with Γ_0, as follows,

$$\mathcal{L}_2 = \{\, \Gamma_1\Gamma_{2n},\ \Gamma_2\Gamma_{2n-1},\ \ldots, \Gamma_{n-2}\Gamma_{n+3},\ \Gamma_{n-1}\Gamma_{n+2}\,\},$$

where $\mathcal{L}_2$ denotes the set of length-2 operators in $\mathcal{C}_1$. Since we have left out the pair $\Gamma_n\Gamma_{n+1}$ in the middle, we get, $|\mathcal{L}_2| = n-1$.

3. Form higher-length operators that commute with $\mathcal{L}_2 \cup \{\Gamma_0\}$, by combining Γ_0 with appropriate combinations of the length-2 operators. Any operator of even length $\ell = 2j$ is created by combining i pairs in $\mathcal{L}_2$. And any operator of odd length $\ell = 2j+1$ is created by appending Γ_0 to a length-$2j$ operator.

   Denoting the sets of length-3 operators as $\mathcal{L}_3$, length-4 operators as $\mathcal{L}_4$, and in general, the set of length-i operators as $\mathcal{L}_i$, we have,

$$\begin{aligned}
|\mathcal{L}_3| &= |\mathcal{L}_2| = n-1, \\
|\mathcal{L}_4| &= \binom{n-1}{2},\ |\mathcal{L}_5| = |\mathcal{L}_4|, \\
|\mathcal{L}_6| &= \binom{n-1}{3},\ |\mathcal{L}_7| = |\mathcal{L}_6|, \\
&\vdots \qquad \vdots \\
|\mathcal{L}_{2n-2}| &= \binom{n-1}{n-1} = 1,\ |\mathcal{L}_{2n-1}| = |\mathcal{L}_{2n-2}|.
\end{aligned}$$

Putting together the operators from steps (1), (2), and (3) we get the desired cardinality for

the class $\mathcal{C}_0$ as follows:

$$\begin{aligned}
|\mathcal{C}_0| &= 1 + (n-1) + \sum_{i=3}^{2n} |\mathcal{L}_i| \\
&= 1 + 2(n-1) + 2\begin{pmatrix} n-1 \\ 2 \end{pmatrix} + 2\begin{pmatrix} n-1 \\ 3 \end{pmatrix} \\
&\quad +\ldots + 2\begin{pmatrix} n-1 \\ n-1 \end{pmatrix} \\
&= 2\sum_{i=0}^{n-1}\begin{pmatrix} n-1 \\ i \end{pmatrix} - 1 = 2(2^{n-1}) - 1 \\
&= 2^n - 1 = d - 1. \qquad \text{(B.11)}
\end{aligned}$$

The rest of the classes are generated by successive applications of the unitary U to the elements of $\mathcal{C}_0$, so that $U : \mathcal{C}_i \rightarrow \mathcal{C}_{(i+1) \mod 2n+1}$.

It is easy to see that the elements of each class satisfy property (**P1**) above—the different length operators have been chosen so as to ensure that they all commute with each other. Similarly, by construction, they satisfy property **(P3)**. It only remains to prove property (**P2**), that the classes are all mutually disjoint.

The elements of $\mathcal{L}_2$ correspond to the following set of spacings,

$$S(\mathcal{L}_2) \equiv \{2n-1, 2n-3, \ldots, 5, 3\},$$

which are all distinct. So by Lemma B.1.1, the elements of $\mathcal{L}_2$ cycle through mutually disjoint sets of length-2 operators.

For higher-length operators, we first show that our construction meets the conditions of Lemma B.1.2. For the class $\mathcal{C}_0$, the elements of $\mathcal{L}_2$ correspond to index sets that satisfy

$$\mathcal{L}_2(\mathcal{C}_0) = \{(i_1, i_2) |\, i_1 + i_2 = 0 \bmod (2n+1)\}.$$

The length-2 operators of a generic class $\mathcal{C}_k$ similarly satisfy

$$\mathcal{L}_2(\mathcal{C}_k) = \{(i_1, i_2) |\, i_1 + i_2 = 2k \bmod (2n+1)\}. \tag{B.12}$$

Since higher-length operators are essentially combinations of length-2 operators and the singleton operator, conditions similar to (B.12) hold for higher-length index sets as well. Since operators of even length $\ell = 2j$ contain j pairs from $\mathcal{L}_2$, the corresponding index sets in $\mathcal{C}_0$ satisfy

$$\begin{aligned} i_1 + i_2 + \ldots + i_{2j} &= 0 \bmod (2n+1), \\ &\forall \quad (i_1, i_2, \ldots, i_{2j}) \in \mathcal{C}_0. \end{aligned}$$

Similarly, since the odd-length operators have Γ_0 appended to the even-length operators, the index sets of length $\ell = 2j+1$ in $\mathcal{C}_0$ satisfy,

$$\begin{aligned} i_1 + i_2 + \ldots + i_{2j+1} &= 0 \bmod (2n+1), \\ &\forall \quad (i_1, i_2, \ldots, i_{2j+1}) \in \mathcal{C}_0. \end{aligned}$$

To sum up, for any $3 \leq \ell \leq 2n$, our construction ensures that index sets of length-ℓ belonging to $\mathcal{C}_0$ sum to the same value. The conditions of Lemma B.1.2 are therefore satisfied, with the quantity c_ℓ in (B.6) taking the value $c_\ell = 0$, for all $\ell = 3, \ldots, 2n$. Now, we can simply evoke Lemma B.1.2 to conclude that, when $2n+1$ is prime, the higher-length operators in $\mathcal{C}_0$ cycle through mutually disjoint sets of operators. ■

B.3 Constructing L|n Classes for Prime Values of L

Next, we show that it is possible to obtain an arrangement of operators into L classes in dimension 2^n, when L is prime and $L|n$, such that the unitary U that cyclically permutes L of the Γ-operators also permutes the corresponding classes.

Theorem B.3.1 (L|n classes for prime L). *Suppose U is a unitary that cycles through sets of L operators from the set $\mathcal{G}^{(full)} \setminus \{\Gamma_{2n}\}$ in dimension 2^n, where L is prime and $L|n$. Then there exist L classes $\mathcal{C}_0, \mathcal{C}_1, \ldots, \mathcal{C}_{L-1}$ that satisfy properties **(P1)** through **(P3)**.*

Proof: Note that since $L|n$ we have $n = rL$ for some positive integer r. The set of $2n$ Clifford generators $\Gamma_0, \Gamma_1, \ldots, \Gamma_{2n-1}$ can then be partitioned into $2r$ sets as follows.

$$\begin{aligned}
\mathcal{G}^{(0)} &= \{\Gamma_0, \Gamma_1, \ldots, \Gamma_{L-1}\}, \\
\mathcal{G}^{(1)} &= \{\Gamma_L, \Gamma_{L+1}, \ldots, \Gamma_{2L-1}\}, \\
\vdots &\quad \vdots \\
\mathcal{G}^{(2r-1)} &= \{\Gamma_{(2r-1)L}, \Gamma_{(2r-1)L+1}, \ldots, \Gamma_{2n-1}\}.
\end{aligned}$$

Without loss of generality, we can assume that the unitary U is so constructed that it cyclically permutes the L operators within each set, as follows.

$$\begin{aligned}
U \;:\; & \Gamma_0 \to \Gamma_1 \to \ldots \to \Gamma_{L-1} \to \Gamma_0, \\
& \Gamma_L \to \ldots \to \Gamma_{2L-1} \to \Gamma_L, \\
& \vdots \\
& \Gamma_{(2r-1)L} \to \ldots \to \Gamma_{2n-1} \to \Gamma_{(2r-1)L}.
\end{aligned}$$

Once again, we begin with an algorithm for picking $d-1$ elements for the class $\mathcal{C}_0$. The algorithm closely follows the one outlined in the previous section, barring minor modifications.

Algorithm

1. The "middle" element from $\mathcal{G}^{(1)}$, $\Gamma_{(L-1)/2}$, is picked as the singleton element of $\mathcal{C}_0$.
2. The $(n-1)$ length-2 operators which commute with $\Gamma_{(L-1)/2}$ are picked as follows.
   (a) $\frac{L-3}{2}$ pairs are picked from $\mathcal{G}^{(0)} \setminus \{\Gamma_{(L-1)/2}\}$

$$\mathcal{L}_2^{(0)} = \{\, \Gamma_1\Gamma_{L-1},\, \Gamma_2\Gamma_{L-2}, \ldots, \Gamma_{(L-3)/2}\Gamma_{(L+3)/2} \,\},$$

   leaving Γ_0 and $\Gamma_{(L+1)/2}$ unused.

(b) $\frac{L-1}{2}$ pairs are picked from each of the sets $\mathcal{G}^{(1)}$ through $\mathcal{G}^{(2r-1)}$,

$$\begin{aligned}
\mathcal{L}_2^{(1)} &= \{\, \Gamma_{L+1}\Gamma_{2L-1},\ \Gamma_{L+2}\Gamma_{2L-2},\ \ldots, \\
&\qquad \ldots,\ \Gamma_{L+(L-1)/2}\Gamma_{L+(L+1)/2}\,\}, \\
&\vdots \qquad \vdots \\
\mathcal{L}_2^{(2r-1)} &= \{\, \Gamma_{(2r-1)L+1}\Gamma_{2n-1},\ \Gamma_{(2r-1)L+2}\Gamma_{2n-2},\ \ldots, \\
&\qquad \Gamma_{(2r-1)L+(L-1)/2}\Gamma_{(2r-1)L+(L+1)/2}\,\},
\end{aligned}$$

leaving the first operator in each set unused.

(c) Finally, the unused Γ-operators from different sets are put together as specified below, to get the remaining r length-2 operators:

$$\mathcal{L}_2^{(2r)} = \{\, \Gamma_0\Gamma_L,\ \Gamma_{2L}\Gamma_{3L}, \ldots, \Gamma_{(2r-2)L}\Gamma_{(2r-1)L}\,\}.$$

The set of length-2 operators is then given by

$$\mathcal{L}_2 = \mathcal{L}_2^{(0)} \cup \mathcal{L}_2^{(1)} \ldots \cup \mathcal{L}_2^{(2r-1)} \cup \mathcal{L}_2^{(2r)},$$

which gives $|\mathcal{L}_2| = \frac{L-3}{2} + (2r-1)\left(\frac{L-1}{2}\right) + r = rL - \frac{2r-2}{2} + r = n-1$.

3. Higher-length operators that commute with $\Gamma_{(L-1)/2}$ and $\mathcal{L}_2$ are then chosen by combining $\Gamma_{(L-1)/2}$ with appropriate combinations of the length-2 operators. As before, any even-length operator of length $\ell = 2i$ is obtained by combining i length-2 operators from $\mathcal{L}_2$. Any operator of odd-length $\ell = 2i+1$, is created by appending $\Gamma_{(L-1)/2}$ to a length-$2i$ operator.

Putting together all the operators created in Steps [1]-[3], we get the desired cardinality for the class (see (B.11)), that is, $|\mathcal{C}_0| = 2^n - 1$.

Proof of properties (P1) through (P3): The different length operators have again been picked in such a way as to ensure that they all commute with each other. Since the remaining $L-1$ classes are generated by successive applications of the unitary U to the elements of $\mathcal{C}_0$, we have $U : \mathcal{C}_i \to \mathcal{C}_{(i+1) \bmod L}$. Thus **(P1)** and **(P3)** are satisfied. It remains to prove that the classes

constructed here also satisfy property (**P2**).

As in the earlier case of $2n+1$ classes, the operators in each of the sets $\{\mathcal{L}_2^{(0)}, \mathcal{L}_2^{(1)}, \ldots, \mathcal{L}_2^{(2r-1)}\}$ correspond to unique values of the spacing function:

$$S(\mathcal{L}_2^{(i)}) \equiv \{L-2, L-4, \ldots, 1\}, \forall i \in [0, 2r-1].$$

This guarantees by Lemma B.1.1 that these operators cycle through mutually disjoint sets under U. Since the operators in $\mathcal{L}_2^{(2r)}$ are formed by combining Γ-operators from different sets $\mathcal{G}^{(i)}$, each of them cycles through a different set of operators under U. Thus we see that all the length-2 operators in $\mathcal{C}_0$ cycle through mutually disjoint sets.

Before we proceed to discuss the higher-length operators, we make one further observation about the length-2 operators. The operators in $\mathcal{L}_2$ correspond to index sets which satisfy

$$\mathcal{L}_2(\mathcal{C}_1) = \{\Gamma_{i_1}\Gamma_{i_2} |\, i_1 + i_2 = 0 \bmod L\}. \tag{B.13}$$

In particular, the length-2 operators in the set $\mathcal{L}^{(2r)}$ have been picked carefully so as to ensure that the above constraint is satisfied. In fact, this was the rationale behind leaving out the first operator in each of the sets $\mathcal{G}^{(i)}$ while choosing the corresponding length-2 elements in $\mathcal{L}_2^{(i)}$.

The higher-length operators in $\mathcal{C}_0$ can be of two types:

(a) Products of Γ-operators from a single set $\mathcal{G}^{(i)}$ alone, and,

(b) Products of Γ-operators from more than one set.

Since an operator of type (a) cannot cycle into one of type (b) under the action of U, these two cases can be examined separately.

Operators of type (a): The maximum length that an operator of type (a) can have, as per our construction, is $L-1$. We have ensured this by leaving at least one operator of each of the sets $\mathcal{G}^{(i)}$ unused in constructing the length-2 operators. Furthermore, the constraint in (B.13) implies that the index sets corresponding to such higher length operators in $\mathcal{C}_0$, sum to the same value modulo L. More precisely, any even-length index set of length $\ell = 2j$, where the indices are all drawn from

a given set $\mathcal{G}^{(i)}$, satisfies

$$\begin{aligned} i_1 + i_2 + \ldots + i_l &= 0 \bmod L, \\ &\forall \quad (i_1, i_2, \ldots, i_l) \in \mathcal{C}_0. \end{aligned} \tag{B.14}$$

And any index set of odd length $\ell = 2j + 1$ satisfies

$$\begin{aligned} i_1 + i_2 + \ldots + i_l &= \left(\frac{L-1}{2}\right) \bmod L, \\ &\forall \quad (i_1, i_2, \ldots, i_l) \in \mathcal{C}_0. \end{aligned} \tag{B.15}$$

Then, invoking Lemma B.1.2 with $c_\ell = 0$ for even values of ℓ and $c_\ell = (L-1)/2$ for odd values of ℓ, we see that no operator of type (a) can belong to more than one class, for prime values of L.

Operators of type (b): An operator of type (b) is a product of operators from smaller sets $\mathcal{K}_j \subseteq \mathcal{G}^{(j)}$. Consider a length-$\ell$ operator, $\mathcal{O}$ which comprises ℓ_0 Γ-operators from $\mathcal{G}^{(0)}$, ℓ_1 operators from $\mathcal{G}^{(1)}$, and in general, ℓ_i from the set $\mathcal{G}^{(i)}$.

$$\mathcal{O} = \underbrace{\Gamma_{i_1} \ldots \Gamma_{i_{\ell_0}}}_{\mathcal{K}_0 \subseteq \mathcal{G}^{(0)}} \underbrace{\Gamma_{j_1} \ldots \Gamma_{j_{\ell_1}}}_{\mathcal{K}_1 \subseteq \mathcal{G}^{(1)}} \ldots \underbrace{\Gamma_{k_1} \ldots \Gamma_{k_{\ell_{2r-1}}}}_{\mathcal{K}_{2r-1} \subseteq \mathcal{G}^{(2r-1)}}$$

Note that by our construction, the operator $\mathcal{O}$ exists in more than one class if and only if, for all $\mathcal{K}_j$ the product of all operators in $\mathcal{K}_j$ also belongs to more than one class. In what follows, we show that our construction ensures that this is not possible. In particular, given a set of length-ℓ operators in $\mathcal{C}_0$ which can be broken down into smaller sets as described above, we will show that there exists at least one set $\mathcal{K}_j$ in every such length-ℓ operator $\mathcal{O}$, such that the products of operators in $\mathcal{K}_j$ corresponding to different length-ℓ operators cycle through mutually disjoint sets, as defined earlier.

Note the following two facts about the subsets $\mathcal{K}_j$. First, our construction ensures that any subset $\mathcal{K}_j \subseteq \mathcal{G}^{(j)}$ of a given size ℓ_j, satisfies either (B.14) or (B.15) depending on ℓ_j being even or odd. Second, note that the maximum size of these subsets is $\ell_j \leq L$. However, in order to invoke Lemma B.1.2, we still require ℓ_j to be strictly less than L. We thus need to show that every

length-ℓ operator must have at least one subset $\mathcal{K}_j$ of size $\ell_j < L$.

Suppose there exists a length-ℓ operator such that every subset is of size L. Then, such an operator has to be of length

$$\ell = \ell_0 + \ell_1 + \ldots + \ell_{2r-1} = 2rL = 2n. \tag{B.16}$$

However the maximum value of ℓ in our construction is $2n - 1$, implying that at least one of the $2r$ subsets must be of a size strictly smaller than L. And, for such a subset of size less than L, constraints (B.14) and (B.15) ensure that the same subset cannot be found in more than one class, provided L is prime. ■

Bibliography

[1] P. Aliferis, F. Brito, D. P. DiVincenzo, J. Preskill, M. Steffen, and B. M. Terhal. Fault-tolerant computing with biased-noise superconducting qubits: a case study. *New Journal of Physics*, 11:013061, 2009.

[2] P. Aliferis, D. Gottesman, and J. Preskill. Quantum accuracy threshold for concatenated distance-3 codes. *Quantum Information and Computation*, 6:97, 2006.

[3] P. Aliferis and J. Preskill. Fault-tolerant quantum computation against biased noise. *Physical Review A*, 78(5):052331, 2008.

[4] A. Ambainis. Limits on entropic uncertainty relations for 3 and more MUBs. *e-print arXiv:0909.3720*, 2009.

[5] D. M. Appleby. Properties of the extended clifford group with applications to sic-povms and mubs. e-print arXiv:0909.5233, 2009.

[6] M. Ballester and S. Wehner. Entropic uncertainty relations and locking: tight bounds for mutually unbiased bases. *Physical Review A*, 75:022319, 2007.

[7] S. Bandyopadhyay, P.O. Boykin, V.P. Roychowdhury, and F. Vatan. A new proof for the existence of mutually unbiased bases. *Algorithmica*, 34(4):512–528, 2002.

[8] H. Barnum and E. Knill. Reversing quantum dynamics with near-optimal quantum and classical fidelity. *Journal of Mathematical Physics*, 43:2097, 2002.

[9] D. Beaver. Precomputing oblivious transfer. In *Advances in Cryptology—EUROCRYPT '95*, volume 963 of *Lecture Notes in Computer Science*, pages 97–109. Springer-Verlag, 1995.

[10] W. Beckner. Inequalities in fourier analysis. *Annals of Mathematics*, 102(1):159–182, 1975.

[11] C. H. Bennett and G. Brassard. Quantum cryptography: Public key distribution and coin tossing. In *Proceedings of the IEEE International Conference on Computers, Systems and Signal Processing*, pages 175–179, 1984.

[12] C. H. Bennett and G. Brassard. The dawn of a new era for quantum cryptography: The experimental prototype is working. *Sigact News*, 20(4):78–82, 1989.

[13] C. H. Bennett, G. Brassard, S. Breidbart, and S. Wiesner. Quantum cryptography, or unforgeable subway tokens. In *Advances in Cryptology CRYPTO '82*, pages 267–275, 1982.

[14] C. H. Bennett, G. Brassard, C. Crépeau, and H. Skubiszewska. Practical quantum oblivious transfer. In *Advances in Cryptology—CRYPTO '91*, volume 576 of *Lecture Notes in Computer Science*, pages 351–366. Springer, 1992.

[15] C. H. Bennett, D. P. DiVincenzo, J. A. Smolin, and W. K. Wootters. Mixed state entanglement and quantum error correction. *Phys. Rev. A*, 54:3824, 1996.

[16] C. Bény. Conditions for approximate quantum error correction. *The 4th Workshop on Theory of Quantum Computation, Communication, and Cryptography*, 2009.

[17] C. Bény and O. Oreshkov. General conditions for approximate quantum error correction and near-optimal recovery channels. *Physical Review Letters*, 104(12):120501, 2010.

[18] I. Bialynicki-Birula and J. Mycielski. Uncertainty relations for information entropy in wave mechanics. *Communications in Mathematical Physics*, 44:129–132, 1975.

[19] R. Blume-Kohout, H. K. Ng, D. Poulin, and L. Viola. Characterizing the structure of preserved information in quantum processes. *Physical Review Letters*, 100:030501, 2008.

[20] F. Buscemi. Entanglement measures and approximate quantum error correction. *Physical Review A*, 77:012309, 2008.

[21] C. Cachin and U. M. Maurer. Unconditional security against memory-bounded adversaries. In *Proceedings of CRYPTO 1997*, pages 292–306, 1997.

[22] A. R. Calderbank and P. W. Shor. Good quantum error-correcting codes exist. *Physical Review A*, 54:1098–1106, 1995. quant-ph/9512032.

[23] A. Casaccino, E. F. Galvao, and S. Severini. Extrema of discrete wigner functions and applications. *Physical Review A*, 78:022310, 2008.

[24] H. F. Chau. Unconditionally secure quantum key distribution in higher dimensions. *IEEE Transactions on Information Theory*, 51:1451, 2005.

[25] M. D. Choi. Completely positive linear maps on complex matrices. *Linear algebra and its applications*, 10(3):285–290, 1975.

[26] I. L. Chuang, D. W. Leung, and Y. Yamamoto. Bosonic quantum codes for amplitude damping. *Physical Review A*, 56(2):1114, 1997.

[27] P. M. Cohn. *Basic Algebra – Groups, Rings and Fields*. Springer, 2003.

[28] T. M. Cover and J. A. Thomas. *Elements of Information Theory*. Wiley, 1991.

[29] C. Crépeau. Quantum oblivious transfer. *Journal of Modern Optics*, 41(12):2455–2466, 1994.

[30] I. B. Damgård, S. Fehr, R. Renner, L. Salvail, and C. Schaffner. A tight high-order entropic quantum uncertainty relation with applications. In *Advances in Cryptology—CRYPTO '07*, volume 4622 of *Lecture Notes in Computer Science*, pages 360–378. Springer-Verlag, 2007.

[31] I. B. Damgård, S. Fehr, L. Salvail, and C. Schaffner. Cryptography in the Bounded-Quantum-Storage Model. In *Proceedings of 46th IEEE FOCS*, pages 449–458, 2005.

[32] I. B. Damgård, S. Fehr, L. Salvail, and C. Schaffner. Secure identification and QKD in the bounded-quantum-storage model. In *Advances in Cryptology—CRYPTO '07*, volume 4622 of *Lecture Notes in Computer Science*, pages 342–359. Springer-Verlag, 2007.

[33] G. D'Ariano, D. Kretschmann, D. Schlingemann, and R. F. Werner. Quantum bit commitment revisited: the possible and the impossible. quant-ph/0605224, 2006.

[34] D. Deutsch. Uncertainty in quantum measurements. *Physical Review Letters*, 50:631–633, 1983.

[35] D. Deutsch, A. Ekert, R. Jozsa, C. Macchiavello, S. Popescu, and A. Sanpera. Quantum privacy amplification and the security of quantum cryptography over noisy channels. *Physical Review Letters*, 77(13):2818–2821, 1996.

[36] K. Dietz. Generalized bloch spheres for m-qubit states. *Journal of Physics A: Mathematical and General*, 36(6):1433–1447, 2006.

[37] Y. Z. Ding, D. Harnik, A. Rosen, and R. Shaltiel. Constant round oblivious transfer in the bounded-storage model. In *Proceedings of TCC*, volume 2951 of *Lecture Notes in Computer Science*, pages 446–472, 2004.

[38] D. P. DiVincenzo, M. Horodecki, D. Leung, J. Smolin, and B. Terhal. Locking classical correlation in quantum states. *Physical Review Letters*, 92(067902), 2004.

[39] D. P. DiVincenzo and P. W. Shor. Fault-tolerant error correction with efficient quantum codes. *Physical Review Letters*, 77(15):3260–3263, 1996.

[40] T. Durt, B. G. Englert, I. Bengtsson, and K. Życzkowski. On mutually unbiased bases. *International Journal of Quantum Information*, 8:535–640, 2010.

[41] S. Dziembowski and U. Maurer. On generating the initial key in the bounded-storage model. In *Proceedings of EUROCRYPT*, Lecture Notes in Computer Science, pages 126–137, 2004.

[42] A. Ekert. Quantum cryptography based on Bell's theorem. *Physical Review Letters*, 67:661–663, 1991.

[43] A. Ekert and C. Macchiavello. Error correction in quantum communication. *Physical Review Letters*, 77:2585, 1996.

[44] S. Fehr and C. Schaffner. Composing quantum protocols in a classical environment. e-print arXiv:0804.1059, 2008.

[45] A. S. Fletcher. Channel-adapted quantum error correction. 2007. PhD Thesis Massachusetts Institute of Technology, e-print arXiv:0706.3400.

[46] A. S. Fletcher, P. W. Shor, and M. Z. Win. Optimum quantum error recovery using semidefinite programming. *Physical Review A*, 75:021338, 2007.

[47] A. S. Fletcher, P. W. Shor, and M. Z. Win. Channel-adapted quantum error correction for the amplitude damping channel. *IEEE Transactions on Information Theory*, 54:5705, 2008.

[48] A. S. Fletcher, P. W. Shor, and M. Z. Win. Structured near-optimal channel-adapted quantum error correction. *Physical Review A*, 77:012320, 2008.

[49] C. W. Gardiner and P. Zoller. *Quantum Noise*. Springer, 1991.

[50] K. S. Gibbons, M. J. Hoffman, and W. K. Wootters. Discrete phase space based on finite fields. *Physical Review A*, 062101, 2004.

[51] D. Gottesman. Stabilizer codes and quantum error correction. 1997. PhD Thesis California Institute of Technology, e-print arXiv quant-ph/9705052.

[52] M. Grassl. On SIC-POVMs and MUBs in dimension 6. In *Proceedings ERATO Conference on Quantum Information Science*, pages 60–61, 2004.

[53] O. Guehne. Characterizing entanglement via uncertainty relations. *Physical Review Letters*, 92:117903, 2004.

[54] P. Hayden. Personal communication, 2010.

[55] P. Hayden, R. Jozsa, D. Petz, and A. Winter. Structure of states which satisfy strong subadditivity of quantum entropy with equality. *Communications in Mathematical Physics*, 246(2):359–374, 2004.

[56] P. Hayden, D. Leung, P. Shor, and A. Winter. Randomizing quantum states: Constructions and applications. *Communications in Mathematical Physics*, 250(2):371–391, 2004.

[57] W. Heisenberg. Über den anschaulichen inhalt der quantentheoretischen kinematik und mechanik. *Zeitschrift für Physik*, 43:172–198, 1927.

[58] I. I. Hirschmann. A note on entropy. *American Journal of Mathematics*.

[59] P. A. Hiskett, D. Rosenberg, C. G. Peterson, R. J. Hughes, S. Nam, A. E. Lita, A. J. Miller, and J. E. Nordholt. Long-distance quantum key distribution in optical fibre. *New Journal of Physics*, 8:193, 2006.

[60] W. Hoeffding. Probability inequalities for sums of bounded random variables. *Journal of the American Statistical Association*, 58(301):13–30, March 1963.

[61] I. D. Ivanovic. An inequality for the sum of entropies of unbiased quantum measurements. *Journal of Physics A: Math General*, 25(7):363–364, 1992.

[62] P. Jordan and E. Wigner. Über das paulische äquivalenzverbot. *Zeitschrift für Physik*, 47:631, 1928.

[63] K. Khodjasteh and D. A. Lidar. Fault-tolerant quantum dynamical decoupling. *Physical Review Letters*, 95(18):180501, 2005.

[64] J. Kilian. Founding cryptography on oblivious transfer. In *Proceedings of 20th ACM STOC*, pages 20–31, 1988.

[65] A.Y. Kitaev. Fault-tolerant quantum computation by anyons. *e-print arXiv quant-ph/9707021*, 1997.

[66] A.Y. Kitaev. Quantum computations: algorithms and error correction. *Russian Mathematical Surveys*, 52(6):1191–1249, 1997.

[67] F. Kittneh. Norm inequalities for certain operator sums. *Journal of Functional Analysis*, 143:337–348, 1997.

[68] A. Klappenecker and M. Rötteler. Constructions of mutually unbiased bases. In *International Conference on Finite Fields and Applications (Fq7)*, volume 2948 of *Lecture Notes in Computer Science*, pages 137–144. Springer, 2004.

[69] R. Klesse. Approximate quantum error correction, random codes, and quantum channel capacity. *Physical Review A*, 75:062315, 2007.

[70] E. Knill and R. Laflamme. A theory of quantum error-correcting codes. *Physical Review A*, 55:900, 1997.

[71] E. Knill, R. Laflamme, and W. H. Zurek. Resilient quantum computation. *Science*, 279(5349):342.

[72] M. Koashi. Simple security proof of quantum key distribution via uncertainty principle. e-print arXiv quant-ph/0505108, 2005.

[73] R. König and R. Renner. Sampling of min-entropy relative to quantum knowledge. e-print arXiv:0712.4291, 2007.

[74] R. König, R. Renner, and C. Schaffner. The operational meaning of min- and max-entropy. e-print arXiv:0807.1338, 2008.

[75] R. König and S. Wehner. A strong converse for classical channel coding using entangled inputs. e-print arXiv:0903.2838, 2009.

[76] R. König, S. Wehner, and J. Wullschleger. Unconditional security from noisy quantum storage. e-print arXiv:0906.1030v3, 2009.

[77] R. L. Kosut and D. A. Lidar. Quantum error correction via convex optimization. *Quantum Information Processing*, 8:443, 2009.

[78] R. L. Kosut, A. Shabani, and D. A. Lidar. Robust quantum error correction via convex optimization. *Physics Review Letters*, 100:020502, 2008.

[79] K. Kraus. General state changes in quantum theory. *Annals of Physics*, 64(2):311–335, 1971.

[80] K. Kraus. *States, Effects and Operations: Fundamental Notions of Quantum Theory (Lecture Notes in Physics vol 190)*. Berlin: Springer, 1983.

[81] K. Kraus. Complementary observables and uncertainty relations. *Physical Review D*, 35(10):3070–3075, 1987.

[82] D. W. Kribs, R. Laflamme, and D. Poulin. Unified and generalized approach to quantum error correction. *Physical Review Letters*, 94:180501, 2005.

[83] A. K. Kwasniewski. Clifford and Grassman algebras—old and new. *Journal of Mathematical Physics*, 26:2234, 1985.

[84] R. Laflamme, C. Miquel, J. P. Paz, and W. H. Zurek. Perfect quantum error correcting code. *Physical Review Letters*, 77(1):198–201, 1996.

[85] D. W. Leung, M. A. Nielsen, I. L. Chuang, and Y. Yamamoto. Approximate quantum error correction can lead to better codes. *Physical Review A*, 56:2567, 1997.

[86] H-K. Lo. Insecurity of quantum secure computations. *Physical Review A*, 56:1154, 1997.

[87] H-K. Lo and H. F. Chau. Unconditional security of quantum key distribution over arbitrarily long distances. *Science*, 283:2050–2056, 1999.

[88] P. Lounesto. *Clifford Algebras and Spinors*. Cambridge University Press, 2001.

[89] H. Maassen and J. Uffink. Generalized entropic uncertainty relations. *Physical Review Letters*, 60(1103), 1988.

[90] P. Mandayam and D. Poulin. *unpublished*, 2009.

[91] P. Mandayam and S. Wehner. Achieving the physical limits of the bounded-storage model. *Physical Review A*, 83(2):022329, 2011.

[92] P. Mandayam, S. Wehner, and N. Balachandran. A transform of complementary aspects with applications to entropic uncertainty relations. *Journal of Mathematical Physics*, 51:082201, 2010.

[93] U. Maurer. A provably-secure strongly-randomized cipher. In *Advances in Cryptology—EUROCRYPT*, volume 473 of *Lecture Notes in Computer Science*, pages 361–373, 1990.

[94] U. Maurer. Conditionally perfect secrecy and a provably-secure randomized cipher. *Journal of Cryptology*, 5(1):53–66, 1992.

[95] D. Mayers. Quantum key distribution and string oblivious transfer in noisy channels. In *Proceedings of Advances in Cryptology—CRYPTO '96*, pages 343–357, 1996.

[96] D. Mayers. Unconditionally secure quantum bit commitment is impossible. *Physical Review Letters*, 78:3414–3417, 1997.

[97] H. K. Ng and P. Mandayam. Simple approach to approximate quantum error correction based on the transpose channel. *Physical Review A*, 81(6):62342, 2010.

[98] M. Nielsen and I. Chuang. *Quantum Computation and Quantum Information*. Cambridge University Press, 2000.

[99] M. Ohaya and D. Petz. *Quantum Entropy and Its use*. Springer-Verlag, Berlin, Heidelberg, 1993.

[100] R. Ostrovsky, R. Venkatesan, and M. Yung. Fair games against an all-powerful adversary. In *AMS DIMACS Series in Discrete Mathematics and Theoretical Computer Science*, pages 155–169, 1991.

[101] D. Petz. Monotonicity of quantum relative entropy revisited. *Reviews in Mathematical Physics*, 15:79, 2003.

[102] J. Preskill. Fault-tolerant quantum computation. In H-K. Lo, S. Popescu, and T. P. Spiller, editors, *Introduction to Quantum Computation*. 1998.

[103] J. Preskill. Lecture notes for a course on quantum computation. Unpublished. Available at http://www.theory.caltech.edu/people/preskill/ph229/, 1998–1999.

[104] M. Rabin. How to exchange secrets by oblivious transfer. Technical report, Aiken Computer Laboratory, Harvard University, 1981. Technical Report TR-81.

[105] M. Reimpell and R. F. Werner. Iterative optimization of quantum error correcting codes. *Physical Review Letters*, 94:080501, 2005.

[106] J. M. Renes and J.-C. Boileau. Physical underpinnings of privacy. *Physical Review A*, 78:032335, 2008.

[107] R. Renner. Security of quantum key distribution. 2005. PhD Thesis ETH Zurich, e-print arXiv: quant-ph/0512258.

[108] R. Renner and R. König. Universally composable privacy amplification against quantum adversaries. In *Proceedings of TCC 2005*, volume 3378 of *Lecture Notes in Computer Science*, pages 407–425. Springer, 2005.

[109] R. Renner and S. Wolf. Smooth Rényi entropy and applications. In *Proceedings of ISIT 2004*, page 233. IEEE, 2004.

[110] A. Rényi. On measures of information and entropy. In *Proceedings of the 4th Berkeley Symposium on Mathematics, Statistics and Probability*, pages 547–561, 1960.

[111] H.P. Robertson. The uncertainty principle. *Physical Review*, 34:163–164, 1929.

[112] J. Sanchez. Entropic uncertainty and certainty relations for complementary observables. *Physics Letters A*, 173:233–239, 1993.

[113] G. Savvides. Interactive hashing and reductions between oblivious transfer variants. 2007. PhD Thesis McGill University, Montreal.

[114] C. Schaffner. Cryptography in the bounded-quantum storage model. 2007. PhD Thesis University of Aarhus, Denmark e-print arXiv:0709.0289.

[115] C. Schaffner, B. Terhal, and S. Wehner. Robust cryptography in the noisy-quantum-storage model. *Quantum Information and Computation*, 9(11):0963–0996, 2009.

[116] T. Schmitt-Manderbach, H. Weier, M. Fürst, R. Ursin, F. Tiefenbacher, T. Scheidl, J. Perdigues, Z. Sodnik, C. Kurtsiefer, J.G. Rarity, et al. Experimental demonstration of free-space decoy-state quantum key distribution over 144 km. *Physical Review Letters*, 98(1):10504, 2007.

[117] B. Schumacher. Sending entanglement through noisy quantum channels. *Physical Review A*, 54:2614, 1996.

[118] B. Schumacher and M. D. Westmoreland. Approximate quantum error correction. *Quantum Information Processing*, 1:5, 2002.

[119] C. E. Shannon. A mathematical theory of communication. *Bell System Technical Journal*, 27:379–423, 623–656, 1948.

[120] P. W. Shor. Scheme for reducing decoherence in quantum computer memory. *Physical Review A*, 52(4):2493–2496, 1995.

[121] P. W. Shor. Fault-tolerant quantum computation. In *Foundations of Computer Science'96*, page 56. Published by the IEEE Computer Society, 1996.

[122] P. W. Shor and J. Preskill. Simple proof of security of the BB84 quantum key distribution protocol. *Physical Review Letters*, 85(2):441–444, 2000.

[123] A. M. Steane. Error correcting codes in quantum theory. *Physical Review Letters*, 77(5):793–797, 1996.

[124] J. Tyson. Two-sided bounds on minimum-error quantum measurement, on the reversibility of quantum dynamics, and on maximum overlap using directional iterates. *Journal of Mathematical Physics*, 51:092204, 2010.

[125] S. Vadhan. On constructing locally computable extractors and cryptosystems in the bounded storage model. In *Advances in Cryptology—CRYPTO '03*, Lecture Notes in Computer Science, pages 61–77. Springer, 2003.

[126] L. Viola, E. Knill, and S. Lloyd. Dynamical decoupling of open quantum systems. *Physical Review Letters*, 82(12):2417–2421, 1999.

[127] S. Wehner. Cryptography in a quantum world. 2008. PhD Thesis, University of Amsterdam, e-print arXiv:0806.3483.

[128] S. Wehner, C. Schaffner, and B. M. Terhal. Cryptography from noisy storage. *Physical Review Letters*, 100(22):220502, 2008.

[129] S. Wehner and A. Winter. Higher entropic uncertainty relations for anti-commuting observables. *Journal of Mathematical Physics*, 49:062105, 2008.

[130] S. Wehner and A. Winter. Entropic uncertainty relations—a survey. *New Journal of Physics*, page 025009, 2010. Special Issue on Quantum Information and Many-Body Systems.

[131] P. Wocjan and T. Beth. New construction of mutually unbiased bases in square dimensions. *Quantum Information and Computation*, 5(2):93–101, 2005.

[132] W. K. Wootters and D. M. Sussman. Discrete phase space and minimum-uncertainty states. 2007. e-print arXiv:0704.1277.

[133] W. K. Wootters and W. H. Zurek. A single quantum cannot be copied. *Nature*, 299:802–803, 1982.

[134] W.K. Wootters and B. Fields. Optimal state-determination by mutually unbiased measurements. *Annals of Physics*, 191(368), 1989.

[135] N. Yamamoto, S. Hara, and K. Tsumura. Suboptimal quantum-error correcting procedure based on semidefinite programming. *Physical Review A*, 71:022322, 2005.

[136] A. C. Yao. Protocols for secure computations. In *Proceedings of the 23rd Annual IEEE FOCS*, pages 160–164, 1982.

[137] G. Zauner. Quantendesigns—grundzüge einer nichtkommutativen designtheorie. 1999. PhD Thesis Universität Wien.