Applications of Semidefinite Programming in Quantum Cryptography

by

Jamie W. J. Sikora

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Master of Mathematics
in
Combinatorics and Optimization

Waterloo, Ontario, Canada, 2007

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

Coin-flipping is the cryptographic task of generating a random coin-flip between two mistrustful parties. Kitaev [19] discovered that the security of quantum coin-flipping protocols can be analyzed using semidefinite programming. This lead to his result that one party can force a desired coin-flip outcome with probability at least $1/\sqrt{2}$.

We give sufficient background in quantum computing and semidefinite programming to understand Kitaev's semidefinite programming formulation for coin-flipping cheating strategies. These ideas are specialized to a specific class of protocols singled out by Nayak and Shor [29]. We also use semidefinite programming to solve for the maximum cheating probability of a particular protocol which has the best known security.

Furthermore, we present a family of protocols where one party has a greater probability of forcing an outcome of 0 than an outcome of 1. We also discuss a computer search to find specific protocols which minimize the maximum cheating probability.

Acknowledgements

I would like to deeply thank my supervisor Dr. Levent Tunçel for his guidance and direction throughout my undergraduate and graduate studies. He taught me that if I want something in life it is worth working toward. He convinced me to look into the Combinatorics and Optimization Summer Research Program at the University of Waterloo which was my first step into research.

I would also like to deeply thank my other supervisor Dr. Ashwin Nayak. His patience and teachings were a huge help in learning and understanding quantum computing. He was always willing to answer any of my questions and no question was ever too small.

Many thanks go to the members of my thesis reading committee, Dr. John Watrous and Dr. Andris Ambainis, for helpful comments and suggestions.

I would also like to thank Dr. Henry Wolkowicz who, through his enthusiastic lecturing, further inspired me to study continuous optimization.

Furthermore, I thank my good friend Gary Au for always being a caring and supportive friend. I also thank Nathan Krislock for all the help with Matlab.

Most of all, I would like to thank my parents, my brother, and my girlfriend for their endless support and encouragement.

Contents

1 Introduction **1**

2 Quantum Computing **8**
2.1 Background Material 8
2.2 Quantum Computing Definitions and Notation 12
2.3 Ensembles and Density Matrices 14
2.4 Evolution of a Quantum State 16
2.5 Projective Measurements 19

3 The Partial Trace and Partitioned Matrix Theory **23**
3.1 The Partial Trace 23
3.2 The Block Kronecker Product 31

4 Survey of Fidelity **36**
4.1 Fidelity of Quantum States 36
4.2 Positive Definite Matrix Means 40

5 Semidefinite Programming **42**
5.1 Semidefinite Programming and Duality 42
5.2 Applications of Semidefinite Programming in Quantum Computing 44

6 Quantum Coin-Flipping **48**
6.1 Definitions and Notation 48
6.2 An Example of a Coin-Flipping Protocol 52

7 Security Analysis of Quantum Coin-Flipping Protocols **56**
7.1 Semidefinite Programming Formulation 56
7.2 A Lower Bound on the Bias of any Coin-Flipping Protocol 60
7.3 Solving for the Bias with One SDP 62

8 Coin-Flipping Based on Quantum Bit-Commitment **64**
8.1 Quantum Bit-Commitment 64
8.2 Coin-Flipping Based on Quantum Bit-Commitment 65
8.3 Three-Round BCCF Protocols 67
8.4 Semidefinite Programming Formulation 70
8.5 An Upper Bound on the Best Possible Bias 72

9 Asymmetry in Cheating Probabilities **84**
9.1 An Asymmetric Protocol 84
9.1.1 Optimal Cheating Strategy for Forcing Outcome 0 93
9.1.2 Optimal Cheating Strategy for Forcing Outcome 1 95

10 Computational Results **98**
10.1 The Search for 3 and 4-Round Protocols with Low Bias 98
10.1.1 Computational Platform 102
10.2 Three-Round BCCF Search Results 103
10.3 Four-Round BCCF Search Results 103

11 Conclusion and Future Research **105**

Chapter 1

Introduction

Quantum computing is the study of computing using the laws of quantum mechanics. Classical computers, the computers we all use today, do computations on arrays of bits, e.g., variables being in either the state 0 or 1. The power of quantum computing is that it can deal with quantum states, an example being a *qubit*, which is the quantum analogue of a bit. A qubit has the possibility of being in the 0 state and the 1 state simultaneously, and we then say that a qubit is in a *superposition* of the two bits. The ability to *interfere* superpositions makes quantum computing a powerful generalization of classical computing and can make many problems much easier. Indeed there are quantum algorithms which provide an exponential speedup over the best classical algorithms. One of the first breakthroughs for quantum computing is a discovery by Shor. He discovered a quantum algorithm for factoring integers in polynomial-time [32], a problem for which no classical polynomial-time algorithms are yet known and many believe that none exists. Since many cryptographic schemes rely on the assumed intractability of factoring integers, an implementation of the Shor Algorithm on a quantum computer would be able to break these schemes efficiently. This raises the question of which classical cryptographic schemes would not be obsolete in the quantum computing age. If we had access to quantum computers, then how would communication and the transfer of information change? How could we view cryptography if we had access to quantum messages? These are some of the important issues in quantum information theory.

Quantum information theory is the study of how information can be sent over quantum

communication channels. This is a very rich theory and is fundamental when studying quantum cryptography since communication is the underlying task in many cryptographic schemes. Quantum information theory deals with many of the same issues as classical information theory, such as data compression and channel capacities. It also has some interesting results which are not true in classical information theory. One of these results, which is a fundamental theorem in quantum information theory, is the *no-cloning theorem*, see [11] and [42]. The no-cloning theorem states that arbitrary unknown quantum states can not be copied. This is not the case in classical computing since a classical computer can easily copy as many bit values as needed. Another important result is that it is not always possible to distinguish between quantum states, and attempting to do so typically disturbs the state. Suppose Alice sends a qubit to Bob but Eve, the eavesdropper, intercepts the qubit. If Eve does not know what state the qubit is in, then it may not be possible for her to completely determine its state. If she wants to extract any information about the qubit, she can do so but at the price of disturbing the superposition of the qubit. If the superposition is disturbed, then she no longer has a copy of the qubit. We see that the no-cloning theorem is important here since Eve is not able to copy the qubit she has in possession. If she could copy the qubit, she could determine its state to any level of accuracy. As a consequence of these two results, the amount of information an eavesdropper can intercept is limited. This is the basis for many cryptographic schemes such as in *quantum key distribution*.

Quantum cryptography is the study of cryptography using quantum computers and quantum communication channels. In quantum cryptography, often we only need to assume the laws of quantum mechanics. A scenario in quantum cryptography is Alice sending an encrypted message to Bob which is protected from Eve determining the content of the message. Many of these schemes use the results mentioned above to limit the information an eavesdropper can obtain. An example of this is quantum key distribution. Key distribution is the task of two communicating parties trying to establish a *private key*. Two parties who share a private key can use it to encrypt and decrypt messages. The issue is how to securely exchange private keys when the private key needs to be changed. This is very hard to implement classically. In the past, private keys were sealed in suitcases and hand-delivered, which raises security issues of its own. There are quantum protocols for secure key distribution which resolves the issue of exchanging private keys. The first popular cryptographic scheme was created by Bennett and Brassard in 1984. They developed a

protocol for quantum key distribution called BB84, see [8]. This was a huge breakthrough and it illustrates how quantum computing will change the future of cryptography. The idea is Alice sends Bob messages through a quantum channel and then they can test to see if any errors occurred in the transmission. The more information Eve tries to obtain from the messages the higher the error rate is, alerting Alice and Bob to the eavesdropper and allowing them to abort the protocol. If they detect that few errors have occurred, they use the messages to create a private key.

Another scenario in cryptography is multiple parties performing a computation where the parties are remotely located. The security now sought is to protect against dishonest parties altering the computation to force a desired result. This type of security differs from the other scenario since instead of seeking security against third parties gaining information, we now seek security from a dishonest party attempting to change the information. Classically, these protocols are implemented with the security based on assumed computational limitations. In quantum computing, there are tests each party can perform to see if the other parties have tried to alter or change the information. This limits how much one party can effectively 'cheat.' An important first step in understanding how such security can be implemented is to examine how two parties can generate a random coin-flip, a problem known as *coin-flipping*.

Coin-flipping is the cryptographic primitive of two mistrustful people trying to generate a fair coin-flip, where the two parties are not in the same location. The two parties communicate through sets of messages and then output a bit which has equal probability of being a 0 or 1. A typical scenario is a recently divorced couple who have moved apart and enter into a coin-flipping protocol where the resulting bit determines who gets ownership of their house. This is an example of *weak coin-flipping*, where the participants desire opposing outcomes. There is another version, called *strong coin-flipping*, where no desired outcomes are assumed. Call the two parties Alice and Bob and suppose that Bob wants the outcome 0. He may attempt to 'cheat' by deviating from the rules of the protocol trying to force the outcome to be 0 with high probability. As mentioned before, classical coin-flipping protocols are based on computational limitations. In quantum coin-flipping, each party is able to test whether or not the other has affected the information, trying to disturb the probabilities of the outcomes. There are a few different types of security one can study in coin-flipping, an example being *cheat-sensitivity* [37]. A cheat-sensitive

quantum coin-flipping protocol is one in which there is a positive probability of detecting a cheating attempt of a dishonest party. This type of security can be useful if there is a high penalty for cheating. Another kind of security studied in coin-flipping is the extent to which one party can affect the outcome of the protocol. Analyzing this type of security of strong coin-flipping protocols is the focus of this thesis.

A reason quantum coin-flipping is an interesting problem is the fact that classical coin-flipping is impossible (without assuming the validity of complexity theoretic conjectures). It is known that one party can cheat perfectly in a classical coin-flipping protocol by forcing an outcome with probability 1, see, for instance [24]. When implemented in the quantum computing framework, we can design coin-flipping protocols where neither party can force an outcome with probability 1, see [1]. Thus, there is a provable advantage to having quantum coin-flipping protocols over the classical variety. When designing such protocols, we seek the security to protect against both parties' attempts to force either outcome. Suppose Alice can deviate from the rules and force an outcome of 0 with a maximum probability of $P^*_{A,0}$. We can similarly define $P^*_{A,1}$, $P^*_{B,0}$, and $P^*_{B,1}$ for the other cheating probabilities. Any protocol is as insecure as the maximum of these four quantities since we do not know beforehand who is cheating and towards what outcome. We define the *bias*, denoted ε, of a strong coin-flipping protocol as

$$\varepsilon := \max\{P^*_{A,0}, P^*_{A,1}, P^*_{B,0}, P^*_{B,1}\} - \frac{1}{2}.$$

The bias is the maximum a party can force a desired outcome above the honest outcome probability of 1/2. Thus, we have turned the notion of security into an optimization problem. Due to the underlying convex structure of the state space in the setting of quantum mechanics, we can analyze the security with *semidefinite optimization*.

Semidefinite optimization is the study of optimizing a linear function over a positive semidefinite matrix variable which is subject to affine constraints. This has been applied to many combinatorial problems, a popular example being the MAX-CUT problem [13]. The MAX-CUT problem is the problem of finding a cut of maximum weight in an edge-weighted graph. No polynomial-time algorithm for this problem is known to exist. Using semidefinite optimization, we can solve the MAX-CUT problem in polynomial-time so that the expected value of the solution is at least 0.87856 times the optimal value, as proved by Goemans and Williamson [13]. This illustrates that semidefinite optimization is an

indispensable tool for analyzing some combinatorial problems.

It is significant that semidefinite optimization plays a role in quantum cryptography since we often examine worst-case scenarios. We attempt to design protocols to nullify worst cases. This is very important in coin-flipping since a coin-flipping protocol is as insecure as the maximum any of the two parties can cheat. If a party is going to cheat, there is no advantage for him/her to cheat in a non-optimal way. This makes it important to be able to quantify the extent of the cheating capabilities, which is something we can compute using semidefinite optimization. In fact, semidefinite optimization can also be used to describe the optimal cheating strategies for each party trying to force either outcome. It is imperative to keep this in mind when designing new protocols to avoid such opportunities for either party to cheat.

Coin-flipping is an interesting problem to study for several reasons. One reason is that coin-flipping is the most basic task in computation involving multiple parties. A solid understanding of how two people can compute a random coin-flip is key to understanding how more parties can compute more general functions. Secondly, coin-flipping is often used as a subroutine in a larger protocol. Thus, it is important to have a coin-flipping protocol with low bias since it could affect the security of the larger protocol. In fact, there are cryptographic protocols where the outcome of a coin-flip determines the next step of the protocol. These reasons illustrate why coin-flipping is an interesting and important problem to study.

There are some important results known about strong coin-flipping. Mayers [23] and Lo and Chau [22] proved that a protocol with 0 bias is impossible. Aharonov et al. [1] showed there exists a protocol with bias strictly less than 1, which is not possible in the classical case. The upper bound on the best possible bias was improved to $\varepsilon = 1/4$ by Ambainis [2] and independently by Spekkens and Rudolph [36]. Ambainis [2] showed that if a protocol has bias ε, then the protocol requires $\Omega\left(\log\log\frac{1}{\varepsilon}\right)$ rounds of communication. This result holds for weak coin-flipping as well where the bias of a weak coin-flipping protocol is defined analogously. Kitaev [19] used semidefinite optimization to show the best known lower bound for strong coin-flipping. He showed that any strong coin-flipping protocol has bias at least $\varepsilon \geq 1/\sqrt{2} - 1/2$. Ambainis et al. [5] give a detailed proof of Kitaev's lower bound and Watrous and Gutoski [14] give a different proof of the same lower bound. There

are also some important results about weak coin-flipping. Kerenidis and Nayak [18] showed that there is a weak coin-flipping protocol achieving a bias of less than 1/4. Spekkens and Rudolph [37] found a protocol achieving a bias of $1/\sqrt{2} - 1/2$, independently found by Ambainis, see [3]. Mochon [27] showed there exists a protocol with bias 1/6 which is the best known upper bound for the optimal bias. Ambainis' lower bound on the number of rounds of communication is the best known lower bound for the bias in weak coin-flipping.

Our goal is to examine a class of strong coin-flipping protocols based on *quantum bit-commitment*, singled out by Nayak and Shor [29]. Examining this class of protocols may give insight towards the general case of coin-flipping. This class includes the protocols by Ambainis and Spekkens and Rudolph which have bias $\varepsilon = 1/4$. We analyze these protocols using semidefinite optimization and illustrate how they also satisfy Kitaev's lower bound. We show that asymmetric cheating probabilities arise when there are more than three rounds of communication, the content of Chapter 9. We also present computational results for finding the optimal bias in special cases of these protocols.

The next chapter introduces the key concepts of quantum computing. It presents properties of quantum states and gives their most general representation. This chapter discusses basic concepts of quantum mechanics such as the evolution of quantum states and projective measurements. Chapter 3 discusses a linear operator called the *partial trace* which is an operator used extensively in quantum computing and appears in semidefinite programs (the content of later chapters). We investigate a generalization of the Kronecker product and show its close connection to the partial trace. Chapter 4 discusses a 'closeness' measure for quantum states called the *fidelity*. This chapter shows properties of the fidelity, which play an important role in the analysis of coin-flipping protocols and bit-commitment, as shown in Chapter 8. We also briefly discuss positive definite matrix means. Chapter 5 introduces semidefinite programming. We discuss duality theory which is essential to the analysis in later chapters. We also give some applications of semidefinite programming in quantum computing. In Chapter 6, we introduce quantum coin-flipping which is the focus of the following chapters. We define coin-flipping protocols, cheating strategies, and the bias of a protocol. Chapter 7 shows the analysis of quantum coin-flipping protocols by formulating cheating strategies as semidefinite programs. We examine Kitaev's lower bound which is the best known lower bound for a general coin-flipping protocol. In Chapter 8, we define a class of coin-flipping protocols based on quantum bit-commitment. A protocol

with the best known bias belongs to this class and we solve for its bias using semidefinite optimization. Chapter 9 presents a family of protocols with asymmetric cheating probabilities. We analyze and find cheating strategies for these protocols and show how semidefinite optimization plays a role. In Chapter 10, we discuss our computational search for protocols with low bias. We define an equivalence on protocols and show that our search performs well on a known example.

Chapter 2

Quantum Computing

2.1 Background Material

A Hilbert space is a vector space equipped with an inner product such that the norm induced by the inner product is a complete metric. We are primarily concerned with complex Hilbert spaces and denote the *set of linear operators* $\mathcal{L} : \mathcal{H} \to \mathcal{H}$ by $\mathcal{L}(\mathcal{H})$.

The set of $m \times n$ matrices with complex entries is denoted by $\mathbb{M}^{m,n}$ and the set of $n \times n$ matrices with complex entries is denoted by $\mathbb{M}^n$.

Given two matrices A and B we define their *Kronecker* or *tensor product* as

$$A \otimes B = \begin{bmatrix} A_{1,1}B & A_{1,2}B & \cdots & A_{1,n}B \\ A_{2,1}B & A_{2,2}B & \cdots & A_{2,n}B \\ \vdots & \vdots & \ddots & \vdots \\ A_{n,1}B & A_{n,2}B & \cdots & A_{n,n}B \end{bmatrix}.$$

Properties of the Kronecker product are well-known in literature. We present here some basic properties of the Kronecker product.

Properties 2.1. *Assume A, B, C, and D are matrices and $k \in \mathbb{C}$ is a scalar. Then we have the following properties of the Kronecker product.*

1. $(A + B) \otimes C = A \otimes C + B \otimes C$,
2. $A \otimes (B + C) = A \otimes B + A \otimes C$,

3. $k(A \otimes B) = kA \otimes B = A \otimes kB$ *(Bilinearity),*
4. $(A \otimes B) \otimes C = A \otimes (B \otimes C)$ *(Associativity),*
5. $(A \otimes B)(C \otimes D) = AC \otimes BD$ *(Mixed-Product Property),*
6. $\mathrm{Tr}\,(A \otimes B) = \mathrm{Tr}\,(A)\mathrm{Tr}\,(B)$*, if A and B are square,*
7. $\det(A \otimes B) = \det(A)^m \det(B)^n$*, if $A \in \mathbb{M}^n, B \in \mathbb{M}^m$,*
8. $(A \otimes B)^{-1} = A^{-1} \otimes B^{-1}$*, if A and B are invertible,*
9. $(A \otimes B)^* = A^* \otimes B^*$*,*
10. $\mathrm{rank}\,(A \otimes B) = \mathrm{rank}\,(A) \cdot \mathrm{rank}\,(B)$*.*

Given two Hilbert spaces $\mathcal{H}_1$, having orthonormal basis $\mathcal{B}_1$, and $\mathcal{H}_2$, having orthonormal basis $\mathcal{B}_2$, we can create another Hilbert space $\mathcal{H}_1 \otimes \mathcal{H}_2$ having the orthonormal basis

$$\{v \otimes w : v \in \mathcal{B}_1, w \in \mathcal{B}_2\}.$$

A tensor product of Hilbert spaces is the setting for quantum mechanics, discussed in Section 2.2.

We say that an operator $A \in \mathcal{L}(\mathcal{H})$ is a *nonnegative operator* if

$$\langle h, Ah\rangle \geq 0,\ \forall h \in \mathcal{H},$$

and a *positive operator* if

$$\langle h, Ah\rangle > 0,\ \forall h \in \mathcal{H} \setminus \{0\}.$$

If a matrix is a nonnegative operator we say it is *positive semidefinite* and if it is a positive operator we say it is *positive definite*.

We define

$$\Sigma^n := \{X \in \mathbb{M}^n : X = X^*\}$$

as the *set of Hermitian matrices*,

$$\Sigma^n_+ := \{X \in \Sigma^n : h^* X h \geq 0, \forall h \in \mathbb{C}^n\}$$

as the *set of positive semidefinite matrices*, and

$$\Sigma^n_{++} := \{X \in \Sigma^n : h^*Xh > 0, \forall h \in \mathbb{C}^n \setminus \{0\}\}$$

as the *set of positive definite matrices*. Note we have the following containments

$$\Sigma^n_{++} \subset \Sigma^n_+ \subset \Sigma^n.$$

A positive semidefinite matrix is denoted as $X \succeq 0$ and a positive definite matrix is denoted as $X \succ 0$ when the dimensions are clear from context. We write $A \succeq B$ meaning $A - B \succeq 0$ and $A \succ B$ meaning $A - B \succ 0$.

The standard inner product on $\mathbb{M}^{m,n}$ is the *trace inner product* defined as

$$\langle X, Y \rangle = \mathrm{Tr}\,(X^*Y).$$

Given $A, B \in \Sigma^n$, we know $A = A^*$ and $B = B^*$, so we can write

$$\langle A, B \rangle^* = \langle B, A \rangle = \mathrm{Tr}\,(B^*A) = \mathrm{Tr}\,(AB^*) = \mathrm{Tr}\,(A^*B) = \langle A, B \rangle .$$

Therefore, $\langle B, A \rangle = \langle A, B \rangle \in \mathbb{R},\ \forall A, B \in \Sigma^n$.

For a matrix $X \in \mathbb{M}^n$, we index its eigenvalues in descending order and denote them as

$$\lambda_1(X) \geq \lambda_2(X) \geq \cdots \geq \lambda_n(X)$$

and collect them in a vector and denote it as

$$\lambda(X) := [\lambda_1(X), \lambda_2(X), \ldots, \lambda_n(X)]^T .$$

A basic linear algebra result is that $\lambda(X) \in \mathbb{R}^n,\ \forall X \in \Sigma^n$. For a vector $x \in \mathbb{R}^n$, we write $x \geq 0$ meaning that each element of x is nonnegative and $x > 0$ meaning each element is positive. Using this notation, we also have that

$$\Sigma^n_+ = \{X \in \Sigma^n : \lambda(X) \geq 0\} \quad \text{and} \quad \Sigma^n_{++} = \{X \in \Sigma^n : \lambda(X) > 0\}.$$

The above description of positive semidefinite and positive definite matrices serve as alternative definitions. Notice that every positive definite matrix is invertible by the absence of zero eigenvalues.

We now present some well-known results about positive semidefinite matrices and tensor products. The following can be found in, for example, [38].

Lemma 2.2. *For any two positive semidefinite matrices $X, S \in \Sigma_+^n$, we have that*

$$\langle X, S \rangle \geq 0.$$

The following corollary is immediate from linearity of inner products.

Corollary 2.3. *For any positive semidefinite matrix $X \in \Sigma_+^n$, and matrices $A, B \in \Sigma^n$ satisfying $A \succeq B$, we have that*

$$\langle A, X \rangle \geq \langle B, X \rangle .$$

The following lemma characterizes the eigenvalue-eigenvector pairs of tensor products, see for instance [39].

Lemma 2.4. *Suppose $A \in \mathbb{M}^n$ has eigenvalue-eigenvector pair (λ, v) and $B \in \mathbb{M}^m$ has eigenvalue-eigenvector pair (μ, w). Then $A \otimes B$ has eigenvalue-eigenvector pair $(\lambda\mu, v \otimes w)$.*

Note that we may not have $\lambda(A \otimes B) = \lambda(A) \otimes \lambda(B)$ since it may not list the eigenvalues in descending order.

Next is a very important result relating positive semidefinite matrices and tensor products.

Corollary 2.5. *For any two positive semidefinite matrices $X \in \Sigma_+^n$ and $S \in \Sigma_+^m$, we have that*

$$X \otimes S \in \Sigma_+^{m \cdot n}.$$

From bilinearity of the tensor product, we have the following corollary.

Corollary 2.6. *For any positive semidefinite matrix $X \in \Sigma_+^n$, and matrices $A, B \in \Sigma^n$ satisfying $A \succeq B$, we have that*

$$A \otimes X \succeq B \otimes X.$$

We present two well-known lemmas which show properties of Σ_+^n we use extensively.

Lemma 2.7. *Σ_+^n is a convex cone. That is,*

$$X \in \Sigma_+^n, \lambda \geq 0 \implies \lambda X \in \Sigma_+^n \quad \textit{and} \quad X, S \in \Sigma_+^n \implies X + S \in \Sigma_+^n.$$

Next, we mention an important set of automorphisms of the cone of positive semidefinite matrices.

Lemma 2.8. *Given any invertible matrix $M \in \mathbb{M}^n$, we have $T(\cdot) = M(\cdot)M^* \in \mathrm{Aut}\,(\Sigma_+^n)$ and $T(\cdot) = M(\cdot)M^* \in \mathrm{Aut}\,(\Sigma_{++}^n)$. That is,*

$$X \in \Sigma_+^n \iff MXM^* \in \Sigma_+^n, \quad \textit{for all invertible } M \in \mathbb{M}^n,$$

and

$$X \in \Sigma_{++}^n \iff MXM^* \in \Sigma_{++}^n, \quad \textit{for all invertible } M \in \mathbb{M}^n.$$

We now define *orthogonal projectors* which we use in quantum computing.

Definition 2.9 (Orthogonal Projector)**.** *Let $\mathcal{H}$ be an inner product space and let $\mathcal{G} \subseteq \mathcal{H}$ be a subspace of $\mathcal{H}$. Then we define $\Pi \in \mathcal{L}(\mathcal{H})$ to be the* orthogonal projector onto $\mathcal{G}$ *if it satisfies*

$$\Pi(v) = v, \ \forall v \in \mathcal{G}, \quad \textit{and} \quad \Pi(w) = 0, \ \forall w \in \mathcal{G}^\perp.$$

Immediate from the definition is that orthogonal projectors are idempotent, that is, any orthogonal projector $\Pi \in \mathcal{L}(\mathcal{H})$ satisfies $\Pi^2 = \Pi$. Also, every orthogonal projector $\Pi \in \mathcal{L}(\mathcal{H})$ satisfies $\lambda(\Pi) \in \{0,1\}^{\dim(\mathcal{H})}$.

2.2 Quantum Computing Definitions and Notation

This section presents some basic concepts of quantum computing. A reference for many of these results and related linear algebra results is [31].

A *pure quantum state* $|\psi\rangle$ is a vector with unit norm in a Hilbert space $\mathcal{H}$ and $\mathcal{H}$ is the corresponding *state space*. In quantum computing we use Dirac notation for vectors which includes $|\psi\rangle$, read as "*ket* ψ," and $\langle\psi| := |\psi\rangle^*$, read as "*bra* ψ." We use $\langle\psi|\phi\rangle$ to denote the inner product on $\mathcal{H}$ of the vectors $|\psi\rangle, |\phi\rangle \in \mathcal{H}$ and $|\psi\rangle\langle\phi|$ to denote the outer product.

Using this notation, we have that

$$\Sigma_+^{\mathcal{H}} = \{A \in \Sigma^n : \langle\psi| A |\psi\rangle \geq 0, \ \forall |\psi\rangle \in \mathcal{H}\}$$

and

$$\Sigma_{++}^{\mathcal{H}} = \{A \in \Sigma^n : \langle\psi| A |\psi\rangle > 0, \ \forall |\psi\rangle \in \mathcal{H} \setminus \{0\}\}.$$

Notice that $\operatorname{Tr}|\psi\rangle\langle\psi| = \langle\psi|\psi\rangle = \||\psi\rangle\|^2 = 1$ for any quantum state $|\psi\rangle$ and $|\psi\rangle\langle\psi|$ is a matrix with rank and trace equal to 1. Another useful fact to note is the *resolution of the identity*, which states that

$$\sum_{|\phi\rangle\in B} |\phi\rangle\langle\phi| = I_{\mathcal{H}},$$

for any orthonormal basis B of the Hilbert space $\mathcal{H}$.

The quantum analogue of a *bit* is a *qubit* and the quantum analogue of a *trit* is a *qutrit*. A bit can take two values so we define a qubit to be a unit vector in the two-dimensional Hilbert space $\mathbb{C}^2$. A qubit $|\psi\rangle$ can thus be represented in the form

$$|\psi\rangle = \begin{bmatrix} \alpha \\ \beta \end{bmatrix},$$

where

$$|\alpha|^2 + |\beta|^2 = 1.$$

Notice if we let

$$|0\rangle := \begin{bmatrix} 1 \\ 0 \end{bmatrix} \text{ and } |1\rangle := \begin{bmatrix} 0 \\ 1 \end{bmatrix},$$

we can represent a qubit in the form

$$\alpha|0\rangle + \beta|1\rangle, \text{ where } |\alpha|^2 + |\beta|^2 = 1,$$

and we say $|\psi\rangle$ is in a *superposition* of the states $|0\rangle$ and $|1\rangle$.

The power of quantum computing comes from the fact that it can deal with quantum states being in a superposition of many basis states. We say that $\{|0\rangle, |1\rangle\}$ is the *computational basis* for $\mathbb{C}^2$. Not every quantum state is a qubit however. For example, we could have a quantum state of multiple qubits. For this we need to be able to describe the corresponding state space. It turns out that the state space for multiple states is the tensor product of the individual state spaces. I.e., if we have a state of three qubits, the corresponding state space is $\mathbb{C}^2 \otimes \mathbb{C}^2 \otimes \mathbb{C}^2 = \mathbb{C}^{2^3}$.

Given a matrix $M \in \mathbb{M}^{m,n}$, we have the correspondence between linear algebra notation and Dirac notation given as

$$M = [m_{i,j}] = \sum_{i=1}^{m}\sum_{j=1}^{n} m_{i,j}|i\rangle\langle j|.$$

Notice that $|i\rangle\langle j| = E_{i,j}$ where $\{E_{i,j} : i \in \{1,\ldots,m\}, j \in \{1,\ldots,n\}\}$ is the standard basis for $\mathbb{M}^{m,n}$.

There is more notation which makes things easier. Notice $|1\rangle \otimes |1\rangle \in \mathbb{C}^2 \otimes \mathbb{C}^2$ is another quantum state. When we write this state, we usually leave out the tensor products or take this a step further and include it in one large ket, i.e., $|1\rangle\,|1\rangle$ or $|11\rangle$. Note that this could be confused whether this is "11" in binary or another basis state such as "11" in state space $\mathbb{C}^{12}$, but the former should be assumed unless the context suggests otherwise. Suppose we have a multiple qubit state space $\mathcal{H} := \mathbb{C}^2 \otimes \cdots \mathbb{C}^2$. We use the notation $|\mathbf{0}\rangle_{\mathcal{H}} := |0\cdots 0\rangle$ to denote multiple qubits initialized to the $|0\rangle$ state. Usually only qubits not in the $|\mathbf{0}\rangle$ state are explicitly mentioned.

With these conventions, consider the following two-qubit quantum states

$$|\psi_1\rangle := \frac{1}{\sqrt{2}}\,|00\rangle + \frac{1}{\sqrt{2}}\,|01\rangle \quad \text{and} \quad |\psi_2\rangle := \frac{1}{\sqrt{2}}\,|00\rangle + \frac{1}{\sqrt{2}}\,|11\rangle\,.$$

We see that $|\psi_1\rangle$ has the tensor product factorization $|\psi_1\rangle = |0\rangle \left(\frac{1}{\sqrt{2}}\,|0\rangle + \frac{1}{\sqrt{2}}\,|1\rangle\right)$. A quick check shows that $|\psi_2\rangle$ does not have such a factorization, the two qubits are then said to be *entangled*.

2.3 Ensembles and Density Matrices

We define an *ensemble* as a probability distribution over pure states in state space $\mathcal{H}$, denoted as

$$\mathcal{E} := \{(p_i, |\psi_i\rangle) : i \in \{1,\ldots,k\}, |\psi_i\rangle \in \mathcal{H}\}.$$

This is also referred to as a *mixed state*. As a convention, we use $\{p_i, |\psi_i\rangle\}$ to denote an ensemble when the dimension is clear from context. For an ensemble $\{p_i, |\psi_i\rangle\}$, we define its corresponding *density matrix* as

$$\sum_i p_i |\psi_i\rangle\langle\psi_i|.$$

We also say that $X \in \Sigma^n$ is a density matrix if there exists an ensemble for which it is the density matrix. This is the most general description of a quantum state. Suppose ρ is the density matrix of the ensemble $\{p_i, |\psi_i\rangle\}$. It could be the case that ρ is also the density

matrix of a different ensemble $\{p'_i, |\psi'_i\rangle\}$. If two ensembles have the same density matrix they are said to be *equivalent*. Equivalent ensembles behave the same way under quantum operations and quantum measurements, see Section 2.5. Thus, they are effectively the same quantum state.

Suppose we have two parties, Alice and Bob, who have respective state spaces $\mathcal{A}$ and $\mathcal{B}$. If Alice and Bob share a quantum state $|\psi\rangle \in \mathcal{A} \otimes \mathcal{B}$, then neither have total control over $|\psi\rangle$. Mixed states are used to describe parts of quantum states over a tensor product when the other part of the state space is not available. This is discussed further in Chapter 3.

The following theorem is a fundamental result which is essential to the discussion in later chapters. It can be found in abundance in the literature, for example, see [31].

Theorem 2.10 (Density Matrix Equivalence)**.** *Given a matrix $X \in \Sigma^{\mathcal{H}}$, we have that X is a density matrix if and only if $X \in \Sigma_+^{\mathcal{H}}$ and* $\mathrm{Tr}\,(X) = 1$.

Proof. Suppose X is a density matrix of the ensemble $\{p_i, |\psi_i\rangle\}$, that is, $X := \sum_i p_i |\psi_i\rangle\langle\psi_i|$. Then we have

$$\mathrm{Tr}\,(X) = \mathrm{Tr}\left(\sum_i p_i |\psi_i\rangle\langle\psi_i|\right) = \sum_i p_i \mathrm{Tr}\,|\psi_i\rangle\langle\psi_i| = \sum_i p_i \langle\psi_i|\psi_i\rangle = \sum_i p_i = 1.$$

For any $|h\rangle \in \mathcal{H}$, we have

$$\langle h|\, X\, |h\rangle = \langle h| \left(\sum_i p_i |\psi_i\rangle\langle\psi_i|\right) |h\rangle = \sum_i p_i \langle h|\psi_i\rangle\langle\psi_i|h\rangle = \sum_i p_i |\langle h|\psi_i\rangle|^2 \geq 0.$$

Thus, $X \in \Sigma_+^{\mathcal{H}}$ and $\mathrm{Tr}\,(X) = 1$.
Now suppose we are given $X \in \Sigma_+^{\mathcal{H}}$ where $\mathrm{Tr}\,(X) = 1$. The spectral decomposition of X is given as

$$X = \sum_{i=1}^{n} \lambda_i(X) |v_i\rangle\langle v_i|,$$

where $(\lambda_i(X), |v_i\rangle)$ is an eigenvalue-eigenvector pair of X with $\||v_i\rangle\| = 1$, for all i in the set $\{1, \ldots, \dim(\mathcal{H})\}$. Thus, X is the density matrix of the ensemble $\{\lambda_i(X), |v_i\rangle\}$. □

Using this equivalence, we define the *set of density matrices over state space* $\mathcal{H}$ as

$$\mathbb{D}^{\mathcal{H}} := \{X \in \Sigma_+^{\mathcal{H}} : \mathrm{Tr}\,(X) = 1\}.$$

For ease of notation, we denote $\mathbb{D}^{\mathbb{C}^n}$ simply as $\mathbb{D}^n$.

One may think a more general description of a quantum state is being in a probabilistic mixture of density matrices. We can describe a quantum state in this manner by defining the density matrix of an ensemble of density matrices $\{p_i, \rho_i\}$ as

$$\sum_i p_i \rho_i.$$

We show that this is a density matrix as defined earlier, thus we have not abused notation.

The following is a well-known result.

Lemma 2.11. *The density matrix of a probabilistic mixture of density matrices is a density matrix.*

Proof. Consider the ensemble of density matrices $\{p_i, \rho_i\}$ and its density matrix

$$X := \sum_i p_i \rho_i.$$

Then

$$\mathrm{Tr}\,(X) = \mathrm{Tr}\left(\sum_i p_i \rho_i\right) = \sum_i p_i \mathrm{Tr}\,(\rho_i) = \sum_i p_i = 1.$$

For any $|h\rangle \in \mathcal{H}$, we have that

$$\langle h| X |h\rangle = \langle h| \left(\sum_i p_i \rho_i\right) |h\rangle = \sum_i p_i \langle h| \rho_i |h\rangle \geq 0,$$

since $\rho_i \in \Sigma_+^{\mathcal{H}}$ and $p_i \geq 0$. The result now follows from Theorem 2.10. □

The above lemma shows that $\mathbb{D}^{\mathcal{H}}$ is a convex set, i.e.,

$$\lambda \in [0,1] \text{ and } \rho_1, \rho_2 \in \mathbb{D}^{\mathcal{H}} \implies \lambda\rho_1 + (1-\lambda)\rho_2 \in \mathbb{D}^{\mathcal{H}}.$$

2.4 Evolution of a Quantum State

Suppose we have an operator $M \in \mathcal{L}(\mathcal{H})$ acting on a quantum state $|\psi\rangle \in \mathcal{H}$. If $\mathcal{H}$ is finite dimensional, what conditions must M satisfy? Suppose we want any such action to produce another quantum state, then it must preserve unit norm, i.e., we require

$$\|M|\psi\rangle\|_2 = 1, \ \forall\, |\psi\rangle \in \mathcal{H}.$$

From this we see that M must be unitary. In fact, any unitary matrix is a valid operation on a quantum state.

Suppose we are given a density matrix $\rho \in \mathbb{D}^n$, how does one apply a unitary to this quantum state? It turns out applying the unitary U to the density matrix ρ yields the density matrix $U\rho U^*$, we now show why this is the case. Suppose ρ is the density matrix of the ensemble $\mathcal{E} := \{p_i, |\psi_i\rangle\}$ and we apply the unitary U to each of the states in $\mathcal{E}$. This yields a new ensemble, namely $\mathcal{E}' := \{p_i, U|\psi_i\rangle\}$. By definition, the density matrix of $\mathcal{E}'$ is

$$\sum_i p_i (U|\psi_i\rangle)(U|\psi_i\rangle)^* = \sum_i p_i U|\psi_i\rangle\langle\psi_i|U^* = U\left(\sum_i p_i|\psi_i\rangle\langle\psi_i|\right)U^* = U\rho U^*,$$

where ρ is the density matrix of $\mathcal{E}$.

There are more physically possible operations on a quantum state and these operations are called *quantum operations*.

Definition 2.12 (Quantum Operation)**.** *Suppose we are given the following set of matrices* $A := \{A_i : i \in \{1, \ldots, k\}, A_i : \mathcal{H} \to \mathcal{G}\}$ *which satisfy* $\sum_{i=1}^k A_i^* A_i = I_{\mathcal{H}}$*. We define a* quantum operation $G_A : \mathcal{L}(\mathcal{H}) \to \mathcal{L}(\mathcal{G})$ *as*

$$G_A(X) := \sum_{i=1}^k A_i X A_i^*.$$

The elements of the set A in the above definition are called "*Kraus operators*." Quantum operations are defined to act on any matrix $X \in \mathcal{L}(\mathcal{H})$, though we are concerned with their actions on density matrices.

We now show the following well-known results that quantum operations are trace-preserving and preserve positive semidefiniteness.

Lemma 2.13. *Given a set of matrices* $A := \{A_i : i \in \{1, \ldots, k\}, A_i : \mathcal{H} \to \mathcal{G}\}$ *satisfying* $\sum_{i=1}^k A_i^* A_i = I_{\mathcal{H}}$*, we have*

$$\mathrm{Tr}\,(G_A(X)) = \mathrm{Tr}\,(X),\ \forall X \in \mathcal{L}(\mathcal{H}).$$

Proof. Suppose we are given a set of matrices $A := \{A_i : i \in \{1, \ldots, k\}, A_i : \mathcal{H} \to \mathcal{G}\}$

satisfying $\sum_{i=1}^{k} A_i^* A_i = I_{\mathcal{H}}$. Then we have

$$\begin{aligned}
\mathrm{Tr}\left(G_A(X)\right) &= \mathrm{Tr}\left(\sum_{i=1}^{k} A_i X A_i^*\right) \\
&= \sum_{i=1}^{k} \mathrm{Tr}\left(A_i X A_i^*\right) \\
&= \sum_{i=1}^{k} \mathrm{Tr}\left(A_i^* A_i X\right) \\
&= \mathrm{Tr}\left(\sum_{i=1}^{k} A_i^* A_i X\right) \\
&= \mathrm{Tr}\left(X\right).
\end{aligned}$$

□

We now show that quantum operations preserve positive semidefiniteness.

Lemma 2.14. *Given a set of matrices* $A := \{A_i : i \in \{1, \ldots, k\}, A_i : \mathcal{H} \to \mathcal{G}\}$ *satisfying* $\sum_{i=1}^{k} A_i^* A_i = I_{\mathcal{H}}$*, we have* $X \succeq 0 \implies G_A(X) \succeq 0$*. Furthermore, if* $\dim(\mathcal{G}) \leq \dim(\mathcal{H})$ *and* A_j *has full row rank for some* $j \in \{1, \ldots k\}$*, then we have* $X \succ 0 \implies G_A(X) \succ 0$*.*

Proof. Suppose we are given $X \succeq 0$ and $A := \{A_i : i \in \{1, \ldots, k\}, A_i : \mathcal{H} \to \mathcal{G}\}$ with $\sum_i A_i^* A_i = I_{\mathcal{H}}$. Then, for an arbitrary $|h\rangle \in \mathcal{G}$, we have

$$\langle h| G_A(X) |h\rangle = \langle h| \left(\sum_i A_i X A_i^*\right) |h\rangle = \sum_i \langle h| A_i X A_i^* |h\rangle = \sum_i \langle h_i| X |h_i\rangle ,$$

where $|h_i\rangle := A_i^* |h\rangle$. Therefore, $\langle h| G_A(X) |h\rangle = \sum_i \langle h_i| X |h_i\rangle \geq 0$, since $X \succeq 0$. Now, further suppose that $h \neq 0$, $X \succ 0$, $\dim(\mathcal{G}) \leq \dim(\mathcal{H})$, and A_j has full row rank for some $j \in \{1, \ldots, k\}$. Since A_j^* has full column rank, $|h_j\rangle = A_j^* |h\rangle \neq 0$. Therefore, we have

$$\langle h| G_A(X) |h\rangle = \sum_i \langle h_i| X |h_i\rangle \geq \langle h_j| X |h_j\rangle > 0,$$

since $X \succ 0$ and $|h_j\rangle \neq 0$. Since $|h\rangle$ was chosen arbitrarily, this proves the result. □

The following corollary is immediate from Theorem 2.10.

Corollary 2.15. *Suppose* $A := \{A_i : i \in \{1, \ldots, k\}, A_i : \mathcal{H} \to \mathcal{G}\}$ *is a set of matrices satisfying* $\sum_{i=1}^{k} A_i^* A_i = I_{\mathcal{H}}$. *If* X *is a density matrix, then so is* $G_A(X)$.

The converse does not hold in general. To see this, consider the matrix

$$X = 2|0\rangle\langle 0| - |1\rangle\langle 1| \in \mathcal{L}(\mathbb{C}^2),$$

which is not positive semidefinite by the presence of a negative eigenvalue. Consider the set $A := \{\langle 0|, \langle 1|\}$ which defines a quantum operation since $|0\rangle\langle 0| + |1\rangle\langle 1| = I_{\mathbb{C}^2}$. Notice $G_A(X) = [2] + [-1] = [1]$, thus $G_A(X)$ is a density matrix in $\mathcal{L}(\mathbb{C})$, but X is not a density matrix in $\mathcal{L}(\mathbb{C}^2)$.

2.5 Projective Measurements

Another operation on quantum states which is physically possible is a *measurement*. A measurement is when you "observe" a quantum state to obtain information about its state. A quantum state remains in superposition as long as it is in a closed system. A measurement consists of anything that obtains information about the state. If we introduce a measuring device then the system is no longer closed and this typically disturbs the superposition of the state. For example, a qubit $|\psi\rangle$ in a superposition of the states $|0\rangle$ and $|1\rangle$ can be written as

$$|\psi\rangle = \sum_{i=0}^{1} \alpha_i |i\rangle, \text{ where } \sum_{i=0}^{1} |\alpha_i|^2 = 1.$$

If we were to measure the qubit, then the possible outcomes are

$$\begin{cases} 0, & \text{with probability } |\alpha_0|^2, \\ 1, & \text{with probability } |\alpha_1|^2. \end{cases}$$

The sum of the probabilities is 1 due to the norm condition on quantum states. This is an example of measuring a qubit in the computational basis. When the state space has multiple tensor factors, it is possible to measure only part of a state. It is also possible to measure in other orthonormal bases as well, demonstrated in the following example. Suppose we measure the subspace $\mathcal{K}$ from the quantum state $|\psi\rangle \in \mathcal{H} \otimes \mathcal{K}$ in the orthonormal basis $\{|\nu_i\rangle\}$ of $\mathcal{K}$. This can be done by writing $|\psi\rangle$ as

$$|\psi\rangle = \sum_i \sqrt{\alpha_i}\,|\xi_i\rangle\,|\nu_i\rangle\,,$$

for some states $|\xi_i\rangle \in \mathcal{H}$, where $\alpha_i \geq 0$ satisfy $\sum_i \alpha_i = 1$. Then the measurement produces outcome $|\nu_i\rangle$ with probability α_i and "collapses" $|\psi\rangle$ to the state consistent with the outcome, i.e., $|\xi_i\rangle\,|\nu_i\rangle$.

We can also describe the process of measuring in a basis using a *projective measurement*.

Definition 2.16 (Projective Measurement)**.** *A* projective measurement *on a Hilbert space $\mathcal{H}$ is a set of mutually orthogonal projectors $\{\Pi_i \in \mathcal{L}(\mathcal{H}) : i \in \{1,\ldots,n\}\}$, such that*

$$\sum_{i=1}^{n} \Pi_i = I_{\mathcal{H}}.$$

Suppose we have a state $|\psi\rangle \in \mathcal{H}$ and we measure it using the projective measurement $\{\Pi_i \in \mathcal{L}(\mathcal{H}) : i \in \{1,\ldots,n\}\}$. The probability of outcome i is given by

$$p_i := \langle \Pi_i, |\psi\rangle\langle\psi| \rangle\,,$$

leaving the post-measurement state

$$\frac{1}{\sqrt{p_i}}\Pi_i\,|\psi\rangle\,.$$

We can also measure mixed states in the same way. This is done by measuring each state in the ensemble. Suppose we have the ensemble $\mathcal{E} := \{q_j, |\psi_j\rangle\}$. The probability of outcome i is given by

$$\begin{aligned} p_i &= \sum_j q_j \,\langle \Pi_i, |\psi_j\rangle\langle\psi_j| \rangle \\ &= \left\langle \Pi_i, \sum_j q_j |\psi_j\rangle\langle\psi_j| \right\rangle \\ &= \langle \Pi_i, \rho \rangle\,, \end{aligned}$$

where ρ is the density matrix of $\mathcal{E}$. This leaves the post-measurement mixed state

$$\left\{ q_j p_i, \frac{1}{\sqrt{p_i}}\Pi_i\,|\psi_j\rangle \right\}.$$

This ensemble has the density matrix

$$\sum_j q_j \left(\frac{1}{p_i} \Pi_i |\psi_j\rangle\langle\psi_j| \Pi_i \right) = \frac{\Pi_i \rho \Pi_i}{\langle \Pi_i, \rho \rangle}.$$

We again consider measuring the subspace $\mathcal{K}$ from a state $|\psi\rangle \in \mathcal{H} \otimes \mathcal{K}$ in the basis $\{|\nu_i\rangle\}$. We can do this by defining the projective measurement

$$\{\Pi_i = I_{\mathcal{H}} \otimes |\nu_i\rangle\langle\nu_i|\}.$$

This agrees with the definition of a projective measurement since $\{I_{\mathcal{H}} \otimes |\nu_i\rangle\langle\nu_i|\}$ is a set of mutually orthogonal projectors and the resolution of the identity tells us that

$$\sum_i I_{\mathcal{H}} \otimes |\nu_i\rangle\langle\nu_i| = I_{\mathcal{H}} \otimes I_{\mathcal{K}} = I_{\mathcal{H}\otimes\mathcal{K}}.$$

We verify this with the previous example. If we measure the state

$$|\psi\rangle := \sum_i \sqrt{\alpha_i} \, |\xi_i\rangle \, |\nu_i\rangle \in \mathcal{H} \otimes \mathcal{K}$$

with the projective measurement $\{\Pi_i = I_{\mathcal{H}} \otimes |\nu_i\rangle\langle\nu_i|\}$, we obtain outcome i with probability

$$\begin{aligned}
p_i &:= \langle \Pi_i, |\psi\rangle\langle\psi| \rangle \\
&= \left\langle I_{\mathcal{H}} \otimes |\nu_i\rangle\langle\nu_i|, \sum_{j,k} \sqrt{\alpha_j \alpha_k} \, |\xi_j\rangle \, \langle\xi_k| \otimes |\nu_j\rangle \, \langle\nu_k| \right\rangle \\
&= \sum_{j,k} \sqrt{\alpha_j \alpha_k} \langle\nu_i|\nu_j\rangle \langle\nu_k|\nu_i\rangle \mathrm{Tr} \, |\xi_j\rangle\langle\xi_k| \\
&= \sqrt{\alpha_i \alpha_i} \mathrm{Tr} \, |\xi_i\rangle\langle\xi_i| \\
&= \alpha_i.
\end{aligned}$$

This leaves the post-measurement state

$$\begin{aligned}
\frac{1}{\sqrt{p_i}} \Pi_i \, |\psi\rangle &= \frac{1}{\sqrt{\alpha_i}} \left(I_{\mathcal{H}} \otimes |\nu_i\rangle\langle\nu_i| \right) \left(\sum_j \sqrt{\alpha_j} \, |\xi_j\rangle \, |\nu_j\rangle \right) \\
&= \frac{1}{\sqrt{\alpha_i}} \sum_j \sqrt{\alpha_j} \, |\xi_j\rangle \, |\nu_i\rangle \, \langle\nu_i|\nu_j\rangle \\
&= |\xi_i\rangle \, |\nu_i\rangle \, .
\end{aligned}$$

This agrees with our original description of measuring a subspace of a state in the basis $\{|\nu_i\rangle\}$.

We now return to our discussion of quantum operations. Quantum operations are physically possible operations on quantum states. Suppose we want to perform a quantum operation on the state $|\psi\rangle \in \mathcal{H}$. We pad this state with an *ancilla*, $|0\rangle \in \mathcal{K}$, to create the state $|\psi\rangle_{\mathcal{H}} |0\rangle_{\mathcal{K}}$. Suppose we perform a unitary $U \in \mathcal{L}(\mathcal{H} \otimes \mathcal{K})$ on $|\psi\rangle_{\mathcal{H}} |0\rangle_{\mathcal{K}}$ then measure the subspace $\mathcal{K}$. This process can be described by a quantum operation on $\mathcal{H}$. Conversely, every quantum operation can be described in this manner.

Suppose we have a qubit entangled with another qubit but the other qubit is in another person's control. What statistical information do we have about the qubit? What if the other qubit has been measured? If the other party measured the qubit, then our qubit is in a probability distribution of possible states, as explained above. Thus, our qubit can be described as a density matrix. The operation which obtains this density matrix from the two entangled qubits is called the *partial trace*, which is examined in the next chapter.

Chapter 3

The Partial Trace and Partitioned Matrix Theory

3.1 The Partial Trace

Definition 3.1 (Partial Trace)**.** *Suppose we are given the Hilbert space* $\mathcal{H} := \mathcal{H}_1 \otimes \mathcal{H}_2 \otimes \mathcal{H}_3$ *and a matrix* $X \in \mathcal{L}(\mathcal{H})$*. We define the* partial trace over $\mathcal{H}_2$ *of* X *as*

$$\mathrm{Tr}_{\mathcal{H}_2}(X) := \sum_{|e\rangle \in B} (I_{\mathcal{H}_1} \otimes \langle e| \otimes I_{\mathcal{H}_3}) X (I_{\mathcal{H}_1} \otimes |e\rangle \otimes I_{\mathcal{H}_3}),$$

where B *is* any *orthonormal basis of* $\mathcal{H}_2$*.*

The above definition of the partial trace is general since we can set $\mathcal{H}_1 = \mathbb{C}^1$ and/or $\mathcal{H}_3 = \mathbb{C}^1$. This is because the only unit vector in $\mathbb{C}^1$ is 1 itself, and we can write X as $X = 1 \otimes X = X \otimes 1$.

The following well-known result states that the partial trace is well-defined.

Lemma 3.2. *The partial trace is well-defined, i.e., the definition is independent of the choice of orthonormal basis of* $\mathcal{H}_2$*.*

Proof. Recall the resolution of the identity

$$\sum_{|e\rangle \in B} |e\rangle\langle e| = I_{\mathcal{H}},$$

for any orthonormal basis B of a Hilbert space $\mathcal{H}$. Suppose we have the Hilbert space $\mathcal{H} := \mathcal{H}_1 \otimes \mathcal{H}_2 \otimes \mathcal{H}_3$. For two arbitrary bases B_1 and B_2 for $\mathcal{H}_2$, we can write the following.

$$\begin{aligned}
\operatorname{Tr}_{\mathcal{H}_2}(X) &= \sum_{|e\rangle \in B_1} (I_{\mathcal{H}_1} \otimes \langle e| \otimes I_{\mathcal{H}_3}) X (I_{\mathcal{H}_1} \otimes |e\rangle \otimes I_{\mathcal{H}_3}) \\
&= \sum_{|e\rangle \in B_1} (I_{\mathcal{H}_1} \otimes \langle e| \left(\sum_{|f\rangle \in B_2} |f\rangle\langle f| \right) \otimes I_{\mathcal{H}_3}) X (I_{\mathcal{H}_1} \otimes \left(\sum_{|f'\rangle \in B_2} |f'\rangle\langle f'| \right) |e\rangle \otimes I_{\mathcal{H}_3}) \\
&= \sum_{|e\rangle \in B_1, |f\rangle, |f'\rangle \in B_2} (I_{\mathcal{H}_1} \otimes \langle e|f\rangle \langle f| \otimes I_{\mathcal{H}_3}) X (I_{\mathcal{H}_1} \otimes |f'\rangle \langle f'|e\rangle \otimes I_{\mathcal{H}_3}) \\
&= \sum_{|e\rangle \in B_1, |f\rangle, |f'\rangle \in B_2} \langle e|f\rangle\langle f'|e\rangle (I_{\mathcal{H}_1} \otimes \langle f| \otimes I_{\mathcal{H}_3}) X (I_{\mathcal{H}_1} \otimes |f'\rangle \otimes I_{\mathcal{H}_3}) \\
&= \sum_{|f\rangle, |f'\rangle \in B_2} \left\langle |f\rangle\langle f'|, \sum_{|e\rangle \in B_1} |e\rangle\langle e| \right\rangle (I_{\mathcal{H}_1} \otimes \langle f| \otimes I_{\mathcal{H}_3}) X (I_{\mathcal{H}_1} \otimes |f'\rangle \otimes I_{\mathcal{H}_3}) \\
&= \sum_{|f\rangle, |f'\rangle \in B_2} \langle |f\rangle\langle f'|, I_{\mathcal{H}_2} \rangle (I_{\mathcal{H}_1} \otimes \langle f| \otimes I_{\mathcal{H}_3}) X (I_{\mathcal{H}_1} \otimes |f'\rangle \otimes I_{\mathcal{H}_3}) \\
&= \sum_{|f\rangle, |f'\rangle \in B_2} \langle f'|f\rangle (I_{\mathcal{H}_1} \otimes \langle f| \otimes I_{\mathcal{H}_3}) X (I_{\mathcal{H}_1} \otimes |f'\rangle \otimes I_{\mathcal{H}_3}) \\
&= \sum_{|f\rangle \in B_2} (I_{\mathcal{H}_1} \otimes \langle f| \otimes I_{\mathcal{H}_3}) X (I_{\mathcal{H}_1} \otimes |f\rangle \otimes I_{\mathcal{H}_3}).
\end{aligned}$$

Thus, we can write the partial trace using any choice of orthonormal basis B of $\mathcal{H}_2$. □

Notice that $\operatorname{Tr}_{\mathcal{H}_2} = G_A$, where $A := \{I_{\mathcal{H}_1} \otimes \langle e| \otimes I_{\mathcal{H}_3} : |e\rangle \in B\}$ and B is any orthonormal basis of $\mathcal{H}_2$. Thus the partial trace is a quantum operation. Since $I_{\mathcal{H}_1} \otimes \langle v| \otimes I_{\mathcal{H}_3}$ has full row rank, for any nonzero vector $|v\rangle$, we have the following corollary of Theorem 2.14.

Corollary 3.3. *Given a Hilbert space $\mathcal{H} := \mathcal{H}_1 \otimes \mathcal{H}_2 \otimes \mathcal{H}_3$ and $X \in \mathcal{L}(\mathcal{H})$, we have the following.*

1. $\operatorname{Tr}(X) = \operatorname{Tr}(\operatorname{Tr}_{\mathcal{H}_2}(X))$,
2. $X \succeq 0 \implies \operatorname{Tr}_{\mathcal{H}_2}(X) \succeq 0$, *and*
3. $X \succ 0 \implies \operatorname{Tr}_{\mathcal{H}_2}(X) \succ 0$.

Therefore, taking the partial trace of a density matrix yields another density matrix.

Suppose we have a quantum state $|\psi\rangle \in \mathcal{H} \otimes \mathcal{K}$. If we measure the subspace $\mathcal{K}$ of the state and discard the outcome, the remaining state is a mixed state over $\mathcal{H}$. Let ρ be the corresponding density matrix of this mixed state. Then the partial trace is defined so that

$$\rho = \mathrm{Tr}_{\mathcal{K}} |\psi\rangle\langle\psi|.$$

We call ρ the *reduced density operator of the state* $|\psi\rangle$ *on* $\mathcal{H}$. Thus, the partial trace is the operation of measuring a subspace and discarding the outcome. We say that a subspace of a quantum state has been "traced out" when we examine its reduced density matrix.

Since the basis for the subspace being traced out can be arbitrarily chosen, we have

$$\mathrm{Tr}_{\mathcal{H}_2}(\rho) = \mathrm{Tr}_{\mathcal{H}_2}[(I_{\mathcal{H}_1} \otimes U_{\mathcal{H}_2} \otimes I_{\mathcal{H}_3})\rho(I_{\mathcal{H}_1} \otimes U^*_{\mathcal{H}_2} \otimes I_{\mathcal{H}_3})],$$

for any unitary matrix $U_{\mathcal{H}_2}$ and any density matrix $\rho \in \mathbb{D}^{\mathcal{H}_1 \otimes \mathcal{H}_2 \otimes \mathcal{H}_3}$.

The following is a well-known result which is important in the discussion of coin-flipping (the content of Chapter 6).

Lemma 3.4 ([22, 23])**.** *Suppose we are given a density matrix* $\rho \in \mathbb{D}^{\mathcal{H}}$ *and two states* $|\psi\rangle, |\phi\rangle \in \mathcal{H} \otimes \mathcal{K}$ *satisfying* $\mathrm{Tr}_{\mathcal{K}} |\psi\rangle\langle\psi| = \mathrm{Tr}_{\mathcal{K}} |\phi\rangle\langle\phi| = \rho$*. Then there exists a unitary matrix* $U \in \mathcal{L}(\mathcal{K})$ *such that*

$$(I_{\mathcal{H}} \otimes U) |\psi\rangle = |\phi\rangle .$$

We now show that the partial trace cannot be written in a certain way.

Theorem 3.5. *Suppose we have a Hilbert space* $\mathcal{H} := \mathcal{H}_1 \otimes \mathcal{H}_2 \otimes \mathcal{H}_3$ *satisfying*

$$\dim(\mathcal{H}_1) \cdot \dim(\mathcal{H}_3) \geq 2 \quad \textit{and} \quad \dim(\mathcal{H}_2) \geq 2.$$

Then $\mathrm{Tr}_{\mathcal{H}_2}$ *cannot be written as* $\mathrm{Tr}_{\mathcal{H}_2}(\cdot) = M(\cdot)M^*$*, for a matrix* $M : \mathcal{L}(\mathcal{H}) \to \mathcal{L}(\mathcal{H}_1 \otimes \mathcal{H}_3)$*.*

Proof. Suppose we are given a Hilbert space $\mathcal{H} := \mathcal{H}_1 \otimes \mathcal{H}_2 \otimes \mathcal{H}_3$ where $h_1 := \dim(\mathcal{H}_1)$ and $h_3 := \dim(\mathcal{H}_3)$ satisfy $h_1 h_3 \geq 2$ and $h_2 := \dim(\mathcal{H}_2) \geq 2$. Assume the partial trace can be written as $M(\cdot)M^*$, for some matrix $M : \mathcal{L}(\mathcal{H}_1 \otimes \mathcal{H}_3) \to \mathcal{L}(\mathcal{H})$. Define the quantum state

$$|\psi\rangle := \frac{1}{\sqrt{2}} \left(|000\rangle + |h_1 - 1\rangle\, |h_2 - 1\rangle\, |h_3 - 1\rangle \right).$$

By assumption we know that

$$\operatorname{Tr}_{\mathcal{H}_2}|\psi\rangle\langle\psi| = M|\psi\rangle\langle\psi|M^* = (M|\psi\rangle)(M|\psi\rangle)^*,$$

which is a rank 1 matrix. We also have that

$$\begin{aligned}\operatorname{Tr}_{\mathcal{H}_2}|\psi\rangle\langle\psi| &= \sum_{i=0}^{h_2-1}(I_{\mathcal{H}_1}\otimes\langle i|\otimes I_{\mathcal{H}_3})|\psi\rangle\langle\psi|(I_{\mathcal{H}_1}\otimes|i\rangle\otimes I_{\mathcal{H}_3}) \\ &= \frac{1}{2}|0\rangle\langle 0|_{\mathcal{H}_1}\otimes|0\rangle\langle 0|_{\mathcal{H}_3}+\frac{1}{2}|h_1-1\rangle\langle h_1-1|_{\mathcal{H}_1}\otimes|h_3-1\rangle\langle h_3-1|_{\mathcal{H}_3}.\end{aligned}$$

Since $h_1h_3 \geq 2$, we know that $\frac{1}{2}|0\rangle\langle 0|_{\mathcal{H}_1}\otimes|0\rangle\langle 0|_{\mathcal{H}_3}+\frac{1}{2}|h_1-1\rangle\langle h_1-1|_{\mathcal{H}_1}\otimes|h_3-1\rangle\langle h_3-1|_{\mathcal{H}_3}$ is a rank 2 matrix, which is a contradiction. □

To give an explicit characterization of the partial trace, we need the following definition.

Definition 3.6 (p-Block Submatrix)**.** *Suppose we are given $J \subseteq \{0,1,\ldots,m-1\}$ and a matrix $X \in \mathbb{M}^{m\cdot p}$. We define the* p-block submatrix of X*, denoted $X[J]_p$, as*

$$X[J]_p := \{[X_{i,j}] : i,j \in J\} = \sum_{i,j\in J}|i\rangle\langle j|\otimes X_{i,j},$$

where $X_{i,j}$ is the i,j'th $p\times p$ block of X.

Note in the above definition we start indexing from 0 instead of 1. This eases notation since the computational basis for $\mathbb{C}^m$ is $\{|0\rangle, |1\rangle, \ldots, |m-1\rangle\}$.

Recall the definition of the *adjoint* of a linear operator.

Definition 3.7. *Given a linear operator $\mathcal{A}: V \to W$, where V and W are Hilbert spaces, the* adjoint *of $\mathcal{A}$, denoted $\mathcal{A}^*$, is defined to be the unique linear operator satisfying*

$$\langle\mathcal{A}v, w\rangle_W = \langle v, \mathcal{A}^*w\rangle_V,\ \forall v\in V,\ \forall w\in W.$$

Define $T: \mathbb{M}^{m\cdot p} \to \mathbb{M}^{p\cdot|J|}$ to be $T(X) = X[J]_p$, for some $J \subseteq \{0,1,\ldots,m-1\}$. Let $Z[J]_p^* := T^*(Z)$ and $j_0, j_1, \ldots, j_{|J|-1}$ be an ordering of indices in J. Then we can write $X = Z[J]_p^*$ as

$$[X_{i,j}] = \begin{cases}[0], & \text{if } i\notin J \text{ or } j\notin J,\\ [Z_{l,k}], & \text{if } i = j_l \text{ and } j = j_k.\end{cases}$$

We can alternatively write $X = Z[J]_p^*$ as $\sum_{i,j=0}^{m\cdot p-1}|i\rangle\langle j|\otimes X_{i,j}$, for $X_{i,j}$ above.

Example 3.8. *Given the matrix X below, we can find the 2-block submatrix $X[0,2]_2$. (Recall we start indexing from 0.)*

$$X := \left[\begin{array}{cc|cc|cc} X_{11} & X_{12} & X_{13} & X_{14} & X_{15} & X_{16} \\ X_{21} & X_{22} & X_{23} & X_{24} & X_{25} & X_{26} \\ \hline X_{31} & X_{32} & X_{33} & X_{34} & X_{35} & X_{36} \\ X_{41} & X_{42} & X_{43} & X_{44} & X_{45} & X_{46} \\ \hline X_{51} & X_{52} & X_{53} & X_{54} & X_{55} & X_{56} \\ X_{61} & X_{62} & X_{63} & X_{64} & X_{65} & X_{66} \end{array}\right], \quad X[0,2]_2 = \left[\begin{array}{cc|cc} X_{11} & X_{12} & X_{15} & X_{16} \\ X_{21} & X_{22} & X_{25} & X_{26} \\ \hline X_{51} & X_{52} & X_{55} & X_{56} \\ X_{61} & X_{62} & X_{65} & X_{66} \end{array}\right].$$

Given the matrix Z below, we can find $Z[0,2]_2^*$.

$$Z := \left[\begin{array}{cc|cc} Z_{11} & Z_{12} & Z_{13} & Z_{14} \\ Z_{21} & Z_{22} & Z_{23} & Z_{24} \\ \hline Z_{31} & Z_{32} & Z_{33} & Z_{34} \\ Z_{41} & Z_{42} & Z_{43} & Z_{44} \end{array}\right], \quad Z[0,2]_2^* = \left[\begin{array}{cc|cc|cc} Z_{11} & Z_{12} & 0 & 0 & Z_{13} & Z_{14} \\ Z_{21} & Z_{22} & 0 & 0 & Z_{23} & Z_{24} \\ \hline 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ \hline Z_{31} & Z_{32} & 0 & 0 & Z_{33} & Z_{34} \\ Z_{41} & Z_{42} & 0 & 0 & Z_{43} & Z_{44} \end{array}\right].$$

The following result is well-known and widely used in practice.

Theorem 3.9. *Suppose we are given $\mathcal{H} := \mathbb{C}^{m_1} \otimes \mathbb{C}^{m_2} \otimes \mathbb{C}^{m_3}$, where m_1,m_2, and m_3 are positive integers. Let (k,i) be the base m_2 representation of km_2+i. Then we can explicitly write the partial trace over $\mathbb{C}^{m_2}$ as*

$$\mathrm{Tr}_{\mathbb{C}^{m_2}}(X) = \sum_{i=0}^{m_2-1} X[J_i]_{m_3},$$

where $J_i := \{(k,i) : k \in \{0,1,\ldots,m_1-1\}\}$ for each $i \in \{0,1,\ldots,m_2-1\}$.

Proof. Suppose we have $\mathcal{H} := \mathbb{C}^{m_1} \otimes \mathbb{C}^{m_2} \otimes \mathbb{C}^{m_3}$. Then we can write

$$I_{m_1} \otimes \langle 0| \otimes I_{m_3} = I_{m_1} \otimes \left[\begin{array}{c|c} I_{m_3} & 0 \end{array}\right] = \left[\begin{array}{cc|cc|c} I_{m_3} & 0 & 0 & 0 & \cdots \\ \hline 0 & 0 & I_{m_3} & 0 & \cdots \\ \hline \vdots & \vdots & \vdots & \vdots & \ddots \end{array}\right],$$

and

$$I_{m_1} \otimes |0\rangle \otimes I_{m_3} = I_{m_1} \otimes \left[\begin{array}{c} I_{m_3} \\ \hline 0 \end{array}\right] = \left[\begin{array}{c|c|c} I_{m_3} & 0 & \cdots \\ 0 & 0 & \cdots \\ \hline 0 & I_{m_3} & \cdots \\ 0 & 0 & \cdots \\ \hline \vdots & \vdots & \ddots \end{array}\right].$$

By block matrix multiplication, we have that

$$(I_{m_1} \otimes \langle 0| \otimes I_{m_3}) X (I_{m_1} \otimes |0\rangle \otimes I_{m_3}) = X[J_0]_{m_3},$$

where $J_i := \{(k, i) : k \in \{0, 1, \ldots, m_1 - 1\}\}$ for each $i \in \{0, 1, \ldots, m_2 - 1\}$. Similarly, for any $i \in \{0, \ldots, m_2 - 1\}$, we have that

$$(I_{m_1} \otimes \langle i| \otimes I_{m_3}) X (I_{m_1} \otimes |i\rangle \otimes I_{m_3}) = X[J_i]_{m_3},$$

giving the result. □

We use the notation (k, i) throughout this discussion. We now give an alternative proof of Theorem 3.9 using Dirac notation.

Proof. Suppose we are given $X \in \mathcal{L}(\mathbb{C}^{m_1} \otimes \mathbb{C}^{m_2} \otimes \mathbb{C}^{m_3})$. We can write X as

$$X = \sum_{k,l=0}^{m_1-1} |k\rangle\langle l| \otimes X_{k,l},$$

where $X_{k,l}$ is the k, l'th $m_2 \cdot m_3 \times m_2 \cdot m_3$ block of X. Then, for any $i \in \{0, 1, \ldots, m_2 - 1\}$, we have

$$\begin{aligned}
&(I_{m_1} \otimes \langle i| \otimes I_{m_3}) X (I_{m_1} \otimes |i\rangle \otimes I_{m_3}) \\
&= (I_{m_1} \otimes \langle i| \otimes I_{m_3}) \left(\sum_{k,l=0}^{m_1-1} |k\rangle\langle l| \otimes X_{k,l} \right) (I_{m_1} \otimes |i\rangle \otimes I_{m_3}) \\
&= \sum_{k,l=0}^{m_1-1} |k\rangle\langle l| \otimes [(\langle i| \otimes I_{m_3}) X_{k,l} (|i\rangle \otimes I_{m_3})] \\
&= \sum_{k,l=0}^{m_1-1} |k\rangle\langle l| \otimes X_{(k,i),(l,i)} \\
&= \sum_{(k,i),(l,i) \in J_i} |k\rangle\langle l| \otimes X_{(k,i),(l,i)} \\
&= X[J_i]_{m_3},
\end{aligned}$$

where $J_i := \{(k,i) : k \in \{0,1,\ldots,m_1-1\}\}$ for each $i \in \{0,1,\ldots,m_2-1\}$. Therefore, we have

$$\mathrm{Tr}_{\mathbb{C}^{m_2}} = \sum_{i=0}^{m_2-1} (I_{m_1} \otimes \langle i| \otimes I_{m_3}) X (I_{m_1} \otimes |i\rangle \otimes I_{m_3}) = \sum_{i=0}^{m_2-1} X[J_i]_{m_3}.$$

□

To verify this result, let us examine the three-qubit state $\frac{1}{\sqrt{2}}|000\rangle + \frac{1}{\sqrt{2}}|111\rangle$. This has the 8×8 density matrix

$$X := \begin{bmatrix} 1/2 & 0 & 0 & 0 & 0 & 0 & 0 & 1/2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1/2 & 0 & 0 & 0 & 0 & 0 & 0 & 1/2 \end{bmatrix}.$$

Tracing out the second qubit yields the state described by the following density matrix

$$\sum_{i=0}^{1} X[i,i+2]_2 = \begin{bmatrix} 1/2 & & & \\ & 0 & & \\ & & 0 & \\ & & & 0 \end{bmatrix} + \begin{bmatrix} 0 & & & \\ & 0 & & \\ & & 0 & \\ & & & 1/2 \end{bmatrix} = \begin{bmatrix} 1/2 & & & \\ & 0 & & \\ & & 0 & \\ & & & 1/2 \end{bmatrix}.$$

This is precisely the density matrix of the ensemble $\{(1/2, |00\rangle), (1/2, |11\rangle)\}$. Notice we can obtain the same ensemble by measuring the middle qubit of $\frac{1}{\sqrt{2}}|000\rangle + \frac{1}{\sqrt{2}}|111\rangle$ in the computational basis.

We now state an immediate corollary of Theorem 3.9.

Corollary 3.10. *Suppose we are given the Hilbert space* $\mathcal{H} := \mathbb{C}^{m_1} \otimes \mathbb{C}^{m_2} \otimes \mathbb{C}^{m_3}$. *Then we can write the adjoint of the partial trace* $\mathrm{Tr}_{\mathbb{C}^{m_2}}$ *as*

$$\mathrm{Tr}^*_{\mathbb{C}^{m_2}}(Z) = \sum_{i=0}^{m_2-1} Z[J_i]^*_{m_3},$$

where $J_i := \{i + km_2 : k \in \{0,1,\ldots,m_1-1\}\}$ *for each* $i \in \{0,1,\ldots,m_2-1\}$.

Notice that tracing out the first part of a composite system $\mathcal{H} \otimes \mathcal{K}$ has the form $\mathrm{Tr}_{\mathcal{H}}(\cdot) := \sum_{|e\rangle \in B} (\langle e| \otimes I_{\mathcal{K}})(\cdot)(|e\rangle \otimes I_{\mathcal{K}})$, for any orthonormal basis B of $\mathcal{H}$. The adjoint of this is simply

$$\begin{aligned} \mathrm{Tr}^*_{\mathcal{H}} : Z \in \Sigma^n &\rightarrow \sum_{|e\rangle \in B} (|e\rangle \otimes I_{\mathcal{K}}) Z (\langle e| \otimes I_{\mathcal{K}}) \\ &= \sum_{|e\rangle \in B} (|e\rangle \otimes I_{\mathcal{K}})(I_1 \otimes Z)(\langle e| \otimes I_{\mathcal{K}}) \\ &= \sum_{|e\rangle \in B} |e\rangle\langle e| \otimes Z \\ &= I_{\mathcal{H}} \otimes Z. \end{aligned}$$

Therefore, we have a nice expression for the adjoint, namely $\mathrm{Tr}^*_{\mathcal{H}}(Z) = I_{\mathcal{H}} \otimes Z$. Similarly, for the partial trace over $\mathcal{K}$, we have that $\mathrm{Tr}^*_{\mathcal{K}}(Z) = Z \otimes I_{\mathcal{K}}$. We show a similar way of expressing the adjoint of the partial trace over an arbitrary tensor factor, but first some notation is needed.

Definition 3.11 (Block Kronecker Product)**.** *For a matrix A and a matrix $B \in \mathbb{M}^{m \cdot p}$, we define the* p-block Kronecker product*, denoted $A \otimes^p B$, as*

$$A \otimes^p B := \begin{bmatrix} A \otimes B_{1,1} & A \otimes B_{1,2} & A \otimes B_{1,3} & \cdots & A \otimes B_{1,m} \\ A \otimes B_{2,1} & A \otimes B_{2,2} & A \otimes B_{2,3} & \cdots & A \otimes B_{2,m} \\ A \otimes B_{3,1} & A \otimes B_{3,2} & A \otimes B_{3,3} & \cdots & A \otimes B_{3,m} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ A \otimes B_{m,1} & A \otimes B_{m,2} & A \otimes B_{m,3} & \cdots & A \otimes B_{m,m} \end{bmatrix},$$

where $B_{i,j}$ is the i,j'th $p \times p$ block of B.

Notice that the Kronecker product is a permutation of the rows and columns of the block Kronecker product. If we index the blocks of B starting from 0, we can write B as $B = \sum_{i,j=0}^{m-1} |i\rangle\langle j| \otimes B_{i,j}$, where $B_{i,j}$ is the i,j'th block of B. Then we have

$$A \otimes B = \sum_{i,j=0}^{m-1} A \otimes |i\rangle\langle j| \otimes B_{i,j}.$$

By definition of the block Kronecker product, we have that

$$A \otimes^p B = \sum_{i,j=0}^{m-1} |i\rangle\langle j| \otimes A \otimes B_{i,j}.$$

This leads to a characterization of the adjoint of the partial trace.

Theorem 3.12. *For* $\mathcal{H} := \mathbb{C}^{m_1} \otimes \mathbb{C}^{m_2} \otimes \mathbb{C}^{m_3}$*, we have*

$$\mathrm{Tr}^*_{\mathbb{C}^{m_2}}(Z) = I_{m_2} \otimes^{m_3} Z.$$

Proof. From Corollary 3.10, we know that $\mathrm{Tr}^*_{\mathbb{C}^{m_2}}(Z) = \sum_{i=0}^{m_2-1} Z[J_i]^*_{m_3}$. Let $Z_{k,l}$ be the k,l'th $m_3 \times m_3$ block of Z. Then we have

$$\begin{aligned}
\mathrm{Tr}^*_{\mathbb{C}^{m_2}}(Z) &= \sum_{i=0}^{m_2-1} Z[J_i]^*_{m_3} \\
&= \sum_{i=0}^{m_2-1} \sum_{k,l=0}^{m_1-1} |k,i\rangle\langle l,i| \otimes Z_{k,l} \\
&= \sum_{i=0}^{m_2-1} \sum_{k,l=0}^{m_1-1} |k\rangle\langle l| \otimes |i\rangle\langle i| \otimes Z_{k,l} \\
&= \sum_{k,l=0}^{m_1-1} |k\rangle\langle l| \otimes \left(\sum_{i=0}^{m_2-1} |i\rangle\langle i| \right) \otimes Z_{k,l} \\
&= \sum_{k,l=0}^{m_1-1} |k\rangle\langle l| \otimes I_{m_2} \otimes Z_{k,l} \\
&= I_{m_2} \otimes^{m_3} \left(\sum_{k,l=0}^{m_1-1} |k\rangle\langle l| \otimes Z_{k,l} \right) \\
&= I_{m_2} \otimes^{m_3} Z.
\end{aligned}$$

□

3.2 The Block Kronecker Product

Generalized Kronecker products were first introduced by Singh [33] and Tracy and Singh [34]. Koning, Neudecker, and Wansbeek [20] were the first to study the block Kronecker

product in the way it has been presented above. This section presents generalizations of some well-known results about the Kronecker product.

Definition 3.13 (Vec Operator)**.** *Given $A \in \mathbb{M}^{m,n}$, we define the* vec *operator, denoted by* $\text{vec} : \mathbb{M}^{m,n} \to \mathbb{C}^{m \cdot n}$*, where*

$$\text{vec}(A) := \begin{pmatrix} \text{col}_1(A) \\ \text{col}_2(A) \\ \vdots \\ \text{col}_n(A) \end{pmatrix} = \sum_{i=0}^{n-1} |i\rangle \otimes A\,|i\rangle\,.$$

That is, the vec *operator stacks the columns of A on top of each other in a vector.*

Below we list two properties relating the vec operator and the Kronecker product, see for example [15] or [39].

1. $\text{vec}\left(AXB^T\right) = (B \otimes A)\text{vec}(X)$,
2. $\text{Tr}\left(ABCD^T\right) = \text{vec}\left(A^T\right)^T (D \otimes B)\text{vec}(C)$.

We now define *commutation matrices.*

Definition 3.14 (Commutation Matrices)**.** *The* commutation matrix $K_{m,n} \in \mathbb{M}^{m \cdot n}$ *is the unique real-valued matrix that satisfies*

$$K_{m,n}\text{vec}(A) = \text{vec}\left(A^T\right),\ \forall A \in \mathbb{M}^{m,n}.$$

Notice that commutation matrices are permutation matrices. Since they are also real-valued, we have that

$$K_{n,m} = K_{m,n}^T = K_{m,n}^* = K_{m,n}^{-1}.$$

Commutation matrices arise in quantum computing. Consider the two quantum states $|\psi\rangle \in \mathbb{C}^m$ and $|\phi\rangle \in \mathbb{C}^n$, where

$$|\psi\rangle = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_m \end{bmatrix} \in \mathbb{C}^m.$$

Then we have

$$|\psi\rangle \otimes |\phi\rangle = \begin{bmatrix} a_1 |\phi\rangle \\ a_2 |\phi\rangle \\ \vdots \\ a_m |\phi\rangle \end{bmatrix} \text{ and } |\phi\rangle \cdot |\psi\rangle^T = \left[\begin{array}{c|c|c|c} a_1 |\phi\rangle & a_2 |\phi\rangle & \cdots & a_m |\phi\rangle \end{array} \right].$$

Therefore, $|\psi\rangle \otimes |\phi\rangle = \text{vec}\left(|\phi\rangle \cdot |\psi\rangle^T\right)$. We can write

$$\begin{aligned} K_{m,n}\left(|\psi\rangle \otimes |\phi\rangle\right) &= K_{m,n}\text{vec}\left(|\phi\rangle \cdot |\psi\rangle^T\right) \\ &= \text{vec}\left[\left(|\phi\rangle \cdot |\psi\rangle^T\right)^T\right] \\ &= \text{vec}\left(|\psi\rangle \cdot |\phi\rangle^T\right) \\ &= |\phi\rangle \otimes |\psi\rangle . \end{aligned}$$

Thus, commutation matrices swap tensor factors.

The following results generalize the well-known fact that there exists permutation matrices P_1 and P_2 such that

$$A \otimes B = P_1(B \otimes A)P_2.$$

If A and B are square, then we have $P_2 = P_1^T$.

Theorem 3.15 ([20])**.** *Suppose we have $A \in \mathbb{M}^{m,n}$ and $B \in \mathbb{M}^{a \cdot p}$. Then*

$$A \otimes^p B = (I_a \otimes K_{m,p})(B \otimes A)(I_a \otimes K_{p,n}),$$

where $K_{m,p}$ and $K_{p,n}$ are commutation matrices.

The following corollary is immediate.

Corollary 3.16. *Suppose we have a matrix A and another matrix $B \in \mathbb{M}^m$. Suppose p_1 and p_2 are positive integers dividing m. Then*

$$A \otimes^{p_1} B = P_1(A \otimes^{p_2} B)P_2,$$

for some permutation matrices P_1 and P_2. Furthermore, if A is square, then we have $P_2 = P_1^T$.

This does generalize the previously mentioned result because the Kronecker product is a special case of the block Kronecker product. We see this by writing

$$A \otimes B = A \otimes^m B, \text{ where } B \in \mathbb{M}^m.$$

A generalization of the vec operator is defined below.

Definition 3.17 (Block Vec Operator)**.** *For $A \in \mathbb{M}^{mp,np}$, we define the* block vec operator*, denoted by* $\mathrm{vec}^{\mathrm{p}} : \mathbb{M}^{mp,np} \to \mathbb{C}^{mnp^2}$*, as*

$$\mathrm{vec}^{\mathrm{p}}(A) := \begin{pmatrix} \mathrm{vec}(A_{1,1}) \\ \mathrm{vec}(A_{1,2}) \\ \vdots \\ \mathrm{vec}(A_{1,n}) \\ \mathrm{vec}(A_{2,1}) \\ \mathrm{vec}(A_{2,2}) \\ \vdots \\ \mathrm{vec}(A_{m,n}) \end{pmatrix},$$

where $A_{i,j}$ is the i,j'th $p \times p$ block of A.

Notice the block vec operator lists the vec operator applied to the blocks of A along its partitioned rows. We can similarly define the block vec operator using Dirac notation as follows. Given a matrix $A \in \mathbb{M}^{mp,np}$, we know that

$$A = \sum_{i=1}^{m} \sum_{j=1}^{n} |i\rangle\langle j| \otimes A_{i,j},$$

where $A_{i,j}$ is the i,j'th $p \times p$ block of A. Then we can write

$$\begin{aligned} \mathrm{vec}^{\mathrm{p}}(A) &= \sum_{i=1}^{m} \sum_{j=1}^{n} |i\rangle |j\rangle \otimes \mathrm{vec}(A_{i,j}) \\ &= \sum_{i=1}^{m} \sum_{j=1}^{n} |i\rangle |j\rangle \otimes \left(\sum_{k=0}^{p-1} |k\rangle \otimes A_{i,j} |k\rangle \right) \\ &= \sum_{i=1}^{m} \sum_{j=1}^{n} \sum_{k=0}^{p-1} |i\rangle |j\rangle |k\rangle \otimes A_{i,j} |k\rangle . \end{aligned}$$

We now list some basic properties of the block Kronecker product, most are immediate consequences of Theorem 3.15.

Properties 3.18 ([20])**.** *Given $A, A_1, A_2 \in \mathbb{M}^n$, $B, B_1, B_2 \in \mathbb{M}^{m\cdot p}$, and $X \in \mathbb{M}^{m\cdot p,n}$, we have the following properties.*

1. $\mathrm{vec^p}\,(BXA^T) = (A \otimes^p B)\mathrm{vec^p}\,(X)$, *if* $p|n$,
2. $(A_1 + A_2) \otimes^p B = A_1 \otimes^p B + A_2 \otimes^p B$,
3. $A \otimes^p (B_1 + B_2) = A \otimes^p B_1 + A \otimes^p B_2$,
4. $(A \otimes^p B)^T = (A^T \otimes^p B^T)$,
5. $(A_1 \otimes^p B_1)(A_2 \otimes^p B_2) = (A_1A_2 \otimes^p B_1B_2)$,
6. $(A \otimes^p B) = P(B \otimes^{p'} A)P^T$, *for some permutation matrix* P, *if* $p'|n$,
7. $\det(A \otimes^p B) = \det(A)^{m\cdot p} \cdot \det(B)^n$,
8. $\mathrm{Tr}\,(A \otimes^p B) = \mathrm{Tr}\,(A) \cdot \mathrm{Tr}\,(B)$,
9. *If* A *has eigenvalue* λ *and* B *has eigenvalue* μ *then* $A \otimes^p B$ *has eigenvalue* $\lambda\mu$,
10. $\lambda(A \otimes^p B) = \lambda(A \otimes B) = \lambda(B \otimes A) = \lambda(B \otimes^{p'} A)$, *if* $p'|n$.

As a result of the last property above, we have that for all $A \in \mathbb{M}^n$, $B \in \mathbb{M}^m$,

$$A \otimes^{p_1} B \succeq 0 \iff A \otimes B \succeq 0 \iff B \otimes A \succeq 0 \iff B \otimes^{p_2} A \succeq 0,$$

for all $p_1|m$ and $p_2|n$. Another property to note is $(A \otimes^p B)^* = (A^* \otimes^p B^*)$, which is important since we are examining matrices with complex entries.

Recall that the partial trace is the operator for finding reduced density matrices. In the next chapter we define a "closeness" measure for density matrices called the *fidelity*.

Chapter 4

Survey of Fidelity

4.1 Fidelity of Quantum States

Recall that a Hermitian matrix is always orthogonally diagonalizable. That is, given any $X \in \Sigma^n$, we can write it as

$$X = QD_XQ^*,$$

where Q is unitary and D_X is the diagonal matrix whose diagonal is $\lambda(X)$. If $X \in \Sigma^n_+$, then $\lambda(X) \geq 0$, and we can define the *square root* of X as

$$X^{1/2} := QD_X^{1/2}Q^*,$$

where $D_X^{1/2}$ is defined as the diagonal matrix of the nonnegative square roots of the diagonal elements of D_X. Notice that we do indeed have $X^{1/2}X^{1/2} = X$.

Definition 4.1 (Purification)**.** *A* purification *of a density matrix $\rho \in \mathbb{D}^{\mathcal{H}}$ is a pure state $|\psi\rangle$ in a larger Hilbert space $\mathcal{H} \otimes \mathcal{H}'$ such that*

$$\operatorname{Tr}_{\mathcal{H}'} |\psi\rangle\langle\psi| = \rho.$$

For example, suppose we have the density matrix

$$\rho := \alpha_0|\xi_0\rangle\langle\xi_0| + \alpha_1|\xi_1\rangle\langle\xi_1| \in \mathbb{D}^{\mathcal{H}}.$$

We can define a purification of ρ in $\mathcal{H} \otimes \mathbb{C}^2$ as

$$|\rho^{pure}\rangle := \sqrt{\alpha_0}\,|\xi_0\rangle\,|0\rangle + \sqrt{\alpha_1}\,|\xi_1\rangle\,|1\rangle \in \mathbb{D}^{\mathcal{H}\otimes\mathbb{C}^2},$$

since $\mathrm{Tr}_{\mathbb{C}^2}|\rho^{pure}\rangle\langle\rho^{pure}| = \rho$.

Now we can define the *fidelity*.

Definition 4.2 (Fidelity)**.** *Suppose we are given two density matrices ρ_1, ρ_2 over a Hilbert space $\mathcal{H}$. We define the* fidelity*, denoted $F(\rho_1, \rho_2)$, as*

$$F(\rho_1, \rho_2) := \max_{|\psi_1\rangle, |\psi_2\rangle} |\langle\psi_1|\psi_2\rangle|,$$

where $|\psi_1\rangle$ and $|\psi_2\rangle$ are purifications of ρ_1 and ρ_2 in $\mathcal{H} \otimes \mathcal{K}$, respectively.

The fidelity is a measure of "closeness" for density matrices. Before we illustrate this, we need the following definition.

Definition 4.3. *Given two matrices $A, B \in \Sigma^n$, we say that they are* orthogonal*, denoted $A \perp B$, if*

$$AB = BA = 0.$$

Now we state the following well-known result.

Lemma 4.4. *Given any two density matrices $\rho_1, \rho_2 \in \mathbb{D}^{\mathcal{H}}$, we have*

$$0 \leq F(\rho_1, \rho_2) \leq 1.$$

Moreover, we have

$$F(\rho_1, \rho_2) = 0 \iff \rho_1 \perp \rho_2 \quad \textit{and} \quad F(\rho_1, \rho_2) = 1 \iff \rho_1 = \rho_2.$$

Note that $F(\rho, \rho) = 1$ for every density matrix ρ, thus fidelity is not a metric.

We have the following result which is similar in spirit to Lemma 3.4.

Lemma 4.5 ([17])**.** *Suppose $\rho \in \mathbb{D}^{\mathcal{H}}$ has purification $|\psi\rangle \in \mathcal{H} \otimes \mathcal{K}$ and $\sigma \in \mathbb{D}^{\mathcal{H}}$ has purification $|\phi\rangle \in \mathcal{H} \otimes \mathcal{K}$. Then*

$$F(\rho, \sigma) = \sup_U \langle\psi| I_{\mathcal{H}} \otimes U |\phi\rangle,$$

where U is a unitary matrix in $\mathcal{L}(\mathcal{K})$. Furthermore, if $\dim(\mathcal{K}) \geq \dim(\mathcal{H})$, then the maximum is attained.

If the maximum is attained, we call the unitary in the above lemma the *fidelity achieving unitary.*

An immediate consequence of the definition of the fidelity is the following lemma.

Lemma 4.6. *Given any two density matrices ρ_1 and ρ_2 over a Hilbert space $\mathcal{H} \otimes \mathcal{K}$, we have*

$$F(\rho_1, \rho_2) \leq F(\mathrm{Tr}_{\mathcal{K}}(\rho_1), \mathrm{Tr}_{\mathcal{K}}(\rho_2)).$$

The following well-known result shows that the fidelity is strongly concave, see for instance [31].

Theorem 4.7. *Suppose we have two ensembles of density matrices $\{p_i, \rho_i\}$ and $\{q_i, \sigma_i\}$. Then we have*

$$F\left(\sum_i p_i\rho_i, \sum_i q_i\sigma_i\right) \geq \sum_i \sqrt{p_i q_i} F(\rho_i, \sigma_i).$$

This naturally leads to the weaker concavity conditions below.

Corollary 4.8. *Suppose we have two ensembles of density matrices $\{p_i, \rho_i\}$, $\{p_i, \sigma_i\}$, and a density matrix σ. Then we have*

1. $F\left(\sum_i p_i\rho_i, \sum_i p_i\sigma_i\right) \geq \sum_i p_i F(\rho_i, \sigma_i)$, *and*
2. $F\left(\sum_i p_i\rho_i, \sigma\right) \geq \sum_i \sqrt{p_i} F(\rho_i, \sigma)$.

There is an equivalent way of writing the fidelity, but first we need the following definition.

Definition 4.9 (Trace Norm)**.** *Given a matrix $A \in \mathbb{M}^{m,n}$, we define the* trace norm *of A, denoted $\|A\|_t$, as*

$$\|A\|_t := \mathrm{Tr}\left[(A^*A)^{1/2}\right].$$

We have the following equivalent description of the fidelity by Uhlmann [40] using the trace norm.

Lemma 4.10 ([40]). *Given any two density matrices $\rho_1, \rho_2 \in \mathbb{D}^n$, we have*

$$F(\rho_1, \rho_2) = \mathrm{Tr}\left[\left(\rho_1^{1/2}\rho_2\rho_1^{1/2}\right)^{1/2}\right] = \|\sqrt{\rho_1}\sqrt{\rho_2}\|_t.$$

The above is an alternative definition of the fidelity and is much easier to use in many cases.

We now state a result by Fuchs and van de Graaf.

Lemma 4.11 ([12]). *Given any two density matrices $\rho_1, \rho_2 \in \mathbb{D}^{\mathcal{H}}$, we have*

$$1 - F(\rho_1, \rho_2) \leq \frac{1}{2}\|\rho_1 - \rho_2\|_t \leq \sqrt{1 - F(\rho_1, \rho_2)^2}.$$

The following two properties are well-known.

Theorem 4.12.

1. *Fidelity is multiplicative under Kronecker products, i.e.,*

$$F(X_1 \otimes X_2, Y_1 \otimes Y_2) = F(X_1, Y_1)F(X_2, Y_2),\ \forall X_1, X_2, Y_1, Y_2 \in \mathbb{D}^{\mathcal{H}}.$$

2. *Fidelity is invariant under unitary transformations, i.e.,*

$$F(UXU^*, UYU^*) = F(X, Y),\ \forall X_1, X_2, Y_1, Y_2 \in \mathbb{D}^{\mathcal{H}}.$$

Proof. Assume we are given $X_1, X_2, Y_1, Y_2 \in \mathbb{D}^{\mathcal{H}}$. Then we can write the following.

1. $$\begin{aligned} F(X_1 \otimes X_2, Y_1 \otimes Y_2) &= \mathrm{Tr}\left[\left(\left(X_1^{1/2} \otimes X_2^{1/2}\right)(Y_1 \otimes Y_2)\left(X_1^{1/2} \otimes X_2^{1/2}\right)\right)^{1/2}\right] \\ &= \mathrm{Tr}\left[\left(X_1^{1/2}Y_1X_1^{1/2} \otimes X_2^{1/2}Y_2X_2^{1/2}\right)^{1/2}\right] \\ &= \mathrm{Tr}\left[\left(X_1^{1/2}Y_1X_1^{1/2}\right)^{1/2} \otimes \left(X_2^{1/2}Y_2X_2^{1/2}\right)^{1/2}\right] \\ &= \mathrm{Tr}\left[\left(X_1^{1/2}Y_1X_1^{1/2}\right)^{1/2}\right] \mathrm{Tr}\left[\left(X_2^{1/2}Y_2X_2^{1/2}\right)^{1/2}\right] \\ &= F(X_1, Y_1)F(X_2, Y_2) \end{aligned}$$

$$\begin{aligned}
2.\quad F(UXU^*, UYU^*) &= \mathrm{Tr}\left[\left((UXU^*)^{1/2}(UYU^*)(UXU^*)^{1/2}\right)^{1/2}\right] \\
&= \mathrm{Tr}\left[\left((UX^{1/2}U^*)(UYU^*)(UX^{1/2}U^*)\right)^{1/2}\right] \\
&= \mathrm{Tr}\left[\left(UX^{1/2}YX^{1/2}U^*\right)^{1/2}\right] \\
&= \mathrm{Tr}\left[U\left(X^{1/2}YX^{1/2}\right)^{1/2}U^*\right] \\
&= \mathrm{Tr}\left[\left(X^{1/2}YX^{1/2}\right)^{1/2}\right] \\
&= F(X,Y).
\end{aligned}$$

□

4.2 Positive Definite Matrix Means

Definition 4.13 (Matrix Means)**.** *Given two matrices, $X, S \in \mathbb{M}^n$, we define their* arithmetic mean *as*

$$AM(X,S) := \frac{1}{2}(X+S).$$

Given two positive definite matrices $X, S \in \Sigma_{++}^n$, we define their geometric mean *as*

$$GM(X,S) := X^{1/2}\left(X^{-1/2}SX^{-1/2}\right)^{1/2}X^{1/2},$$

and their harmonic mean *as*

$$HM(X,S) := 2\left(X^{-1}+S^{-1}\right)^{-1}.$$

The following matrix inequalities are well-known in Matrix Theory, see for instance [6].

Lemma 4.14. *Given any positive definite matrices $X, S \in \Sigma_{++}^n$, we have the following* harmonic-geometric-arithmetic mean inequality

$$2\left(X^{-1}+S^{-1}\right)^{-1} \preceq X^{1/2}\left(X^{-1/2}SX^{-1/2}\right)^{1/2}X^{1/2} \preceq \frac{1}{2}(X+S).$$

We note here that the geometric mean of two density matrices is not always a density matrix. Although $X, S \succeq 0$ implies $GM(X,S) \succeq 0$ by the harmonic-geometric-arithmetic mean inequality, it may not be the case that the trace is preserved. Consider the example

$$X := \begin{bmatrix} 3/4 & 0 \\ 0 & 1/4 \end{bmatrix} \quad \text{and} \quad S := \begin{bmatrix} 1/4 & 0 \\ 0 & 3/4 \end{bmatrix}.$$

Both X and S are positive semidefinite and have trace 1. Since X and S are diagonal, we can write the geometric mean as

$$GM(X,S) = X^{1/2}S^{1/2} = \begin{bmatrix} \sqrt{3}/2 & 0 \\ 0 & 1/2 \end{bmatrix} \begin{bmatrix} 1/2 & 0 \\ 0 & \sqrt{3}/2 \end{bmatrix} = \begin{bmatrix} \sqrt{3}/4 & 0 \\ 0 & \sqrt{3}/4 \end{bmatrix}.$$

Therefore $\mathrm{Tr}\,(GM(X,S)) = \sqrt{3}/2 < 1$.

Matrix geometric means and related concepts also arise in *semidefinite optimization* (the topic of the next chapter). For instance, in interior-point methods for semidefinite optimization, given $X, S \in \Sigma_{++}^n$, we look for a $T \in \mathrm{Aut}\,(\Sigma_{++}^n)$, $T = T^*$ such that $T(S) = T^{-1}(X)$. Such a T can be described by a $W \in \Sigma_{++}^n$ such that $T(\cdot) = W(\cdot)W$. I.e., $WSW = W^{-1}XW^{-1}$. The unique solution of this matrix equation is given by $W \in \Sigma_{++}^n$ such that $W^2 = GM(S^{-1}, X)$, which leads to the so-called Nesterov-Todd scaling [30].

Chapter 5

Semidefinite Programming

5.1 Semidefinite Programming and Duality

A semidefinite program, referred to as an SDP, is a linear optimization problem over positive semidefinite matrices. In *standard form*, it is defined as

$$\begin{aligned} (P) \quad \inf \quad & \langle C, X \rangle \\ \text{subject to} \quad & \mathcal{A}(X) = b, \\ & X \succeq 0, \end{aligned}$$

where $X \in \Sigma^n$ is the variable, $C \in \Sigma^n$, $b \in \mathbb{C}^m$, and $\mathcal{A} : \Sigma^n \to \mathbb{C}^m$ is a linear operator. The object being minimized is $\langle C, X \rangle$ which we call the *objective function*. The objective function is real-valued since X and C are Hermitian. The conditions $\mathcal{A}(X) = b$ and $X \succeq 0$ are called the *constraints*.

We say that an SDP is *feasible* if there exists an $\bar{X}$ that satisfies the constraints, that is, $\exists \bar{X} \succeq 0$ such that $\mathcal{A}(\bar{X}) = b$; such an $\bar{X}$ is called a *feasible solution*. If the SDP is not feasible, we say that it is *infeasible*. A *strictly feasible solution* is a feasible solution $\bar{X}$ where $\bar{X}$ also satisfies $\bar{X} \succ 0$. We say that the SDP is *unbounded* if there exists $\{\bar{X}_k : k \in \mathbb{N}, \bar{X}_k \text{ feasible}\}$ such that $\langle C, \bar{X}_k \rangle \to -\infty$ as $k \to \infty$. We say that a feasible solution $\bar{X}$ is *optimal* if $\langle C, \bar{X} \rangle \leq \langle C, X \rangle$ for all feasible X and that the SDP *attains its optimal value* if an optimal solution exists. It may be the case that an SDP is feasible, not unbounded, and does not attain its optimal value; thus we must examine the infimum of

the objective function in this case.

Given an SDP in standard form, we define its *dual*, or *dual problem*, as

$$\begin{array}{lll} (D) & \sup & \langle b, y \rangle \\ \text{subject to} & & \mathcal{A}^*(y) + S = C, \\ & & S \succeq 0, \end{array}$$

where $y \in \mathbb{C}^m$ and $S \in \Sigma^n$ are the variables and $\mathcal{A}^*$ is the adjoint of the linear operator $\mathcal{A}$. The fact that $\langle b, y \rangle$ is real-valued is implied by the proof of Theorem 5.1, below.

When we examine the dual of an SDP, we call the original problem the *primal*, or *primal problem*, and the pair of problems (P, D) a *primal-dual pair*. An easy exercise is to verify that the dual of the dual problem is the primal problem. We similarly define feasible solutions for the dual problem and note for a feasible solution $(\bar{y}, \bar{S})$ to be a strictly feasible solution we require that $\bar{S} \succ 0$.

The following duality theorems and corollaries play a critical role in our analysis of coin-flipping protocols in later chapters. For a good reference of these results see [41, 38].

Theorem 5.1 (Weak Duality)**.** *Suppose (P, D) is a primal-dual pair and we are given $\bar{X}$ which is feasible in the primal problem and $(\bar{y}, \bar{S})$ which is feasible in the dual problem. Then*

$$\langle C, \bar{X} \rangle \geq \langle b, \bar{y} \rangle .$$

Proof. Suppose we are given primal feasible $\bar{X}$ and dual feasible $(\bar{y}, \bar{S})$. Then we have

$$\begin{aligned} \left\langle C, \bar{X} \right\rangle - \langle b, \bar{y} \rangle &= \left\langle C, \bar{X} \right\rangle - \left\langle \mathcal{A}(\bar{X}), \bar{y} \right\rangle \\ &= \left\langle C, \bar{X} \right\rangle - \left\langle \bar{X}, \mathcal{A}^*(\bar{y}) \right\rangle \\ &= \left\langle C, \bar{X} \right\rangle - \left\langle \bar{X}, C - \bar{S} \right\rangle \\ &= \left\langle \bar{S}, \bar{X} \right\rangle . \end{aligned}$$

The last equality above holds since $\langle X, C \rangle = \langle C, X \rangle$ because $X, C \in \Sigma^n$. Note that $\left\langle \bar{S}, \bar{X} \right\rangle \geq 0$ since $\bar{X}, \bar{S} \in \Sigma^n_+$, which completes the proof. □

The above proof shows that given a primal feasible solution $\bar{X}$ and a dual feasible solution $(\bar{y}, \bar{S})$, the difference in objective values is $\left\langle \bar{X}, \bar{S} \right\rangle$. This brings us to the following definition.

Definition 5.2 (Duality Gap)**.** *The* duality gap *of a primal-dual pair* (P, D) *is defined as*

$$\inf\{\langle \bar{X}, \bar{S} \rangle : \bar{X} \text{ is feasible in } (P),\ \exists \bar{y} \text{ such that } (\bar{y}, \bar{S}) \text{ is feasible in } (D)\}.$$

We say that a primal-dual pair have *zero duality gap* if the duality gap, as defined above, is 0. A zero duality gap means that both (P) and (D) have feasible solutions and the optimal values of (P) and (D) are the same.

Theorem 5.3 (Strong Duality)**.** *Suppose* (D) *has a strictly feasible solution and the objective value is bounded from above. Then there is zero duality gap and* (P) *attains its optimal value.*

We have the following two corollaries of the strong duality theorem.

Corollary 5.4. *Suppose* (D) *has a strictly feasible solution and* (P) *has a feasible solution. Then there is zero duality gap and* (P) *attains its optimal value.*

Corollary 5.5. *Suppose* (P) *and* (D) *both have strictly feasible solutions. Then there is zero duality gap and both* (P) *and* (D) *attain their optimal values.*

Strong duality says that under certain conditions, we can solve for the optimal value of an SDP by solving that of its dual, which may be advantageous to analyze instead of the original problem.

5.2 Applications of Semidefinite Programming in Quantum Computing

We now present some applications of semidefinite programming in quantum computing. First, we need some definitions.

Definition 5.6 (Hadamard Product)**.** *We define the* Hadamard product *of* $A, B \in \mathbb{M}^{m,n}$ *as*

$$[(A \circ B)_{i,j}] := [A_{i,j} B_{i,j}]\,.$$

Notice that the Hadamard product is only defined for matrices having the same dimensions. Given $A, B \in \Sigma^n$, we can write $\langle A, B\rangle = \langle e| A \circ B |e\rangle$, where $|e\rangle$ is the n-dimensional vector of all ones. From this, we see that

$$T : X \in \Sigma^n \to A \circ X \in \Sigma^n,$$

is a self-adjoint operator, for any $A \in \Sigma^n$.

We say that a system $\mathcal{A}(Y) \preceq C \in \Sigma^n$ is *almost feasible* if for all $\epsilon > 0$, $\exists E$ with $\|E\| := \sqrt{\langle E, E\rangle} < \epsilon$ such that $\mathcal{A}(Y) \preceq C + E$ is feasible.

The following is a well-known result about alternative systems. See for example [38].

Lemma 5.7. *There exists $D \succeq 0$ satisfying $\mathcal{A}(D) = 0$ and $\langle C, D\rangle < 0$ if and only if $\mathcal{A}^*(y) \preceq C$ is not almost feasible.*

Using this lemma, we have an alternative proof of Lemma 2 in [7].

Theorem 5.8 ([7])**.** *Let $\Sigma^n(\mathbb{R})$ denote the set of real-valued symmetric matrices and let $\Sigma^n_+(\mathbb{R})$ denote the set of real-valued positive semidefinite matrices. Suppose we are given the following two systems.*

$$\begin{aligned}(I)\quad & \sum_{j=1}^{m} A_{i,j} \circ X_j = B,\ \forall i \in \{1, \ldots, k\},\\ & X_j \in \Sigma^n_+(\mathbb{R}),\ \forall j \in \{1, \ldots, m\}.\end{aligned}$$

$$\begin{aligned}(II)\quad & \sum_{i=1}^{k} \langle B, Y_i\rangle > 0,\\ & \sum_{i=1}^{k} Y_i \circ A_{i,j} \preceq 0,\ \forall j \in \{1, \ldots, m\},\\ & Y_i \in \Sigma^n(\mathbb{R}),\ \forall i \in \{1, \ldots, k\}.\end{aligned}$$

Suppose $X_1 = \cdots = X_m = 0$ is the only solution to (I) when $B = 0$. Then exactly one of (I) or (II) has a solution.

Proof. Consider the following primal-dual pair

$$\begin{aligned} (P) \quad \inf \quad & \sum_{j=1}^{m} \langle C, X_j \rangle \\ \text{subject to} \quad & \sum_{j=1}^{m} A_{i,j} \circ X_j = B, \ \forall i \in \{1, \ldots, k\}, \\ & X_j \in \Sigma_+^n(\mathbb{R}), \ \forall j \in \{1, \ldots, m\}, \end{aligned}$$

and

$$\begin{aligned} (D) \quad \sup \quad & \sum_{i=1}^{k} \langle B, Y_i \rangle \\ \text{subject to} \quad & \sum_{i=1}^{k} A_{i,j} \circ Y_i \preceq C, \ \forall j \in \{1, \ldots, m\}, \\ & Y_i \in \Sigma^n(\mathbb{R}), \ \forall i \in \{1, \ldots, k\}. \end{aligned}$$

Since $X_1 = \cdots = X_m = 0$ is the only feasible solution to (P) when $B = 0$, by hypothesis, we know that there does not exist a feasible solution $(X_1, \ldots, X_m)$ such that $\sum_{j=1}^{m} \langle C, X_j \rangle < 0$. Thus, for any choice of C, we have that the dual is almost feasible. Choosing $C = -I_n$ implies that there exists $Y_1, \ldots, Y_k$ such that $\sum_{i=1}^{k} A_{i,j} \circ Y_i \prec 0$, for all $j \in \{1, \ldots, m\}$. Now, choose $C = 0$ in the above primal-dual pair and consider two cases. Case 1, if (II) is infeasible, then there does not exist $Y_1, \ldots, Y_k$ such that $\sum_{i=1}^{k} A_{i,j} \circ Y_i \preceq C$, for all $j \in \{1, \ldots, m\}$, and $\sum_{i=1}^{k} \langle B, Y_i \rangle > 0$. Then the objective value of (D) is bounded from above. This implies that (P) attains its optimal value by Theorem 5.3 which implies the weaker condition that (P), and therefore (I), is feasible. I.e., $\exists X_1, \ldots, X_m \in \Sigma_+^n(\mathbb{R})$ such that $\sum_{j=1}^{m} A_{i,j} \circ X_j = B, \ \forall i \in \{1, \ldots, k\}$. Case 2, if (II) is feasible, then $\exists Y_1, \ldots, Y_k$ such that $\sum_{i=1}^{k} A_{i,j} \circ Y_i \preceq C, \ \forall j \in \{1, \ldots, m\}$, and $\sum_{i=1}^{k} \langle B, Y_i \rangle > 0$. Assuming that (I) is also feasible, we have that the objective value of the primal is 0 since $C = 0$. By weak duality, we get the following contradiction

$$0 = \sum_{j=1}^{m} \langle C, X_j \rangle \geq \sum_{i=1}^{k} \langle B, Y_i \rangle > 0.$$

□

We now discuss an application of semidefinite programming in quantum complexity theory. Quantum complexity theory is the study of finding lower bounds on the complexity

of quantum algorithms. There are many methods for finding lower bounds including the *weighted adversary* method [4] and *spectral adversary* method [7]. Adversary methods give lower bounds on the complexity of calculating a function f. For a function f, the weighted adversary method is posed as an optimization problem having optimal value $\mathrm{WA}(f)$. This implies the complexity of calculating f is $\Omega(\mathrm{WA}(f))$. Spalek and Szegedy [35] prove that all of the adversary methods known are equivalent, that is, they all give the same lower bound. As part of their proof, they show that a specific adversary method, posed as an optimization problem, is equivalent to a semidefinite program. That is, there is a correspondence between feasible solutions and they have the same optimal value. They examine the dual of the SDP and prove a zero duality gap by showing a strictly feasible solution for the dual. They next show that the dual, being an adversary method itself, is equivalent to another adversary method. This proves that these four methods give the same value and, therefore, the same lower bound. The key observation is that by examining the dual, they have a new approach to analyzing the problem.

The next chapter introduces coin-flipping, cheating strategies, and optimal cheating probabilities. We examine how the results in this chapter play a role in the analysis of coin-flipping protocols (the content of Chapter 7).

Chapter 6

Quantum Coin-Flipping

Coin-flipping is the cryptographic task of two mistrustful parties generating a random bit. If one party is dishonest, and deviates from protocol, he/she could try to cheat by trying to increase the probability of a desired outcome. A protocol should be able to withstand such dishonesty by not allowing a party to force a desired outcome with probability 1. This is different than many cryptographic protocols in the sense that the security we seek in a protocol is to protect an honest party from the other dishonest party.

6.1 Definitions and Notation

Before we examine coin-flipping, we set up a general framework in which the protocol takes place, and how the two parties communicate. For this discussion of coin-flipping, we call the two parties *Alice* and *Bob* and the setting is as follows. Alice and Bob have their own private spaces and they each have access to a *message space* which they use to alternate sending messages to each other. One may think of the private spaces as desktop computers and the message space as email. After they finish communicating they determine the result of the coin-flip.

We set this up in the following framework. Alice and Bob's private spaces are the state spaces $\mathcal{A}$ and $\mathcal{B}$, respectively. The message space is the state space $\mathcal{M}$. They initially share a quantum state prepared as

$$|\psi_{\text{start}}\rangle := |0\rangle_{\mathcal{A}} |0\rangle_{\mathcal{M}} |0\rangle_{\mathcal{B}} .$$

Alice has a set of unitary operations $\{U_{A,i} \in \mathcal{L}(\mathcal{A} \otimes \mathcal{M}) : i \in \{1, \ldots, N\}\}$ and Bob has another set $\{U_{B,i} \in \mathcal{L}(\mathcal{M} \otimes \mathcal{B}) : i \in \{1, \ldots, N\}\}$ of unitary operations. They apply these unitary operations alternately to the appropriate part of the state, starting with $U_{A,1}$. These are the messages of the protocol. Once all the unitary operations are applied, Alice and Bob each measure the part of the state in their respective private spaces using a projective measurement. This determines the resulting coin-flip. The projective measurements for each party have the form

$$\{\Pi_0, \Pi_1, \Pi_{Abort} := I - \Pi_{A,0} - \Pi_{A,1}\},$$

where Π_x corresponds to the outcome $x \in \{0, 1\}$, and all projectors are mutually orthogonal.

This formulation also includes classical protocols. A classical coin-flipping protocol has private and message spaces of the form $\mathbb{C}^m$, for some m, and all the unitary operations are commutation or "swap" matrices. (Recall from Section 3.2 that commutation matrices swap tensor factors.)

To fully describe a coin-flipping protocol, we define the following tuple.

Definition 6.1 (Coin-Flipping Tuple)**.** *Given a protocol, we collect all the unitary operations and projections in a* protocol-tuple

$$\mathcal{P} := (U_{A,1}, \ldots, U_{A,N}, U_{B,1}, \ldots, U_{B,N}, \Pi_{A,0}, \Pi_{A,1}, \Pi_{B,0}, \Pi_{B,1}).$$

Not any such tuple is a coin-flipping protocol however. The tuple needs to satisfy two conditions. Firstly, we need Alice and Bob to output the same bit when they follow the protocol. Let $|\psi_{\text{final}}\rangle$ be the final state shared by Alice and Bob at the end of the protocol. We need

$$(\Pi_{A,x} \otimes I_{\mathcal{M}} \otimes I_{\mathcal{B}}) |\psi_{\text{final}}\rangle = (I_{\mathcal{A}} \otimes I_{\mathcal{M}} \otimes \Pi_{B,x}) |\psi_{\text{final}}\rangle \,, \ \forall x \in \{0, 1\},$$

to ensure this condition. Secondly, we need the output bit to be randomly generated, that is, having an equal probability of being a 0 or 1. To ensure this, we require

$$\|(\Pi_{A,0} \otimes I_{\mathcal{M}} \otimes I_{\mathcal{B}}) |\psi_{\text{final}}\rangle\|^2 = \|(\Pi_{A,1} \otimes I_{\mathcal{M}} \otimes I_{\mathcal{B}}) |\psi_{\text{final}}\rangle\|^2 = 1/2.$$

Note that we do not need the analogous condition for Bob since it is redundant assuming the first condition.

We collect the two conditions below.

Conditions 6.2.

1. $(\Pi_{A,x} \otimes I_{\mathcal{M}} \otimes I_{\mathcal{B}}) |\psi_{final}\rangle = (I_{\mathcal{A}} \otimes I_{\mathcal{M}} \otimes \Pi_{B,x}) |\psi_{final}\rangle \, , \; \forall x \in \{0,1\}$,
2. $\|(\Pi_{A,0} \otimes I_{\mathcal{M}} \otimes I_{\mathcal{B}}) |\psi_{final}\rangle\|^2 = \|(\Pi_{A,1} \otimes I_{\mathcal{M}} \otimes I_{\mathcal{B}}) |\psi_{final}\rangle\|^2 = 1/2$.

If Bob and Alice trust each other and follow a protocol honestly, they do not need complicated protocols. Indeed, one party could just let the other flip a coin and share the result. This is obviously not useful if the two parties do not trust each other. In simple protocols it could be the case that Bob or Alice can bias the outcome by forcing a desired outcome with probability higher than 1/2. Before analyzing the security of coin-flipping protocols, we need to introduce some terminology and discuss the limitations on cheating.

We do not assume any computational limitations on cheating, anything within the laws of quantum mechanics is allowed. We do however assume for this discussion that if a party, say Bob, is dishonest then he cannot perform any operation on Alice's private space $\mathcal{A}$ or change $\mathcal{M}$ to another message space. What he can do is prepare any private space $\widehat{\mathcal{B}}$ and apply a possibly different unitary on $\mathcal{M} \otimes \widehat{\mathcal{B}}$ when it is his turn to apply a unitary in the protocol. We also assume that he can apply an extra unitary of his choosing at the beginning of the protocol, before any communication occurs. A cheating strategy for Bob is comprised of his private space $\widehat{\mathcal{B}}$ and his set of unitary operations he applies instead of the ones prescribed in the protocol. Notice he does not need to perform any measurements. Since they must output the same bit, the result of Alice's measurement is the outcome of the protocol, which he accepts. Recall from Chapter 2 that this is equivalent to Bob performing a quantum operation on $\mathcal{M}$.

Definition 6.3 (Cheating Strategy)**.** *Assume Alice is honest and Bob is dishonest. Assume Bob also has a private space $\widehat{\mathcal{B}}$ which is unknown to Alice. Given a protocol-tuple*

$$\mathcal{P} := \{U_{A,1}, \ldots, U_{A,N}, U_{B,1}, \ldots, U_{B,N}, \Pi_{A,0}, \Pi_{A,1}, \Pi_{B,0}, \Pi_{B,1}\}$$

we define a cheating strategy *via a* cheating-tuple*, which is defined as*

$$\mathcal{P}_B := \{\widehat{U}_{B,0}, \widehat{U}_{B,1}, \ldots, \widehat{U}_{B,N}, \Pi_{A,0}, \Pi_{A,1}\},$$

where $\widehat{U}_{B,j}$ is a unitary operator acting on $\mathcal{M} \otimes \widehat{\mathcal{B}}$, for $j \in \{0,1,\ldots,N\}$. During the protocol, Bob changes the starting state of $\mathcal{M} \otimes \widehat{\mathcal{B}}$ using $\widehat{U}_{B,0}$ and then applies $\widehat{U}_{B,j}$ instead of $U_{B,j}$.

In the case of dishonest Alice and honest Bob, we define a cheating strategy for Alice analogously.

In the above description, Bob can apply a unitary transformation to the initial state of the protocol before Alice applies any of her unitary operations. In the next chapter, we show that this does not improve his optimal cheating probabilities. Since it is physically possible, we include it in the description and in the analysis in the next chapter.

If one party were to cheat, they have a desired outcome in mind, so ideally they want to increase the probability of that outcome larger than the honest outcome probability of 1/2. How much one can raise the probability is how we view the security of such protocols.

Definition 6.4 (Maximum Cheating Probability)**.** *Suppose $\mathcal{P}$ is a protocol-tuple. Let $|\psi_{final}(\mathcal{P}_B)\rangle$ be the final state of the protocol when Bob uses cheating strategy $\mathcal{P}_B$. We define the* maximum cheating probability for Bob forcing outcome $x \in \{0,1\}$*, denoted $P^*_{B,x}$, as*

$$P^*_{B,x} := \max_{\mathcal{P}_B} \Pr(\textit{Alice's outcome } = x) = \max_{\mathcal{P}_B} \|\Pi_{A,x} |\psi_{final}(\mathcal{P}_B)\rangle\|^2 .$$

We analogously define $P^*_{A,x}$, for $x \in \{0,1\}$, for Alice's maximum cheating probability.

Since a protocol is designed to protect an honest party from a dishonest party, a protocol is only as good as the maximum of the four cheating probabilities. This leads to the following definition.

Definition 6.5 (Bias)**.** *For a protocol-tuple $\mathcal{P}$, we define the* bias*, denoted ε, as*

$$\varepsilon := \max\{P^*_{A,0}, P^*_{A,1}, P^*_{B,0}, P^*_{B,1}\} - \frac{1}{2}.$$

This is our measure of security of such protocols. We later show how to solve for $P^*_{A,0}$, $P^*_{A,1}$, $P^*_{B,0}$, and $P^*_{B,1}$ using semidefinite programming and thus determining the bias. The central focus of studying coin-flipping protocols is to try to find protocols which have a low bias so that they are secure against dishonest parties.

Naturally, for any coin-flipping protocol, the bias is in the interval $\left[0, \frac{1}{2}\right]$. Ideally, we want a protocol with 0 bias, but this has been proven impossible by Mayers [23] and Lo and Chau [22].

6.2 An Example of a Coin-Flipping Protocol

We illustrate a simple coin-flipping protocol and show it satisfies Conditions 6.2. This protocol attempts to make a quantum state in each of Alice and Bob's private spaces in such a way that if either try to change the content of the last few messages, the other party can detect cheating.

Suppose we have the five-qubit state space $\mathcal{H} := \mathcal{A}_1 \otimes \mathcal{A}_2 \otimes \mathcal{M} \otimes \mathcal{B}_1 \otimes \mathcal{B}_2$ where $\mathcal{A} := \mathcal{A}_1 \otimes \mathcal{A}_2$ and $\mathcal{B} := \mathcal{B}_2 \otimes \mathcal{B}_2$. Define $U_{A,1}$ and $U_{B,1}$ such that

$$U_{A,1}\left(|\mathbf{0}\rangle_{\mathcal{A}}|0\rangle_{\mathcal{M}}\right) = \frac{1}{\sqrt{2}}\sum_{a\in\{0,1\}}|aa\rangle_{\mathcal{A}}|0\rangle_{\mathcal{M}} \quad \text{and} \quad U_{B,1}\left(|0\rangle_{\mathcal{M}}|\mathbf{0}\rangle_{\mathcal{B}}\right) = \frac{1}{\sqrt{2}}|0\rangle_{\mathcal{M}}\sum_{b\in\{0,1\}}|bb\rangle_{\mathcal{B}}.$$

Define $U_{A,2} := I_{\mathcal{A}_1} \otimes SWAP_4$, where $SWAP_4 \in \mathcal{L}(\mathbb{C}^4)$ is the unitary operation of swapping the places of two qubits, i.e., $SWAP_4\left(|x_1\rangle|x_2\rangle\right) = K_{2,2}\left(|x_1\rangle|x_2\rangle\right) = |x_2\rangle|x_1\rangle$, for all basis states $x_1, x_2 \in \{0,1\}$.

It is given explicitly as

$$SWAP_4 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Define $U_{B,2} := SWAP_4 \otimes I_{\mathcal{B}_2}$ and $U_{A,3} := I_{\mathcal{A}_1} \otimes SWAP_4$.

The first two unitary operations create specific states in $\mathcal{A}$ and $\mathcal{B}$. The next three are used to swap qubits between Alice and Bob.

For this protocol, we use the projectors

$$\Pi_{A,0} := \Pi_{B,0} := \sum_{c\in\{0,1\}}|c\rangle\langle c| \otimes |c\rangle\langle c| \quad \text{and} \quad \Pi_{A,1} := \Pi_{B,1} := \sum_{c\in\{0,1\}}|c\rangle\langle c| \otimes |\bar{c}\rangle\langle\bar{c}|,$$

where $\bar{c} := 1 - c$, for $c \in \{0,1\}$.

To fully illustrate this protocol, let us examine an honest run of the protocol by applying all the unitary transformations.

$|\psi_0\rangle := |\mathbf{0}\rangle_{\mathcal{A}}|0\rangle_{\mathcal{M}}|\mathbf{0}\rangle_{\mathcal{B}}$ is the starting state. Alice applies $U_{A,1}$ which creates the state

$$\begin{aligned} |\psi_1\rangle &= (U_{A,1} \otimes I_{\mathcal{B}})|\psi_0\rangle \\ &= \frac{1}{\sqrt{2}}\sum_{a\in\{0,1\}}|aa\rangle|0\rangle|00\rangle. \end{aligned}$$

Bob now applies $U_{B,1}$ to create the state

$$\begin{aligned} |\psi_2\rangle &= (I_{\mathcal{A}} \otimes U_{B,1}) |\psi_1\rangle \\ &= \frac{1}{2} \sum_{a,b \in \{0,1\}} |aa\rangle |0\rangle |bb\rangle . \end{aligned}$$

Continuing in this fashion the protocol generates the states

$$\begin{aligned} |\psi_3\rangle &= (U_{A,2} \otimes I_{\mathcal{B}}) |\psi_2\rangle \\ &= (I_{\mathcal{A}_1} \otimes SWAP_4 \otimes I_{\mathcal{B}_1} \otimes I_{\mathcal{B}_2}) \left(1/2 \sum_{a,b \in \{0,1\}} |aa\rangle |0\rangle |bb\rangle \right) \\ &= \frac{1}{2} \sum_{a,b \in \{0,1\}} |a0\rangle |a\rangle |bb\rangle , \end{aligned}$$

$$\begin{aligned} |\psi_4\rangle &= (I_{\mathcal{A}} \otimes U_{B,2}) |\psi_3\rangle \\ &= (I_{\mathcal{A}_1} \otimes I_{\mathcal{A}_2} \otimes SWAP_4 \otimes I_{\mathcal{B}_2}) \left(1/2 \sum_{a,b \in \{0,1\}} |a0\rangle |a\rangle |bb\rangle \right) \\ &= \frac{1}{2} \sum_{a,b \in \{0,1\}} |a0\rangle |b\rangle |ab\rangle , \text{ and} \end{aligned}$$

$$\begin{aligned} |\psi_5\rangle &= (U_{A,3} \otimes I_{\mathcal{B}}) |\psi_4\rangle \\ &= (I_{\mathcal{A}_1} \otimes SWAP_4 \otimes I_{\mathcal{B}_1} \otimes I_{B_2}) \left(1/2 \sum_{a,b \in \{0,1\}} |a0\rangle |b\rangle |ab\rangle \right) \\ &= \frac{1}{2} \sum_{a,b \in \{0,1\}} |ab\rangle |0\rangle |ab\rangle . \end{aligned}$$

After all the unitary operations have been applied, $|\psi_5\rangle$ is the final state of the protocol.

We still have to check to see if Conditions 6.2 hold. Applying $\Pi_{A,0}$ to $|\psi_5\rangle$ we get

$$\begin{aligned}(\Pi_{A,0}\otimes I_{\mathcal{M}}\otimes I_{\mathcal{B}})\,|\psi_5\rangle &= \left(\sum_{c\in\{0,1\}}|cc\rangle\langle cc|\otimes I_{\mathcal{M}}\otimes I_{\mathcal{B}}\right)\left(\frac{1}{2}\sum_{a,b\in\{0,1\}}|ab\rangle\,|0\rangle\,|ab\rangle\right)\\ &= \frac{1}{2}\sum_{a,b,c\in\{0,1\}}|c\rangle\,\langle c|a\rangle\,|c\rangle\,\langle c|b\rangle\,|0\rangle\,|ab\rangle\\ &= \frac{1}{2}\sum_{a\in\{0,1\}}|aa\rangle\,|0\rangle\,|aa\rangle\,.\end{aligned}$$

A similar check shows us that

$$(I_{\mathcal{A}}\otimes I_{\mathcal{M}}\otimes\Pi_{B,0})\,|\psi_5\rangle=\frac{1}{2}\sum_{a\in\{0,1\}}|aa\rangle\,|0\rangle\,|aa\rangle\,.$$

Therefore, $(\Pi_{A,0}\otimes I_{\mathcal{M}}\otimes I_{\mathcal{B}})\,|\psi_5\rangle=(I_{\mathcal{A}}\otimes I_{\mathcal{M}}\otimes\Pi_{B,0})\,|\psi_5\rangle$. We can similarly show that

$$(\Pi_{A,1}\otimes I_{\mathcal{M}}\otimes I_{\mathcal{B}})\,|\psi_5\rangle=\frac{1}{2}\sum_{a\in\{0,1\}}|a\bar{a}\rangle\,|0\rangle\,|a\bar{a}\rangle=(I_{\mathcal{A}}\otimes I_{\mathcal{M}}\otimes\Pi_{B,1})\,|\psi_5\rangle\,,$$

so the first condition is satisfied. To check the second condition, we have

$$\begin{aligned}\|(\Pi_{A,0}\otimes I_{\mathcal{M}}\otimes I_{\mathcal{B}})\,|\psi_5\rangle\|^2 &= \left\|\frac{1}{2}\sum_{a\in\{0,1\}}|aa\rangle\,|0\rangle\,|aa\rangle\right\|^2\\ &= \left(\frac{1}{2}\sum_{a\in\{0,1\}}\langle aa|\,\langle 0|\,\langle aa|\right)\left(\frac{1}{2}\sum_{b\in\{0,1\}}|bb\rangle\,|0\rangle\,|bb\rangle\right)\\ &= \frac{1}{4}\sum_{a\in\{0,1\}}\langle a|b\rangle\langle a|b\rangle\langle 0|0\rangle\langle a|b\rangle\langle a|b\rangle\\ &= \frac{1}{4}\sum_{a\in\{0,1\}}1\\ &= \frac{1}{2}.\end{aligned}$$

Similarly, we can show that

$$\|(\Pi_{A,1}\otimes I_{\mathcal{M}}\otimes I_{\mathcal{B}})\,|\psi_5\rangle\|^2=\frac{1}{2},$$

so the second condition is satisfied.

This is an example of a coin-flipping protocol based on *bit-commitment*, discussed in detail in Chapter 8.

Chapter 7

Security Analysis of Quantum Coin-Flipping Protocols

7.1 Semidefinite Programming Formulation

In this chapter we show how to analyze the security of coin-flipping protocols by formulating cheating strategies as semidefinite programs. Suppose that Alice is honest and Bob is dishonest. A cheating strategy for Bob is to create a new private space $\widehat{\mathcal{B}}$ and a new set of unitary operations $\widehat{U}_{B,j}$, $j \in \{0, 1, \ldots, N\}$, to apply instead of the unitary operations in the protocol. In our cheating strategy formulation, the variables are the density matrices ρ_j, $j \in \{0, 1, \ldots, N\}$, where ρ_j is the reduced density matrix Alice has after Bob applies the unitary $\widehat{U}_{B,j}$. Bob's unitary operations preserve the reduced density matrix that is in Alice's private space $\mathcal{A}$. Alice's reduced density matrices therefore satisfy

$$\mathrm{Tr}_{\mathcal{M}}(\rho_j) = \mathrm{Tr}_{\mathcal{M}}\left[U_{A,j}\rho_{j-1}U_{A,j}^*\right], \ \forall\, j \in \{1, \ldots, N\}.$$

Bob cannot change the part of the starting state in Alice's space. This is given as the constraint

$$\mathrm{Tr}_{\mathcal{M}}(\rho_0) = |0\rangle\langle 0|_{\mathcal{A}}.$$

Suppose Bob is interested in forcing outcome $x \in \{0, 1\}$. Then the probability of Alice accepting outcome x is given as

$$\langle \Pi_{A,x} \otimes I_{\mathcal{M}}, \rho_N \rangle .$$

This is the objective function which Bob attempts to maximize. We collect the above in a semidefinite program given below.

Theorem 7.1 ([19]). *For a protocol-tuple $\mathcal{P}$, Bob's optimal cheating probability for forcing outcome $x \in \{0,1\}$ is given as the optimal value of the following SDP.*

$$\begin{array}{lrcl} (B) \quad P^*_{B,x} = \sup & \langle \Pi_{A,x} \otimes I_{\mathcal{M}}, \rho_N \rangle & & \\ & \mathrm{Tr}_{\mathcal{M}}(\rho_0) & = & |0\rangle\langle 0|_{\mathcal{A}}, \\ \text{subject to} & \mathrm{Tr}_{\mathcal{M}}(\rho_j) & = & \mathrm{Tr}_{\mathcal{M}}(U_{A,j}\rho_{j-1}U^*_{A,j}),\ \forall j \in \{1,\ldots,N\}, \\ & \rho_j & \succeq & 0,\ \forall j \in \{0,1,\ldots,N\}. \end{array}$$

This semidefinite program is referred to as *Bob's cheating SDP*. Note that the variables are indeed density matrices as they are positive semidefinite and all are forced to have trace equal to $\mathrm{Tr}\,|0\rangle\langle 0|_{\mathcal{A}} = 1$.

What is also interesting about the above formulation is that we can determine an optimal cheating strategy from an optimal solution of the SDP.

Corollary 7.2 ([19]). *For each $x \in \{0,1\}$ and any feasible solution $(\rho_0, \rho_1, \ldots, \rho_N)$ of the above SDP, Bob has a cheating strategy that forces outcome x with a probability of $\langle \Pi_{A,x} \otimes I_{\mathcal{M}}, \rho_N \rangle$.*

Proof. Suppose for $x \in \{0,1\}$ that $(\rho_0, \rho_1, \ldots, \rho_N)$ is a feasible solution for Bob's cheating SDP. Let $|\psi_j\rangle \in \mathcal{A} \otimes \mathcal{M} \otimes \widehat{\mathcal{B}}$ be a purification of ρ_j, for $j \in \{0,1,\ldots,N\}$. Note that $\mathrm{Tr}_{\widehat{\mathcal{B}}}|\psi_j\rangle\langle\psi_j| = \rho_j$, for $j \in \{0,1,\ldots,N\}$. For each $j \in \{1,\ldots,N\}$, we can write

$$\begin{aligned} \mathrm{Tr}_{\mathcal{M}}\left(\mathrm{Tr}_{\widehat{\mathcal{B}}}|\psi_j\rangle\langle\psi_j|\right) &= \mathrm{Tr}_{\mathcal{M}}\left(\rho_j\right) \\ &= \mathrm{Tr}_{\mathcal{M}}\left(U_{A,j}\rho_{j-1}U^*_{A,j}\right) \\ &= \mathrm{Tr}_{\mathcal{M}}\left(U_{A,j}\left(\mathrm{Tr}_{\widehat{\mathcal{B}}}|\psi_{j-1}\rangle\langle\psi_{j-1}|\right)U^*_{A,j}\right) \\ &= \mathrm{Tr}_{\mathcal{M}}\left(\mathrm{Tr}_{\widehat{\mathcal{B}}}\left[(U_{A,j}\otimes I_{\widehat{\mathcal{B}}})|\psi_{j-1}\rangle\langle\psi_{j-1}|(U^*_{A,j}\otimes I_{\widehat{\mathcal{B}}})\right]\right). \end{aligned}$$

Therefore, by Lemma 3.4, there exists a unitary $V_j \in \mathcal{L}(\mathcal{M}\otimes\widehat{\mathcal{B}})$ such that

$$(I_{\mathcal{A}}\otimes V_j)(U_{A,j}\otimes I_{\widehat{\mathcal{B}}})\,|\psi_{j-1}\rangle = |\psi_j\rangle\,.$$

Let $\widehat{U}_{B,j}$ be such a V_j for each $j \in \{1,\ldots,N\}$. Let $\widehat{U}_{B,0}$ be a unitary satisfying

$$\widehat{U}_{B,0}\,|0\rangle_{\mathcal{A}\otimes\mathcal{M}\otimes\widehat{\mathcal{B}}} = |\psi_0\rangle\,.$$

This defines a cheating strategy for Bob. Using this strategy, Alice has the density matrix ρ_j after applying $\widehat{U}_{B,j}$, for $j \in \{0, 1, \ldots, N\}$. Therefore, Alice accepts outcome x with probability $\langle \Pi_{A,x} \otimes I_{\mathcal{M}}, \rho_N \rangle$. □

From the above proof, we see that there is no need for the unitary $\widehat{U}_{B,0}$. That is, we can choose it arbitrarily.

By solving his cheating SDP for a desired outcome, Bob finds his maximum cheating probability and he has an optimal cheating strategy achieving this value.

Theorem 7.3 ([19])**.** *The dual of Bob's cheating SDP is given as the following SDP.*

$$\begin{array}{lll} (D_B) & \inf & \langle Z_0, |0\rangle\langle 0| \rangle \\ & & Z_j \otimes I_{\mathcal{M}} \succeq U^*_{A,j+1}(Z_{j+1} \otimes I_{\mathcal{M}})U_{A,j+1}, \ \forall j \in \{0, 1, \ldots, N-1\}, \\ & \text{subject to} & Z_N = \Pi_{A,x}. \end{array}$$

Proof. For each variable in (B), we look at the linear operators acting on it. Consider ρ_j for $j \in \{0, 1, \ldots, N-1\}$, it is acted on by $\mathrm{Tr}_{\mathcal{M}}(\cdot)$ in one constraint, corresponding to a dual variable Z_i, and by $\mathrm{Tr}_{\mathcal{M}}\left[(U_{A,j+1}(\cdot)U^*_{A,j+1})\right]$ in another constraint, corresponding to a dual variable Z_{i+1}. Recall the adjoint of the linear operator $\mathrm{Tr}_{\mathcal{M}}$ is $\mathrm{Tr}^*_{\mathcal{M}}(Z) = Z \otimes I_{\mathcal{M}}$ and the adjoint of $T(\rho) := U\rho U^*$ is $T^*(Z) = U^*ZU$. Since these variables do not appear in the objective function, and $\rho_i \succeq 0$, for $i \in \{0, 1, \ldots, N-1\}$, we have the dual constraints

$$Z_j \otimes I_{\mathcal{M}} \succeq U^*_{A,j+1}(Z_{j+1} \otimes I_{\mathcal{M}})U_{A,j+1}, \ \forall j \in \{0, 1, \ldots, N-1\}.$$

We can write the constraints in this way since $A \succeq B$ if and only if $\exists S \succeq 0$ such that $A = B + S$. Notice ρ_N appears in the last constraint acted on by $\mathrm{Tr}_{\mathcal{M}}$, corresponding to a dual variable Z_N, and it appears in the objective function $\langle \Pi_{A,x} \otimes I_{\mathcal{M}}, \rho_N \rangle$. This gives a dual constraint $\mathrm{Tr}^*_{\mathcal{M}}(Z_N) = Z_N \otimes I_{\mathcal{M}} \succeq \Pi_{A,x} \otimes I_{\mathcal{M}}$. This is equivalent to $Z_N \succeq \Pi_{A,x}$ from which we may assume any optimal solution satisfies with equality. Since the only non-homogeneous constraint in the primal is $\mathrm{Tr}_{\mathcal{M}}(\rho_0) = |0\rangle\langle 0|$, corresponding to the dual variable Z_0, we have the objective function of the dual $\langle Z_0, |0\rangle\langle 0| \rangle$. □

Notice that the dual variables are positive semidefinite because $\Pi_{A,x} \succeq 0$, for each $x \in \{0, 1\}$, and $U^*_{A,j+1}(\cdot)U_{A,j+1} \in \mathrm{Aut}\,(\Sigma_+^{\mathcal{A}\otimes\mathcal{M}})$.

The next lemma shows that there is zero duality gap.

Lemma 7.4 ([19]). *The primal-dual pair (B, D_B) has zero duality gap and primal attainment.*

Proof. We see that we get a feasible solution to the primal if Bob plays honestly. Eliminating Z_N and setting $Z_j = (N-j+1)I_{\mathcal{A}}$, for each $j \in \{0,1,\ldots,N-1\}$, gives a strictly feasible solution to the dual. Therefore, there is zero duality gap and primal attainment from strong duality. □

Since the primal SDP attains its optimal value, we know for any protocol there is indeed an achievable optimal cheating strategy. The zero duality gap implies that we can solve for the maximum cheating probability by solving the dual SDP. Weak duality implies that the objective value of any feasible solution of the dual gives an upper bound on the optimal cheating probability.

We now consider the case of cheating Alice. This case is similar to cheating Bob except that we let her apply a unitary $\widehat{U}_{A,N+1}$ after Bob's last unitary $U_{B,N}$. Since Alice does not perform any measurements, applying such a unitary does not give her an advantage. However, we include it in our analysis since it is physically possible.

Her cheating SDP is given as

$$\begin{array}{lll} (A) \quad P^*_{A,x} = \sup & \langle I_{\mathcal{M}} \otimes \Pi_{B,x}, \sigma_N \rangle & \\ & \mathrm{Tr}_{\mathcal{M}}(\sigma_0) &= |0\rangle\langle 0|_{\mathcal{B}}, \\ \text{subject to} & \mathrm{Tr}_{\mathcal{M}}(\sigma_j) &= \mathrm{Tr}_{\mathcal{M}}(U_{B,j}\sigma_{j-1}U^*_{B,j}),\ \forall j \in \{1,\ldots,N\}, \\ & \sigma_j &\succeq 0,\ \forall j \in \{0,1,\ldots,N\}. \end{array}$$

The dual of Alice's cheating SDP is given below.

$$\begin{array}{lll} (D_A) \quad \inf & \langle Y_0, |0\rangle\langle 0| \rangle & \\ & I_{\mathcal{M}} \otimes Y_j &\succeq U^*_{B,j+1}(I_{\mathcal{M}} \otimes Y_{j+1})U_{B,j+1},\ \forall j \in \{0,1,\ldots,N-1\}, \\ \text{subject to} & Y_N &= \Pi_{B,x}. \end{array}$$

All the results about Bob's cheating SDP also hold for Alice's cheating SDP. Namely, Alice's cheating SDP and dual have zero duality gap and Alice can extract an optimal cheating strategy from an optimal solution of her cheating SDP.

7.2 A Lower Bound on the Bias of any Coin-Flipping Protocol

Recall that a coin-flipping protocol with 0 bias is impossible [22, 23]. We now present other lower bounds. The next result is by Ambainis.

Theorem 7.5 ([2])**.** *Any coin-flipping protocol achieving bias ε must use $\Omega(\log\log\varepsilon)$ rounds of communication.*

We now examine the best known lower bound for a general coin-flipping protocol by Kitaev [19].

Theorem 7.6 ([19])**.** *Achieving a bias of $\varepsilon < 1/\sqrt{2} - 1/2$ is impossible in any strong coin-flipping protocol. In other words, at least one party in any strong coin-flipping protocol can cheat with probability at least $1/\sqrt{2} - 1/2 \approx 0.2071$.*

Proof. Assume we are given an arbitrary protocol-tuple and we form Alice and Bob's cheating SDP's where each separately tries to force outcome 0 when the other is honest. Zero duality gaps imply for any $\epsilon > 0$ there exist feasible solutions $(Z_0, Z_1, \ldots, Z_N)$ for (D_B) and $(Y_0, \ldots, Y_N)$ for (D_A) such that $P^*_{B,0} + \epsilon > \langle Z_0, |0\rangle\langle 0| \rangle$ and $P^*_{A,0} + \epsilon > \langle Y_0, |0\rangle\langle 0| \rangle$. Since $\langle I_{\mathcal{M}}, |0\rangle\langle 0| \rangle = 1$, we can write the following.

$$\begin{aligned}(P^*_{B,0} + \epsilon)(P^*_{A,0} + \epsilon) &> \langle Z_0, |0\rangle\langle 0| \rangle \langle I_{\mathcal{M}}, |0\rangle\langle 0| \rangle \langle Y_0, |0\rangle\langle 0| \rangle \\ &= \langle Z_0 \otimes I_{\mathcal{M}} \otimes Y_0, |0\rangle\langle 0| \rangle .\end{aligned}$$

Define $|\psi_j\rangle := (I_{\mathcal{A}} \otimes U_{B,j})(U_{A,j} \otimes I_{\mathcal{B}}) \cdots (I_{\mathcal{A}} \otimes U_{B,1})(U_{A,1} \otimes I_{\mathcal{B}}) |0\rangle$, for $j \in \{0, \ldots, N\}$, to be the state after Bob applies his j'th unitary in an honest run of the protocol. Also, define $\Psi_j = \langle Z_j \otimes I_{\mathcal{M}} \otimes Y_j, |\psi_j\rangle\langle\psi_j| \rangle$ for $j \in \{0, \ldots, N\}$. We next show that $\Psi_j \geq \Psi_{j+1}$ for all

$j \in \{0, 1, \ldots, N-1\}$.

$$\begin{aligned}
\Psi_j &= \langle Z_j \otimes I_{\mathcal{M}} \otimes Y_j, |\psi_j\rangle\langle\psi_j| \rangle \\
&\geq \left\langle U^*_{A,j+1}(Z_{j+1} \otimes I_{\mathcal{M}})U^*_{A,j+1} \otimes Y_j, |\psi_j\rangle\langle\psi_j| \right\rangle \\
&= \left\langle Z_{j+1} \otimes I_{\mathcal{M}} \otimes Y_j, (U_{A,j+1} \otimes I_{\mathcal{B}})|\psi_j\rangle\langle\psi_j|(U^*_{A,j+1} \otimes I_{\mathcal{B}}) \right\rangle \\
&\geq \left\langle Z_{j+1} \otimes U^*_{B,j+1}(I_{\mathcal{M}} \otimes Y_{j+1})U_{B,j+1}, (U_{A,j+1} \otimes I_{\mathcal{B}})|\psi_j\rangle\langle\psi_j|(U^*_{A,j+1} \otimes I_{\mathcal{B}}) \right\rangle \\
&= \left\langle Z_{j+1} \otimes I_{\mathcal{M}} \otimes Y_{j+1}, (I_{\mathcal{A}} \otimes U_{B,j+1})(U_{A,j+1} \otimes I_{\mathcal{B}})|\psi_j\rangle\langle\psi_j|(U^*_{A,j+1} \otimes I_{\mathcal{B}})(I_{\mathcal{A}} \otimes U^*_{B,j+1}) \right\rangle \\
&= \langle Z_{j+1} \otimes I_{\mathcal{M}} \otimes Y_{j+1}, |\psi_{j+1}\rangle\langle\psi_{j+1}| \rangle \\
&= \Psi_{j+1}.
\end{aligned}$$

Also notice that

$$\begin{aligned}
\Psi_N &= \langle Z_N \otimes I_{\mathcal{M}} \otimes Y_N, |\psi_N\rangle\langle\psi_N| \rangle \\
&= \langle \Pi_{A,0} \otimes I_{\mathcal{M}} \otimes \Pi_{B,0}, |\psi_N\rangle\langle\psi_N| \rangle \\
&= \langle (I_{\mathcal{A}} \otimes I_{\mathcal{M}} \otimes \Pi_{B,0})(\Pi_{A,0} \otimes I_{\mathcal{M}} \otimes I_{\mathcal{B}}), |\psi_N\rangle\langle\psi_N| \rangle \\
&= \langle\psi_N| \left(I_{\mathcal{A}} \otimes I_{\mathcal{M}} \otimes \Pi_{B,0}\right)(\Pi_{A,0} \otimes I_{\mathcal{M}} \otimes I_{\mathcal{B}}) |\psi_N\rangle \\
&= \langle\psi_N| \left(\Pi_{A,0} \otimes I_{\mathcal{M}} \otimes I_{\mathcal{B}}\right)(\Pi_{A,0} \otimes I_{\mathcal{M}} \otimes I_{\mathcal{B}}) |\psi_N\rangle \\
&= \left\| \left(\Pi_{A,0} \otimes I_{\mathcal{M}} \otimes I_{\mathcal{B}}\right) |\psi_N\rangle \right\|^2 \\
&= 1/2,
\end{aligned}$$

where the last two equalities follow from Conditions 6.2.

Thus, we have

$$(P^*_{B,0} + \epsilon)(P^*_{A,0} + \epsilon) > \Psi_0 \geq \Psi_1 \geq \cdots \geq \Psi_N = 1/2.$$

Taking the limit $\epsilon \to 0$ implies $P^*_{B,0}P^*_{A,0} \geq 1/2$. Therefore, $\max\{P^*_{B,0}, P^*_{A,0}\} \geq 1/\sqrt{2}$. □

7.3 Solving for the Bias with One SDP

We can find the maximum of the four cheating probabilities by solving the following single SDP.

$$
\begin{aligned}
(P_{\max}) \quad \inf \quad & t \\
& \langle Z_0, |0\rangle\langle 0| \rangle \;\leq\; t, \\
& \langle Z_0', |0\rangle\langle 0| \rangle \;\leq\; t, \\
& \langle Y_0, |0\rangle\langle 0| \rangle \;\leq\; t, \\
& \langle Y_0', |0\rangle\langle 0| \rangle \;\leq\; t, \\
& \\
& Z_j \otimes I_{\mathcal{M}} \;\succeq\; U_{A,j+1}^*(Z_{j+1} \otimes I_{\mathcal{M}})U_{A,j+1},\ \forall j \in \{0,1,\ldots,N-1\}, \\
& Z_N \;=\; \Pi_{A,0}, \\
\text{subject to} \quad & \\
& Z_j' \otimes I_{\mathcal{M}} \;\succeq\; U_{A,j+1}^*(Z_{j+1}' \otimes I_{\mathcal{M}})U_{A,j+1},\ \forall j \in \{0,1,\ldots,N-1\}, \\
& Z_N' \;=\; \Pi_{A,1}, \\
& \\
& I_{\mathcal{M}} \otimes Y_j \;\succeq\; U_{B,j+1}^*(I_{\mathcal{M}} \otimes Y_{j+1})U_{B,j+1},\ \forall j \in \{0,1,\ldots,N-1\}, \\
& Y_N \;=\; \Pi_{B,0}, \\
& \\
& I_{\mathcal{M}} \otimes Y_j' \;\succeq\; U_{B,j+1}^*(I_{\mathcal{M}} \otimes Y_{j+1}')U_{B,j+1},\ \forall j \in \{0,1,\ldots,N-1\}, \\
& Y_N' \;=\; \Pi_{B,1}.
\end{aligned}
$$

This clearly solves for the maximum of the four cheating probabilities, which is exactly $\varepsilon + 1/2$. In the proof of Kitaev's lower bound we show that at least one party can cheat towards 0 with probability $1/\sqrt{2}$. The same result holds for the parties cheating towards 1. We see that we can recreate Kitaev's proof using an optimal solution of the above SDP.

Taking the dual of the above SDP yields the following SDP.

$$
\begin{aligned}
(D_{\max}) \quad \sup \quad & \langle \Pi_{A,0}, \rho_N \rangle + \langle \Pi_{A,1}, \rho'_N \rangle + \langle \Pi_{B,0}, \sigma_{N-1} \rangle + \left\langle \Pi_{B,1}, \sigma'_{N-1} \right\rangle \\
& \text{Tr}_{\mathcal{M}}(\rho_0) = \lambda_1 |0\rangle\langle 0|_{\mathcal{A}}, \\
& \text{Tr}_{\mathcal{M}}(\rho_j) = \text{Tr}_{\mathcal{M}}(U_{A,j} \rho_{j-1} U_{A,j}^*), \ \forall j \in \{1, \ldots, N\}, \\
& \rho_j \succeq 0, \ \forall j \in \{0, 1, \ldots, N-1\}, \\
& \\
& \text{Tr}_{\mathcal{M}}(\rho'_0) = \lambda_2 |0\rangle\langle 0|_{\mathcal{A}}, \\
& \text{Tr}_{\mathcal{M}}(\rho'_j) = \text{Tr}_{\mathcal{M}}(U_{A,j} \rho'_{j-1} U_{A,j}^*), \ \forall j \in \{1, \ldots, N\}, \\
& \rho'_j \succeq 0, \ \forall j \in \{0, 1, \ldots, N-1\}, \\
& \\
\text{subject to} \quad & \text{Tr}_{\mathcal{M}}(\sigma_0) = \lambda_3 |0\rangle\langle 0|_{\mathcal{B}}, \\
& \text{Tr}_{\mathcal{M}}(\sigma_j) = \text{Tr}_{\mathcal{M}}(U_{B,j} \sigma_{j-1} U_{B,j}^*), \ \forall j \in \{1, \ldots, N\}, \\
& \sigma_j \succeq 0, \ \forall j \in \{0, 1, \ldots, N-1\}, \\
& \\
& \text{Tr}_{\mathcal{M}}(\sigma'_0) = \lambda_4 |0\rangle\langle 0|_{\mathcal{B}}, \\
& \text{Tr}_{\mathcal{M}}(\sigma'_j) = \text{Tr}_{\mathcal{M}}(U_{B,j} \sigma'_{j-1} U_{B,j}^*), \ \forall j \in \{1, \ldots, N\}, \\
& \sigma'_j \succeq 0, \ \forall j \in \{0, 1, \ldots, N-1\}, \\
& \\
& \lambda_1, \lambda_2, \lambda_3, \lambda_4 \geq 0, \\
& \lambda_1 + \lambda_2 + \lambda_3 + \lambda_4 = 1.
\end{aligned}
$$

Note that some of the variables are not positive semidefinite, namely ρ_N, ρ'_N, σ_N, and σ'_N. Since $\Pi_{A,0}, \Pi_{A,1}, \Pi_{B,0}, \Pi_{B,1} \succeq 0$, we can assume that $\rho_N, \rho'_N, \sigma_N, \sigma'_N \succeq 0$ as well, since we are maximizing the inner products above.

Although this does not help computationally, it does incorporate all of the optimal cheating probabilities and cheating strategies of the entire protocol.

We now examine a specific class of coin-flipping protocols based on *quantum bit-commitment* which includes protocols that have the best known bias of 1/4.

Chapter 8

Coin-Flipping Based on Quantum Bit-Commitment

8.1 Quantum Bit-Commitment

Quantum bit-commitment (BC) is a two-party protocol where one party, Alice say, chooses a bit $a \in \{0, 1\}$ at the beginning of the protocol which is to be unveiled to the other party, say Bob, at the end of the protocol. BC protocols are designed to limit the amount of information Bob can obtain about the bit a and, simultaneously, to protect from Alice changing the value of a before unveiling. We say that Alice is *committed to the bit* a and refer to a as her *commitment*. Mayers [23] and Lo and Chau [22] show that quantum bit-commitment is impossible. This means that there is a positive probability that Alice can change her commitment or Bob can extract information about her commitment.

To illustrate BC protocols, consider the following example.

Protocol 8.1.

1. *Commitment Stage: Alice chooses* $a \in \{0, 1\}$ *and commits to the two-qutrit state*

$$|\psi_a\rangle := \frac{1}{\sqrt{2}} \left(|aa\rangle + |22\rangle\right) \in \mathbb{C}^3 \otimes \mathbb{C}^3$$

by sending the second qutrit to Bob. We call $|\psi_0\rangle$ *and* $|\psi_1\rangle$ *Alice's commitment states.*

2. *Unveiling Stage: Once Alice is ready to unveil a to Bob, she sends a and the first qutrit of $|\psi_a\rangle$. Bob can now measure using the projective measurement*

$$\{\Pi_{accept} := |\psi_a\rangle\langle\psi_a|, \Pi_{abort} := I_{\mathbb{C}^3\otimes\mathbb{C}^3} - |\psi_a\rangle\langle\psi_a|\}$$

to see if Alice tried to change her commitment.

If Alice tries to change her commitment, she risks Bob detecting her cheating attempt. On the other hand, Bob may not be able to extract much information about the bit a from his half of $|\psi_a\rangle$. The amount Alice and Bob can cheat is known in this case and is later examined in a simple coin-flipping protocol.

8.2 Coin-Flipping Based on Quantum Bit-Commitment

We now discuss a type of coin-flipping protocol introduced by Nayak and Shor [29]. The idea is that Alice commits to a random bit $a \in \{0,1\}$ and Bob commits to a random bit $b \in \{0,1\}$. Then after both unveil their bits to each other, they output $a \oplus b$, which is the outcome of the protocol.

An *m-round BC coin-flipping protocol*,[1] referred to as a BCCF protocol, is of the following form.

Protocol 8.2. *Suppose $m = 2n-1$, i.e., m is odd and $n \geq 2$. Suppose Alice starts with the Hilbert space $\mathcal{A} := \mathcal{A}_1 \otimes \cdots \otimes \mathcal{A}_n$ and Bob starts with the Hilbert space $\mathcal{B} := \mathcal{B}_1 \otimes \cdots \otimes \mathcal{B}_{n-1}$.*

- *Preparation: Assume we have states $|\psi_0\rangle, |\psi_1\rangle \in \mathcal{A}$ satisfying $|\psi_0\rangle \perp |\psi_1\rangle$, and states $|\phi_0\rangle, |\phi_1\rangle \in \mathcal{B}$ satisfying $|\phi_0\rangle \perp |\phi_1\rangle$. These states are called the* start states *or* starting states*. Alice creates the state*

$$|A\rangle := \frac{1}{\sqrt{2}} \sum_{a\in\{0,1\}} |a\rangle |\psi_a\rangle \in \mathcal{A}_0 \otimes \mathcal{A} = \mathbb{C}^2 \otimes \mathcal{A}$$

and Bob creates the state

$$|B\rangle := \frac{1}{\sqrt{2}} \sum_{b\in\{0,1\}} |b\rangle |\phi_b\rangle \in \mathcal{B}_0 \otimes \mathcal{B} = \mathbb{C}^2 \otimes \mathcal{B}.$$

[1] We assume $m \geq 3$ since 1 and 2-round coin-flipping protocols are not of much interest.

- *Commitment and Unveiling Stages: Alice first sends $\mathcal{A}_1$ to Bob. Then Bob sends $\mathcal{B}_1$ to Alice. Then Alice sends $\mathcal{A}_2$ to Bob and so on until Alice has all of $\mathcal{B}$ and Bob has all of $\mathcal{A}$.*

- *Measurement Stage: Alice measures the state she has in $\mathcal{A}_0 \otimes \mathcal{B}$ using the projective measurement*

  $$\{\Pi_{A,0} := \sum_{b\in\{0,1\}} |b\rangle\langle b| \otimes |\phi_b\rangle\langle\phi_b|,\ \Pi_{A,1} := \sum_{b\in\{0,1\}} |\bar{b}\rangle\langle\bar{b}| \otimes |\phi_b\rangle\langle\phi_b|,\ I_{\mathcal{A}_0\otimes\mathcal{B}} - \Pi_{A,0} - \Pi_{A,1}\}$$

  to get c_A, which she outputs. Bob measures the state he has in $\mathcal{B}_0 \otimes \mathcal{A}$ using the projective measurement

  $$\{\Pi_{B,0} := \sum_{a\in\{0,1\}} |a\rangle\langle a| \otimes |\psi_a\rangle\langle\psi_a|,\ \Pi_{B,1} := \sum_{a\in\{0,1\}} |\bar{a}\rangle\langle\bar{a}| \otimes |\psi_a\rangle\langle\psi_a|,\ I_{\mathcal{B}_0\otimes\mathcal{A}} - \Pi_{B,0} - \Pi_{B,1}\}$$

  to get c_B, which he outputs.

In the above protocol, if $c_A \neq c_B$, then an honest party detects cheating. We note here that we do not lose any generality by assuming the protocol has an odd number of messages. Suppose we have a 4-round BCCF protocol. This is equivalent to a 5-round BCCF protocol where we append a trivial message at the end. In other words, Alice sends some message in the last round which does not depend on her commitment, for instance, the last message could be $|0\rangle$. We consider BCCF protocols with an even number of rounds in later chapters, but we assume an odd number of rounds here for the analysis.

This type of protocol is a special case of the general protocol discussed in Chapter 7, where the first two unitary operations set up the starting states $|A\rangle$ and $|B\rangle$ and the rest of the unitary operations are commutation or "swap" matrices. (Recall from Section 3.2 that commutation matrices swap tensor factors.) Including a message space in BCCF protocols is redundant since instead of using commutation matrices, we imagine the spaces $\mathcal{A}_i$ and $\mathcal{B}_i$ changing hands.

We next prove that this type of protocol satisfies Conditions 6.2 for any choice of starting states $|\psi_0\rangle, |\psi_1\rangle$, satisfying $|\psi_0\rangle \perp |\psi_1\rangle$, and $|\phi_0\rangle, |\phi_1\rangle$, satisfying $|\phi_0\rangle \perp |\phi_1\rangle$.

Theorem 8.3. *Any BCCF protocol satisfies Conditions 6.2.*

Proof. Let $|\psi_{\text{final}}\rangle = \frac{1}{2} \sum_{a,b\in\{0,1\}} |a\rangle\,|\phi_b\rangle\,|b\rangle\,|\psi_a\rangle$ be the final state at the end of an honest run of the protocol. Then we have

$$\begin{aligned}
(\Pi_{A,0} \otimes I_{\mathcal{A}_0\otimes\mathcal{B}})\,|\psi_{\text{final}}\rangle &= \frac{1}{2} \sum_{a,b,b'\in\{0,1\}} (|b'\rangle\langle b'| \otimes |\phi_{b'}\rangle\langle\phi_{b'}| \otimes I_{\mathcal{A}_0\otimes\mathcal{B}})\,|a\rangle\,|\phi_b\rangle\,|b\rangle\,|\psi_a\rangle \\
&= \frac{1}{2} \sum_{a,b,b'\in\{0,1\}} |b'\rangle\,\langle b'|a\rangle\,|\phi_{b'}\rangle\,\langle\phi_{b'}|\phi_b\rangle\,|b\rangle\,|\psi_a\rangle \\
&= \frac{1}{2} \sum_{a\in\{0,1\}} |a\rangle\,|\phi_a\rangle\,|a\rangle\,|\psi_a\rangle\,.
\end{aligned}$$

Similarly, we have

$$(I_{\mathcal{A}_0\otimes\mathcal{B}} \otimes \Pi_{B,0})\,|\psi_{\text{final}}\rangle = \frac{1}{2} \textstyle\sum_{a\in\{0,1\}} |a\rangle\,|\phi_a\rangle\,|a\rangle\,|\psi_a\rangle\,,$$

$$(\Pi_{A,1} \otimes I_{\mathcal{A}_0\otimes\mathcal{B}})\,|\psi_{\text{final}}\rangle = \frac{1}{2} \textstyle\sum_{a\in\{0,1\}} |\bar{a}\rangle\,|\phi_a\rangle\,|a\rangle\,|\psi_{\bar{a}}\rangle\,,\ \text{and}$$

$$(I_{\mathcal{A}_0\otimes\mathcal{B}} \otimes \Pi_{B,1})\,|\psi_{\text{final}}\rangle = \frac{1}{2} \textstyle\sum_{a\in\{0,1\}} |\bar{a}\rangle\,|\phi_a\rangle\,|a\rangle\,|\psi_{\bar{a}}\rangle\,.$$

Therefore, Conditions 6.2 are satisfied. □

We later discuss how to find optimal cheating probabilities for this type of protocol using semidefinite programming. First, we examine known results about the special case of 3-round BCCF protocols.

8.3 Three-Round BCCF Protocols

For a 3-round protocol, we examine the cheating probabilities for Alice and Bob. Let $|\psi_a\rangle \in \mathcal{A}_1 \otimes \mathcal{A}_2$ be Alice's start states satisfying $|\psi_0\rangle \perp |\psi_1\rangle$ and let $|\phi_b\rangle \in \mathcal{B}_1$ be Bob's start states satisfying $|\phi_0\rangle \perp |\phi_1\rangle$. Recall that Alice sends the first and last messages and Bob sends the second message.

Bob's cheating strategy is straightforward. He only sends one message to Alice and this is his only opportunity to cheat. If he wants to force a desired outcome, what he

effectively does is determine Alice's commitment. If he can determine Alice's commitment with probability p, then he can force either outcome with probability p. This is because the outcome of the protocol is the XOR of Alice's commitment and Bob's commitment, so Bob can choose his commitment accordingly. The only information he can obtain about Alice's commitment is what he can extract from her first message. His optimal measurement can determine the bit to which Alice is committing with probability

$$\frac{1}{2} + \frac{1}{4} \left\| \rho_0 - \rho_1 \right\|_t ,$$

where $\rho_0 := \mathrm{Tr}_{\mathcal{A}_2} |\psi_0\rangle\langle\psi_0|$ and $\rho_1 := \mathrm{Tr}_{\mathcal{A}_2} |\psi_1\rangle\langle\psi_1|$, see [2]. This is the probability of correctly guessing i given ρ_i, $i \in \{0, 1\}$, with equal probability.

Alice's optimal cheating strategy is different. She is required to send a message in the first round, at which point she has no information about Bob's commitment. In the second round, Bob unveils his commitment to Alice, so at the time of Alice's last message she has complete information about Bob's commitment. Her optimal strategy is to send a "non-committal" state in the first round and try to determine her commitment in the third round. Her optimal strategy is to prepare $|\psi\rangle \in \mathcal{A}_1 \otimes \mathcal{A}_2$, where

$$|\psi\rangle = \frac{1}{\sqrt{2}}(|\psi_0\rangle + |\psi_1\rangle),$$

and send $\mathcal{A}_1$ to Bob in the first round. Then, before sending $\mathcal{A}_2$ in the third round, she applies a fidelity achieving unitary operation (see Lemma 4.5) to the state. That is, if she wants to determine her commitment as $a' \in \{0, 1\}$, then she applies the fidelity achieving unitary $U \in \mathcal{L}(\mathcal{A}_2)$ satisfying

$$\langle\psi| I_{\mathcal{A}_1} \otimes U |\psi_{a'}\rangle = F(\rho, \rho_{a'}),$$

where $\rho := \mathrm{Tr}_{\mathcal{A}_2} |\psi\rangle\langle\psi|$. With this strategy, she can successfully cheat with probability

$$\frac{1}{2} + \frac{1}{2} \mathrm{F} \left(\rho_0, \rho_1 \right),$$

see [2, 36, 18]. These cheating strategies are optimal for 3-round BCCF protocols.

We now show that Kitaev's bound can never be attained in a 3-round BCCF protocol.

Lemma 8.4. *Suppose we have two density matrices $\rho, \sigma \in \mathbb{D}^n$. Then*

$$\left(\frac{1}{2} + \frac{\|\rho - \sigma\|_t}{4}\right)\left(\frac{1 + \mathrm{F}(\rho, \sigma)}{2}\right) \geq 1/2,$$

with equality if and only if $\rho \perp \sigma$ or $\rho = \sigma$.

Proof. From Lemmas 4.4 and 4.11, we have that $\|\rho - \sigma\|_t = 2$ if $\rho \perp \sigma$ and $\|\rho - \sigma\|_t = 0$ if $\rho = \sigma$. This proves we have equality above if $\rho = \sigma$ or $\rho \perp \sigma$. Let $F := F(\rho_1, \rho_2)$. Using the first inequality of Lemma 4.11, we have

$$\begin{aligned}
\left(\frac{1}{2} + \frac{\|\rho - \sigma\|_t}{4}\right)\left(\frac{1+F}{2}\right) &\geq \left(\frac{1}{2} + \frac{2(1-F)}{4}\right)\left(\frac{1+F}{2}\right) \\
&= \frac{1}{4}\left(1 + (1-F)\right)(1+F) \\
&= \frac{1}{4}(2-F)(1+F) \\
&= \frac{1}{4}\left(2 + (F - F^2)\right) \\
&\geq 1/2.
\end{aligned}$$

We also see that the last inequality is tight if and only if $F - F^2 = 0$ and that happens if and only if $F = 0$ or $F = 1$. The result now follows from Lemma 4.4. □

This leads to the following corollary.

Corollary 8.5. *$\varepsilon = 1/\sqrt{2} - 1/2$ is impossible in any 3-round BCCF protocol.*

Proof. Suppose there exists a 3-round BCCF protocol with a bias of $\varepsilon = 1/\sqrt{2} - 1/2$. Then this implies $P^*_{A,0} = P^*_{B,0} = 1/\sqrt{2}$. We know Bob can cheat with probability

$$P^*_{B,0} = \frac{1}{2} + \frac{\|\rho_0 - \rho_1\|_t}{4},$$

where $\rho_0 := \mathrm{Tr}_{\mathcal{A}_2}|\psi_0\rangle\langle\psi_0|$ and $\rho_1 := \mathrm{Tr}_{\mathcal{A}_2}|\psi_1\rangle\langle\psi_1|$, and Alice can cheat with probability

$$P^*_{A,0} = \frac{1 + \mathrm{F}(\rho_0, \rho_1)}{2}.$$

From the above lemma, we have $P^*_{A,0}P^*_{B,0} = 1/2$ if and only if $\rho_0 = \rho_1$ or $\rho_0 \perp \rho_1$. If $\rho_0 \perp \rho_1$, Bob could distinguish the two states without error and thus cheat with probability 1. If

$\rho_0 = \rho_1$, Alice can perfectly determine her commitment in the last round since there exists a unitary $U \in \mathcal{L}(\mathcal{A}_2)$ such that

$$(I_{\mathcal{A}_1} \otimes U) |\psi_0\rangle = |\psi_1\rangle ,$$

by Lemma 3.4. In either case, one party can cheat perfectly, a contradiction. □

Thus, the only way we can have $P^*_{A,0}P^*_{B,0} = 1/2$ is if one probability is 1 and the other is 1/2. Notice that $P^*_{B,0} = P^*_{B,1}$ in 3-round BCCF protocols, by nature of the cheating strategies. So we also have $P^*_{A,0}P^*_{B,1} \geq 1/2$ with equality if and only if one probability is 1 and the other is 1/2. Therefore, a maximum cheating probability of $1/\sqrt{2} - 1/2$ is also impossible for 3-round BCCF protocols used for weak coin-flipping.

Using the relationship between fidelity and trace distance, Lemma 4.11, we can show that any 3-round BCCF protocol has a bias of at least 1/4, see [2]. This is better than the $1/\sqrt{2} - 1/2 \approx 0.2071$ bound given in Corollary 8.5.

The optimal cheating strategies for m-round BCCF protocols, with $m \geq 4$, are not known. We can determine them for each specific protocol using semidefinite programming, which brings us to the next section.

8.4 Semidefinite Programming Formulation

The formulation of cheating strategies as an SDP for BCCF protocols is similar to the formulation in the general case. Recall we do not need a message space as we observe it as $\mathcal{A}_i$ and $\mathcal{B}_i$ changing hands. E.g., in the first round, Alice sends $\mathcal{A}_1$ to Bob, so Bob has control of $\mathcal{B}_0 \otimes \mathcal{A}_1 \otimes \mathcal{B}_1 \otimes \cdots \otimes \mathcal{B}_{n-1}$. By what is Bob constrained in this case? Alice starts with her starting state $\rho_0 := |A\rangle\langle A|$. Let ρ_i be Alice's density matrix after Bob sends $\mathcal{B}_i$ to Alice, for $i \in \{1, \ldots, n-1\}$. Bob is constrained by $\operatorname{Tr}_{\mathcal{B}_i}(\rho_i) = \operatorname{Tr}_{\mathcal{A}_i}(\rho_{i-1})$, for all $i \in \{1, \ldots, n-1\}$, since Alice replaces $\mathcal{A}_i$ with $\mathcal{B}_i$, but Bob is not allowed to change the rest of the density matrix. Alice's final density matrix ρ_n is $\operatorname{Tr}_{\mathcal{A}_n}(\rho_{n-1})$ since Alice sends the last message to Bob. Then she measures ρ_n and accepts the outcome $x \in \{0, 1\}$ with probability $\langle \Pi_{A,x}, \rho_n \rangle$. Putting it all together, we have the following result.

Theorem 8.6. *Bob's cheating SDP for a BCCF protocol is the following.*

$$
\begin{array}{lll}
(P_B^x) \quad P_{B,x}^* = \sup & \langle \Pi_{A,x}, \rho_n \rangle & \\
& \rho_0 = |A\rangle\langle A|, & \\
& \mathrm{Tr}_{\mathcal{B}_i}(\rho_i) = \mathrm{Tr}_{\mathcal{A}_i}(\rho_{i-1}),\ \forall i \in \{1,\ldots,n-1\}, & \\
\text{subject to} & \rho_n = \mathrm{Tr}_{\mathcal{A}_n}(\rho_{n-1}), & \\
& \rho_0 \in \Sigma_+^{\mathcal{A}_0 \otimes \mathcal{A}_1 \otimes \cdots \otimes \mathcal{A}_n}, & \\
& \rho_i \in \Sigma_+^{\mathcal{A}_0 \otimes \mathcal{B}_1 \otimes \cdots \otimes \mathcal{B}_i \otimes \mathcal{A}_{i+1} \otimes \cdots \otimes \mathcal{A}_n},\ \forall i \in \{1,\ldots,n-1\}, & \\
& \rho_n \in \Sigma_+^{\mathcal{A}_0 \otimes \mathcal{B}_1 \otimes \cdots \otimes \mathcal{B}_{n-1}}. &
\end{array}
$$

Using adjoint operators and replacing ρ_0 with $|A\rangle\langle A|$, we get the following smaller, yet equivalent, SDP.

$$
\begin{array}{lll}
(P_B^x) \quad P_{B,x}^* = \sup & \left\langle \mathrm{Tr}^*_{\mathcal{A}_n}(\Pi_{A,x}), \rho_{n-1} \right\rangle & \\
& \mathrm{Tr}_{\mathcal{B}_1}(\rho_1) = \mathrm{Tr}_{\mathcal{A}_1}|A\rangle\langle A|, & \\
\text{subject to} & \mathrm{Tr}_{\mathcal{B}_i}(\rho_i) = \mathrm{Tr}_{\mathcal{A}_i}(\rho_{i-1}),\ \forall i \in \{2,\ldots,n-1\}, & \\
& \rho_i \in \Sigma_+^{\mathcal{A}_0 \otimes \mathcal{B}_1 \otimes \cdots \otimes \mathcal{B}_i \otimes \mathcal{A}_{i+1} \otimes \cdots \otimes \mathcal{A}_n},\ \forall i \in \{1,\ldots,n-1\}. &
\end{array}
$$

We can examine its dual, which is obtained in the same manner as in the general coin-flipping case.

$$
\begin{array}{lll}
(D_B^x) \quad \inf & \langle Z_1, \mathrm{Tr}_{\mathcal{A}_1}|A\rangle\langle A| \rangle & \\
& \mathrm{Tr}^*_{\mathcal{B}_i}(Z_i) \succeq \mathrm{Tr}^*_{\mathcal{A}_{i+1}}(Z_{i+1}),\ \forall i \in \{1,\ldots,n-1\}, & \\
\text{subject to} & \mathrm{Tr}^*_{\mathcal{B}_{n-1}}(Z_{n-1}) \succeq \mathrm{Tr}^*_{\mathcal{A}_n}(\Pi_{A,x}), & \\
& Z_1 \in \Sigma^{\mathcal{A}_0 \otimes \mathcal{A}_2 \otimes \cdots \otimes \mathcal{A}_n}, & \\
& Z_i \in \Sigma^{\mathcal{A}_0 \otimes \mathcal{B}_1 \otimes \cdots \otimes \mathcal{B}_{i-1} \otimes \mathcal{A}_{i+1} \otimes \cdots \otimes \mathcal{A}_n}, \forall i \in \{2,\ldots,n-1\}. &
\end{array}
$$

We similarly formulate Alice's cheating SDP as follows.

Theorem 8.7. *Alice's cheating SDP for a BCCF protocol is the following.*

$$
\begin{array}{lll}
(P_A^x) \quad P_{A,x}^* = \sup & \langle \Pi_{B,x}, \sigma_n \rangle & \\
& \mathrm{Tr}_{\mathcal{A}_1}(\sigma_1) = |B\rangle\langle B|, & \\
& \mathrm{Tr}_{\mathcal{A}_{i+1}}(\sigma_{i+1}) = \mathrm{Tr}_{\mathcal{B}_i}(\sigma_i),\ \forall i \in \{1,\ldots,n-1\}, & \\
\text{subject to} & \sigma_i \in \Sigma_+^{\mathcal{B}_0 \otimes \mathcal{A}_1 \otimes \cdots \otimes \mathcal{A}_i \otimes \mathcal{B}_i \otimes \cdots \otimes \mathcal{B}_{n-1}},\ \forall i \in \{1,\ldots,n-1\}, & \\
& \sigma_n \in \Sigma_+^{\mathcal{B}_0 \otimes \mathcal{A}_1 \otimes \cdots \otimes \mathcal{A}_n}. &
\end{array}
$$

Notice Alice sends the first message, which explains the first constraint. From there, she replaces each of Bob's messages with one of her own, explaining the rest of the constraints. The dual of Alice's cheating SDP is given below.

$$\begin{aligned}
(D_A^x) \quad \inf \quad & \langle Y_1, |B\rangle\langle B| \rangle & & \\
& \mathrm{Tr}^*_{\mathcal{A}_i}(Y_i) & \succeq & \mathrm{Tr}^*_{\mathcal{B}_i}(Y_{i+1}),\ \forall i \in \{1,\ldots,n-1\}, \\
& \mathrm{Tr}^*_{\mathcal{A}_n}(Y_n) & \succeq & \Pi_{B,x}, \\
\text{subject to} \quad & Y_1 & \in & \Sigma^{\mathcal{B}_0 \otimes \mathcal{B}_1 \otimes \cdots \otimes \mathcal{B}_{n-1}}, \\
& Y_i & \in & \Sigma^{\mathcal{B}_0 \otimes \mathcal{A}_1 \otimes \cdots \otimes \mathcal{A}_{i-1} \otimes \mathcal{B}_i \otimes \cdots \otimes \mathcal{B}_{n-1}},\ \forall i \in \{2,\ldots,n-1\}, \\
& Y_n & \in & \Sigma^{\mathcal{B}_0 \otimes \mathcal{A}_1 \otimes \cdots \otimes \mathcal{A}_{n-1}}.
\end{aligned}$$

By Lemma 7.4, we know that every coin-flipping protocol has an achievable optimal strategy. This result is therefore true for BCCF protocols as well. We can also check this directly by the following lemma.

Lemma 8.8. *The primal-dual pairs (P_B^x, D_B^x) and (P_A^x, D_A^x) both have zero duality gap and primal attainment.*

Proof. By playing honestly, both Alice and Bob can achieve feasible solutions for their respective primal problems. By setting $Z_i = (n-i+2)I,\ \forall i \in \{1,\ldots,n-1\}$, for the corresponding identity matrix I, yields a strictly feasible solution for (D_B^x). Similarly, setting $Y_i = (n-i+2)I,\ \forall i \in \{1,\ldots,n\}$, for the corresponding identity matrix I, yields a strictly feasible solution for (D_A^x). The result now holds by strong duality. □

We have the immediate corollary of Theorem 7.6.

Corollary 8.9. *We have Kitaev's lower bound for BCCF protocols, i.e., for each $x \in \{0,1\}$, we have*

$$\max\{P_{A,x}^*, P_{B,x}^*\} \geq \frac{1}{\sqrt{2}}.$$

8.5 An Upper Bound on the Best Possible Bias

The first protocol with bias less than 1 was found by Aharonov et al. [1]. This was improved by Ambainis [2] and Spekkens and Rudolph [36] by showing $\varepsilon = 1/4$ is possible. We now give an alternative proof that there exists a protocol with $\varepsilon = 1/4$.

To fully specify a BCCF protocol, we need to specify Alice's message spaces $\mathcal{A}_i$, for $i \in \{1, \ldots, n\}$, Bob's message spaces $\mathcal{B}_i$, for $i \in \{1, \ldots, n-1\}$, and their respective starting states. We examine a 3-round protocol by Kerenidis and Nayak [18] which has the starting states

$$|\psi_a\rangle = \frac{1}{\sqrt{2}}(|aa\rangle + |22\rangle) \in \mathbb{C}^3 \otimes \mathbb{C}^3,$$

for Alice and

$$|\phi_b\rangle = |b\rangle \in \mathbb{C}^2,$$

for Bob.

We append an extra qubit $|a\rangle$ onto the starting state $|\psi_a\rangle$, for each $a \in \{0, 1\}$, so that they satisfy the condition $|\psi_0\rangle \perp |\psi_1\rangle$. Thus, we have the states now being

$$|\psi_a\rangle = \frac{1}{\sqrt{2}}(|aa\rangle + |22\rangle)\,|a\rangle\,,$$

where $\mathcal{A}_1 = \mathbb{C}^3$ and $\mathcal{A}_2 = \mathbb{C}^3 \otimes \mathbb{C}^2$.

We show that this protocol has a bias of $1/4$. Firstly, we show that $\varepsilon = 1/4$ is a lower bound on the bias (similarly done in [18]).

From Section 8.3, we know Bob can successfully cheat with probability

$$\frac{1}{2} + \frac{1}{4}\left\|\rho_0 - \rho_1\right\|_t,$$

where

$$\rho_0 = \mathrm{Tr}_{\mathcal{A}_2}|\psi_0\rangle\langle\psi_0| = \frac{1}{2}\begin{bmatrix} 1 & & \\ & 0 & \\ & & 1 \end{bmatrix} \quad \text{and} \quad \rho_1 = \mathrm{Tr}_{\mathcal{A}_2}|\psi_1\rangle\langle\psi_1| = \frac{1}{2}\begin{bmatrix} 0 & & \\ & 1 & \\ & & 1 \end{bmatrix}.$$

Thus, he can cheat with probability

$$\begin{aligned} \frac{1}{2} + \frac{1}{4}\left\|\rho_0 - \rho_1\right\|_t &= \frac{1}{2} + \frac{1}{4}\left\|\begin{bmatrix} 1/2 & & \\ & -1/2 & \\ & & 0 \end{bmatrix}\right\|_t \\ &= \frac{1}{2} + \frac{1}{4}\left(\frac{1}{2} + \frac{1}{2}\right) \\ &= \frac{3}{4}. \end{aligned}$$

Now that we know that Bob can cheat with probability at least 3/4, we use semidefinite programming to show that neither party can bias the outcome with probability higher than 3/4. We do this by finding dual solutions to the four dual SDP's each with objective value 3/4.

The dual of Bob's cheating SDP in this case is the following.

$$\begin{array}{lll} (D_B^x) & P^*_{B,x} = \inf & \langle \mathrm{Tr}_{\mathcal{A}_1} |A\rangle\langle A|, Z\rangle \\ & \text{subject to} & \begin{array}{rcl} \mathrm{Tr}^*_{\mathcal{B}_1}(Z) & \succeq & \mathrm{Tr}^*_{\mathcal{A}_2}(\Pi_{A,x}), \\ Z & \in & \Sigma^{\mathcal{A}_0 \otimes \mathcal{A}_2}. \end{array} \end{array}$$

The following lemma helps prove positive semidefiniteness of certain matrices.

Lemma 8.10. *Suppose we are given a matrix $X \in \Sigma^n$ and there exists a permutation matrix P such that PXP^T is block diagonal with blocks of size 1×1 and 2×2. Then $X \in \Sigma^n_+$ if X has nonnegative diagonal and the 2×2 blocks (upon permutation) have nonnegative determinant.*

Proof. Suppose we are given such a matrix $X \in \Sigma^n$. A block diagonal matrix is positive semidefinite if and only if the blocks are positive semidefinite, this follows from definition. Since $P(\cdot)P^T \in \mathrm{Aut}\,(\Sigma^n_+)$, we only need to check if the blocks in the diagonal of PXP^T are positive semidefinite. Since X has nonnegative diagonal, the 1×1 blocks are positive semidefinite. Arbitrarily choose a 2×2 block of PXP^T. Suppose this block has eigenvalues λ_1 and λ_2. Since it has nonnegative determinant, we know $\lambda_1 \lambda_2 \geq 0$. Also, since X has nonnegative diagonal, we know the trace of this block is also nonnegative. Thus, $\lambda_1 + \lambda_2 \geq 0$. This implies $\lambda_1, \lambda_2 \geq 0$, showing it is positive semidefinite. □

The operation $P(\cdot)P^T$, where P is a permutation matrix, is called a *symmetric permutation*.

Consider the following parametrized matrix.

$$Z(t) := \begin{bmatrix} 1 & & & & & & & & & & & & & & & & & & & & & & & \\ & 1 & & & & & & & & & & & & & & & & & & & & & & \\ & & 1 & & & & & & & & & & & & & & & & & & & & & \\ & & & 1 & & & & & & & & & & & & & & & & & & & & \\ & & & & t & & & & & & & & & & & & & & & & & & -\sqrt{t(t-1)} \\ & & & & & & & & 1 & & & & & & & & & & & & & & \\ & & & & & & & & & 1 & & & & & & & & & & & & & \\ & & & & & & & & & & 1 & & & & & & & & & & & & \\ & & & & & & & & & & & 1 & & & & & & & & & & & \\ & & & & & & & & & & & & 1 & & & & & & & & & & \\ & & & & & & & & & & & & & 1 & & & & & & & & & \\ & & & & & -\sqrt{t(t-1)} & & & & & & & & & & & & & & & & t & \end{bmatrix},$$

where $t \geq 1$. We now prove it is feasible for (D_B^0) and (D_B^1). We can write $\mathrm{Tr}^*_{\mathcal{A}_2}(\Pi_{A,0}) \in \mathcal{A}_0 \otimes \mathcal{B}_1 \otimes \mathcal{A}_2$ as

$$\mathrm{Tr}^*_{\mathcal{A}_2}(\Pi_{A,0}) = \mathrm{Diag}\left([1,1,1,1,1,1,0,0,0,0,0,0,0,0,0,0,0,0,1,1,1,1,1,1]^T\right)$$

and $\mathrm{Tr}^*_{\mathcal{A}_2}(\Pi_{A,1}) \in \mathcal{A}_0 \otimes \mathcal{B}_1 \otimes \mathcal{A}_2$ as

$$\mathrm{Tr}^*_{\mathcal{A}_2}(\Pi_{A,1}) = I_{24} - \mathrm{Tr}^*_{\mathcal{A}_2}(\Pi_{A,0}),$$

where $I_n := I_{\mathbb{C}^n}$. Expanding, we can write $\mathrm{Tr}^*_{\mathcal{B}_1}(Z(t)) \in \mathcal{A}_0 \otimes \mathcal{B}_1 \otimes \mathcal{A}_2$ as

$$\left[\begin{array}{cccccccccccccccc}
1 & & & & & & & & & & & & & & & \\
 & 1 & & & & & & & & & & & & & & \\
 & & 1 & & & & & & & & & & & & & \\
 & & & 1 & & & & & & & & & & & & \\
 & & & & t & & & & & & & & & & s & \\
 & & & & & 1 & & & & & & & & & & \\
 & & & & & & 1 & & & & & & & & & \\
 & & & & & & & 1 & & & & & & & & \\
 & & & & & & & & 1 & & & & & & & \\
 & & & & & & & & & 1 & & & & & & \\
 & & & & & & & & & & t & & & & & s \\
 & & & & & & & & & & & 1 & & & & \\
 & & & & & & & & & & & & 1 & & & \\
 & & & & & & & & & & & & & 1 & & \\
 & & & & & & & & & & & & & & 1 & \\
 & & & & & & & & & & & & & & & 1 \\
 & & & & s & & & & & & & t & & & & \\
 & & & & & & & & & & & & 1 & & & \\
 & & & & & & & & & & & & & 1 & & \\
 & & & & & & & & & & & & & & 1 & \\
 & & & & & & & & & & & & & & & 1 \\
 & & & & & & & & & & & & & & & \\
 & & & & & & & & & & s & & & & & t
\end{array}\right],$$

where $s := -\sqrt{t(t-1)}$. (The s's align with the t's in the matrix.) By Lemma 8.10, we have $\mathrm{Tr}^*_{\mathcal{B}_1}(Z) - \mathrm{Tr}^*_{\mathcal{A}_2}(\Pi_{A,0}) \succeq 0$ and $\mathrm{Tr}^*_{\mathcal{B}_1}(Z) - \mathrm{Tr}^*_{\mathcal{A}_2}(\Pi_{A,1}) \succeq 0$ since both have nonzero diagonal and

$$\det\left(\begin{bmatrix} t & -\sqrt{t(t-1)} \\ -\sqrt{t(t-1)} & t-1 \end{bmatrix}\right) = 0,$$

for $t \geq 1$. Thus, $Z(t)$ is feasible in (D^0_B) and in (D^1_B).

Note we have

$$\mathrm{Tr}_{\mathcal{A}_1}|A\rangle\langle A| := \frac{1}{4}\begin{bmatrix} 1 & & & & & & & & & & & & & & & \\ & 0 & & & & & & & & & & & & & & \\ & & 0 & & & & & & & & & & & & & \\ & & & 0 & & & & & & & & & & & & \\ & & & & 1 & & & & & & & & & & 1 & \\ & & & & & 0 & & & & & & & & & & \\ & & & & & & 0 & & & & & & & & & \\ & & & & & & & 0 & & & & & & & & \\ & & & & & & & & 0 & & & & & & & \\ & & & & & & & & & 1 & & & & & & \\ & & & & & & & & & & 0 & & & & & \\ & & & & 1 & & & & & & & & & & 1 & \end{bmatrix},$$

so the objective value of this feasible solution is

$$\begin{aligned} \lim_{t\to\infty}\langle \mathrm{Tr}_{\mathcal{A}_1}|A\rangle\langle A|, Z(t)\rangle &= \lim_{t\to\infty}\frac{1}{4}(1+1+2t-\sqrt{t(t-1)}) \\ &= \frac{1}{2}+\frac{1}{2}\lim_{t\to\infty}\left[t-\sqrt{t(t-1)}\right] \\ &= \frac{3}{4}. \end{aligned}$$

Therefore, Bob can cheat towards either outcome with a probability bounded above by 3/4. We now bound Alice's cheating probabilities.

The dual of Alice's cheating SDP is given as the following SDP.

$$\begin{aligned} (D_A^x) \quad P^*_{A,x} = \inf \quad & \langle Y_1, |B\rangle\langle B|\rangle & & \\ & \mathrm{Tr}^*_{\mathcal{A}_1}(Y_1) & \succeq & \mathrm{Tr}^*_{\mathcal{B}_1}(Y_2), \\ \text{subject to} \quad & \mathrm{Tr}^*_{\mathcal{A}_2}(Y_2) & \succeq & \Pi_{B,x}, \\ & Y_1 & \in & \Sigma^{\mathcal{B}_0\otimes\mathcal{B}_1}, \\ & Y_2 & \in & \Sigma^{\mathcal{B}_0\otimes\mathcal{A}_1}. \end{aligned}$$

Consider the parametrized pair $(Y_1(t), Y_2)$ given as

$$(Y_1(t), Y_2) = \left(\begin{bmatrix} t & & & -\sqrt{t\left(t-\frac{3}{2}\right)} \\ & \frac{3}{2} & & \\ & & \frac{3}{2} & \\ -\sqrt{t\left(t-\frac{3}{2}\right)} & & & t \end{bmatrix}, \begin{bmatrix} \frac{3}{2} & & & & & & \\ & 0 & & & & & \\ & & \frac{3}{4} & & & & \\ & & & 0 & & & \\ & & & & \frac{3}{2} & & \\ & & & & & \frac{3}{4} & \end{bmatrix} \right),$$

for $t \geq 3/2$. We now show this is feasible in (D_A^0). We can write $\mathrm{Tr}^*_{\mathcal{A}_1}(Y_1(t)) \in \mathcal{B}_0 \otimes \mathcal{A}_1 \otimes \mathcal{B}_1$ as

$$\mathrm{Tr}^*_{\mathcal{A}_1}(Y_1(t)) := \begin{bmatrix} t & & & & & & s & & & & & \\ & \frac{3}{2} & & & & & & & & & & \\ & & t & & & & & & s & & & \\ & & & \frac{3}{2} & & & & & & & & \\ & & & & t & & & & & & s & \\ & & & & & \frac{3}{2} & & & & & & \\ & & & & & & \frac{3}{2} & & & & & \\ s & & & & & & & t & & & & \\ & & & & & & & & \frac{3}{2} & & & \\ & & s & & & & & & & t & & \\ & & & & & & & & & & \frac{3}{2} & \\ & & & & s & & & & & & & t \end{bmatrix},$$

where $s := -\sqrt{t(t-3/2)}$, and $\mathrm{Tr}^*_{\mathcal{B}_1}(Y_2) \in \mathcal{B}_0 \otimes \mathcal{A}_1 \otimes \mathcal{B}_1$ as

$$\mathrm{Tr}^*_{\mathcal{B}_1}(Y_2) := \begin{bmatrix} \frac{3}{2} & & & & & & & & & & & & & & & \\ & \frac{3}{2} & & & & & & & & & & & & & & \\ & & 0 & & & & & & & & & & & & & \\ & & & 0 & & & & & & & & & & & & \\ & & & & \frac{3}{4} & & & & & & & & & & & \\ & & & & & \frac{3}{4} & & & & & & & & & & \\ & & & & & & 0 & & & & & & & & & \\ & & & & & & & 0 & & & & & & & & \\ & & & & & & & & \frac{3}{2} & & & & & & & \\ & & & & & & & & & \frac{3}{2} & & & & & & \\ & & & & & & & & & & \frac{3}{4} & & & & & \\ & & & & & & & & & & & \frac{3}{4} & & & & \end{bmatrix}.$$

By Lemma 8.10, $\mathrm{Tr}^*_{\mathcal{A}_1}(Y_1(t)) - \mathrm{Tr}^*_{\mathcal{A}_2}(Y_2) \succeq 0$, for $t \geq 3/2$, since it has nonnegative diagonal and, noting $s = -\sqrt{t(t-3/2)}$, we have

$$\det\left(\begin{bmatrix} t-3/2 & s \\ s & t \end{bmatrix}\right), \det\left(\begin{bmatrix} t & s \\ s & t-3/2 \end{bmatrix}\right), \det\left(\begin{bmatrix} t-3/4 & s \\ s & t-3/4 \end{bmatrix}\right) \geq 0.$$

Thus, the first constraint is satisfied. To show the second constraint is satisfied, we need to show that $\mathrm{Tr}^*_{\mathcal{A}_2}(Y_2) \succeq \Pi_{B,0}$. We know

$$\mathrm{Tr}^*_{\mathcal{A}_2}(Y_2) = Y_2 \otimes I_3 \otimes I_2 \in \mathcal{B}_0 \otimes \mathcal{A}_1 \otimes \mathcal{A}_2$$

and

$$\Pi_{B,0} = |0\rangle\langle 0| \otimes |\psi'_0\rangle\langle\psi'_0| \otimes |0\rangle\langle 0| + |1\rangle\langle 1| \otimes |\psi'_1\rangle\langle\psi'_1| \otimes |1\rangle\langle 1| \in \mathcal{B}_0 \otimes \mathcal{A}_1 \otimes \mathcal{A}_2,$$

where $|\psi'_a\rangle = \frac{1}{\sqrt{2}}(|aa\rangle + |22\rangle)$. Therefore, we can write $\mathrm{Tr}^*_{\mathcal{A}_1}(Y_2) - \Pi_{B,0}$ as

$$\begin{aligned}
\mathrm{Tr}^*_{\mathcal{A}_1}(Y_2) - \Pi_{B,0} &= Y_2 \otimes I_3 \otimes I_2 - |0\rangle\langle 0| \otimes |\psi'_0\rangle\langle\psi'_0| \otimes |0\rangle\langle 0| - |1\rangle\langle 1| \otimes |\psi'_1\rangle\langle\psi'_1| \otimes |1\rangle\langle 1| \\
&= Y_2 \otimes I_3 \otimes |0\rangle\langle 0| + Y_2 \otimes I_3 \otimes |1\rangle\langle 1| \\
&\quad -|0\rangle\langle 0| \otimes |\psi'_0\rangle\langle\psi'_0| \otimes |0\rangle\langle 0| - |1\rangle\langle 1| \otimes |\psi'_1\rangle\langle\psi'_1| \otimes |1\rangle\langle 1| \\
&= (Y_2 \otimes I_3 - |0\rangle\langle 0| \otimes |\psi'_0\rangle\langle\psi'_0|) \otimes |0\rangle\langle 0| \\
&\quad +(Y_2 \otimes I_3 - |1\rangle\langle 1| \otimes |\psi'_1\rangle\langle\psi'_1|) \otimes |1\rangle\langle 1|.
\end{aligned}$$

Thus, we have $\mathrm{Tr}^*_{\mathcal{A}_1}(Y_2) - \Pi_{B,0} \succeq 0$ if and only if

$$Y_2 \otimes I_3 - |0\rangle\langle 0| \otimes |\psi'_0\rangle\langle\psi'_0| \succeq 0 \quad \text{and} \quad Y_2 \otimes I_3 - |1\rangle\langle 1| \otimes |\psi'_1\rangle\langle\psi'_1| \succeq 0.$$

Since Y_2 is diagonal, we can write $Y_2 \otimes I_3$ as $\begin{bmatrix} Y' & 0 \\ 0 & Y'' \end{bmatrix}$, where $Y', Y'' \in \mathbb{M}^9$. We need to check that $Y' \succeq |\psi'_0\rangle\langle\psi'_0|$ and $Y'' \succeq |\psi'_1\rangle\langle\psi'_1|$. Notice that

$$Y' = \begin{bmatrix}
\frac{3}{2} & & & & & & & & \\
& \frac{3}{2} & & & & & & & \\
& & \frac{3}{2} & & & & & & \\
& & & 0 & & & & & \\
& & & & 0 & & & & \\
& & & & & 0 & & & \\
& & & & & & \frac{3}{4} & & \\
& & & & & & & \frac{3}{4} & \\
& & & & & & & & \frac{3}{4}
\end{bmatrix}$$

and

$$|\psi_0'\rangle\langle\psi_0'| = \begin{bmatrix} \frac{1}{2} & & & & & & & & \frac{1}{2} \\ & 0 & & & & & & & \\ & & 0 & & & & & & \\ & & & 0 & & & & & \\ & & & & 0 & & & & \\ & & & & & 0 & & & \\ & & & & & & 0 & & \\ & & & & & & & 0 & \\ \frac{1}{2} & & & & & & & & \frac{1}{2} \end{bmatrix}.$$

Thus, $Y' \succeq |\psi_0'\rangle\langle\psi_0'|$ by Lemma 8.10. We can write the other two matrices as

$$Y'' = \begin{bmatrix} 0 & & & & & & & & \\ & 0 & & & & & & & \\ & & 0 & & & & & & \\ & & & \frac{3}{2} & & & & & \\ & & & & \frac{3}{2} & & & & \\ & & & & & \frac{3}{2} & & & \\ & & & & & & \frac{3}{4} & & \\ & & & & & & & \frac{3}{4} & \\ & & & & & & & & \frac{3}{4} \end{bmatrix}$$

and

$$|\psi_1'\rangle\langle\psi_1'| = \begin{bmatrix} 0 & & & & & & & & \\ & 0 & & & & & & & \\ & & 0 & & & & & & \\ & & & 0 & & & & & \\ & & & & \frac{1}{2} & & & \frac{1}{2} & \\ & & & & & 0 & & & \\ & & & & & & 0 & & \\ & & & & & & & 0 & \\ & & & & \frac{1}{2} & & & \frac{1}{2} & \end{bmatrix}.$$

Therefore, $Y'' \succeq |\psi_1'\rangle\langle\psi_1'|$, by Lemma 8.10, showing the second constraint is satisfied.

To find the objective value, note that

$$|B\rangle\langle B| = \frac{1}{2}\begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix}.$$

The objective value can now be written as

$$\begin{aligned} \langle Y_1, |B\rangle\langle B|\rangle &= \lim_{t\to\infty} \frac{1}{2}\left(2t - 2\sqrt{t(t-3/2)}\right) \\ &= \lim_{t\to\infty} \left[t - \sqrt{t(t-3/2)}\right] \\ &= \frac{3}{4}. \end{aligned}$$

Thus, Alice can not cheat towards 0 with probability higher than 3/4.

We can similarly prove that

$$(Y_1'(t), Y_2') = \left(\begin{bmatrix} t & & -\sqrt{t\left(t-\frac{3}{2}\right)} \\ & \frac{3}{2} & \\ & \frac{3}{2} & \\ -\sqrt{t\left(t-\frac{3}{2}\right)} & & t \end{bmatrix}, \begin{bmatrix} 0 & & & & & \\ & \frac{3}{2} & & & & \\ & & \frac{3}{4} & & & \\ & & & \frac{3}{2} & & \\ & & & & 0 & \\ & & & & & \frac{3}{4} \end{bmatrix} \right),$$

for $t \geq 3/2$, is a feasible solution for (D_A^1). Since $Y_1'(t) = Y_1(t)$, it also has an objective value of $3/4$ as $t \to \infty$. Therefore, Alice can not cheat towards 1 with probability higher than $3/4$.

We have just proved the following result.

Theorem 8.11. *The above protocol has a bias of* $1/4$.

Recall that Kitaev's lower bound on any protocol is $\varepsilon \geq 1/\sqrt{2} - 1/2 \approx 0.2071$. The difference between the best known upper bound and best known lower bound for the bias is $1/4 - (1/\sqrt{2} - 1/2) \approx 0.0429$, which is quite small, but there is room for improvement.

Chapter 9

Asymmetry in Cheating Probabilities

9.1 An Asymmetric Protocol

In every 3-round BCCF protocol we have that $P^*_{A,0} = P^*_{A,1}$ and $P^*_{B,0} = P^*_{B,1}$ since the optimal cheating strategies do not depend on the desired outcome. However, while solving for the bias of 4-round protocols, we noticed that sometimes it is the case that $P^*_{A,0} \neq P^*_{A,1}$ and $P^*_{B,0} \neq P^*_{B,1}$. We have isolated a family of protocols which we noticed had this asymmetry in the cheating probabilities.

Consider the following class of 4-round BCCF protocols where all the messages are qubits.

$$|\psi_0\rangle := \sqrt{m}\,|00\rangle + \sqrt{1-m}\,|11\rangle\,, \quad |\psi_1\rangle := |01\rangle\,,$$

$$|\phi_0\rangle := \sqrt{y}\,|00\rangle + \sqrt{1-y}\,|11\rangle\,, \quad |\phi_1\rangle := |10\rangle\,,$$

where $m, y \geq 0$ satisfy $m + y \leq 1$.

We argue that $P^*_{B,0} = 1$ and $P^*_{B,1} \leq 1$, provided that $m + y \leq 1$, and $P^*_{B,1} < 1$ if we further assume $m \neq 0$ and $y \neq 0$.

This is interesting since the optimal cheating strategies for 3-round BCCF protocols were independent of the desired outcome and now we claim that there exist protocols where cheating strategies do depend on the desired outcome.

Note that this BCCF protocol has 4-rounds and our SDP formulations are for an odd number of rounds. We formulate the cheating SDP's for 4-round BCCF protocols in the

same manner and are given below.

$$\begin{array}{lll} (P_B^x) \quad P_{B,x}^* = \sup & \langle \Pi_{A,x}, \rho_2 \rangle & \\ & \mathrm{Tr}_{\mathcal{B}_1}(\rho_1) &= \mathrm{Tr}_{\mathcal{A}_1} |A\rangle\langle A|, \\ \text{subject to} & \mathrm{Tr}_{\mathcal{B}_2}(\rho_2) &= \mathrm{Tr}_{\mathcal{A}_2}(\rho_1), \\ & \rho_1 &\in \Sigma_+^{\mathcal{A}_0 \otimes \mathcal{B}_1 \otimes \mathcal{A}_2}, \\ & \rho_2 &\in \Sigma_+^{\mathcal{A}_0 \otimes \mathcal{B}_1 \otimes \mathcal{B}_2}. \end{array}$$

The dual of Bob's cheating SDP is

$$\begin{array}{lll} (D_B^x) \quad \inf & \langle X_1, \mathrm{Tr}_{\mathcal{A}_1} |A\rangle\langle A| \rangle & \\ & \mathrm{Tr}^*_{\mathcal{B}_1}(X_1) &\succeq \mathrm{Tr}^*_{\mathcal{A}_2}(X_2), \\ \text{subject to} & \mathrm{Tr}^*_{\mathcal{B}_2}(X_2) &\succeq \Pi_{A,x}, \\ & X_1 &\in \Sigma^{\mathcal{A}_0 \otimes \mathcal{A}_2}, \\ & X_2 &\in \Sigma^{\mathcal{A}_0 \otimes \mathcal{B}_1}. \end{array}$$

Assume $m + y \leq 1$, $m \neq 0$, and $y \neq 0$. To show Bob can force outcome 0 with probability 1, we find optimal solutions for the primal and dual each having objective value 1. Since the objective value of the primal is a probability, we know it is bounded above by 1. We find the optimal dual solution to verify the zero duality gap.

Consider (ρ_1, ρ_2) defined below. We show it is a feasible solution for (P_B^0).

$$\rho_1 := \frac{1}{2} \begin{bmatrix} 0 & & & & & & & \\ & y & & & & & & \\ & & m & & & & & \sqrt{m} \\ & & & 1-m-y & & & & \\ & & & & 0 & & & \\ & & & & & 0 & & \\ & & & & & & 0 & \\ & & \sqrt{m} & & & & & 1 \end{bmatrix}$$

and

$$\rho_2 := \frac{1}{2}\begin{bmatrix} y & & & \sqrt{y(1-y)} & & & \\ & 0 & 0 & & & & \\ & 0 & 0 & & & & \\ \sqrt{y(1-y)} & & & 1-y & & & \\ & & & & 0 & & \\ & & & & & 0 & \\ & & & & & & 1 \\ & & & & & & & 0 \end{bmatrix}.$$

By Lemma 8.10, we see that the above two matrices are positive semidefinite.

We note that

$$\mathrm{Tr}_{\mathcal{A}_2}(\rho_1) = \frac{1}{2}\begin{bmatrix} y & & & \\ & 1-y & & \\ & & 0 & \\ & & & 1 \end{bmatrix} = \mathrm{Tr}_{\mathcal{B}_2}(\rho_2),$$

so the second constraint is satisfied. Also

$$\mathrm{Tr}_{\mathcal{B}_1}(\rho_1) = \frac{1}{2}\begin{bmatrix} m & & & \sqrt{m} \\ & 1-m & & \\ & & 0 & \\ \sqrt{m} & & & 1 \end{bmatrix} = \mathrm{Tr}_{\mathcal{A}_1}|A\rangle\langle A|,$$

where

$$|A\rangle\langle A| = \frac{1}{2}\begin{bmatrix} m & & & \sqrt{m(1-m)} & & \sqrt{m} & & \\ & 0 & 0 & & & & & \\ & 0 & 0 & & & & & \\ \sqrt{m(1-m)} & & & 1-m & & \sqrt{1-m} & & \\ & & & & 0 & & & \\ \sqrt{m} & & & \sqrt{1-m} & & 1 & & \\ & & & & & & 0 & \\ & & & & & & & 0 \end{bmatrix}.$$

This shows the first constraint is satisfied. Therefore, (ρ_1, ρ_2) is feasible for (P_B^0). Since

$\rho_2 = \frac{1}{2}\Pi_{A,0}$, the objective value is

$$\langle \Pi_{A,x}, \rho_2 \rangle = \frac{1}{2}\mathrm{Tr}\left(\Pi_{A,0}\Pi_{A,0}\right) = \frac{1}{2}\mathrm{Tr}\left(\Pi_{A,0}\right) = \frac{1}{2}\mathrm{rank}\left(\Pi_{A,0}\right) = 1.$$

To prove this is an optimal solution for (P_B^0), note that $(X_1, X_2) = (I_{\mathcal{A}_0\otimes\mathcal{A}_1}, I_{\mathcal{A}_0\otimes\mathcal{B}_1})$ is a feasible solution to (D_B^0) giving objective value

$$\langle I, \mathrm{Tr}_{\mathcal{A}_1}|A\rangle\langle A|\rangle = \mathrm{Tr}\left(\mathrm{Tr}_{\mathcal{A}_1}|A\rangle\langle A|\right) = 1.$$

Therefore, both the primal and dual solutions above are optimal by weak duality.

We now find $P_{B,1}^*$ by finding an explicit formula for the optimal value of (P_B^1). Define $\gamma := \sqrt{(m+1)^2 - 4my}$. Recall that $m + y \leq 1$ and $m, y \neq 0$. Under these assumptions, we have that $(m+1)^2 - 4my > 0$. Thus, γ is well-defined and positive.

We show that (ρ_1', ρ_2'), below, is a feasible solution to (P_B^1).

$$\rho_1' := \begin{bmatrix}
\alpha & & & & & & \frac{\alpha}{\sqrt{m}} & & \\
& 0 & & & & & & & \\
& & \frac{m}{2} - \alpha & & & & & \frac{\sqrt{m}}{2} - \frac{\alpha}{\sqrt{m}} & \\
& & & \frac{1}{2} - \frac{m}{2} & & & & & \\
& & & & 0 & & & & \\
\frac{\alpha}{\sqrt{m}} & & & & & & \frac{\alpha}{m} & & \\
& & & & & & & 0 & \\
& & \frac{\sqrt{m}}{2} - \frac{\alpha}{\sqrt{m}} & & & & & & \frac{1}{2} - \frac{\alpha}{m}
\end{bmatrix}$$

and

$$\rho_2' := \begin{bmatrix} \frac{\alpha}{2} & & & & & & & \\ & \frac{\alpha}{2} & & & & & & \\ & & \frac{1}{2}-\alpha & & & & & \\ & & & 0 & & & & \\ & & & & \frac{\alpha}{m} & & & \sqrt{\frac{\alpha}{m}\left(\frac{1}{2}-\frac{\alpha}{m}\right)} \\ & & & & & 0 & 0 & \\ & & & & & 0 & 0 & \\ & & & & \sqrt{\frac{\alpha}{m}\left(\frac{1}{2}-\frac{\alpha}{m}\right)} & & & \frac{1}{2}-\frac{\alpha}{m} \end{bmatrix},$$

where $\alpha = \frac{m}{4}\left(1 - \frac{m+1-2y}{\gamma}\right)$. By Lemma 8.10, the above two matrices are positive semidefinite if we can show the following four conditions.

1. $\alpha \leq \frac{m}{2}$,
2. $m \leq 1$,
3. $\alpha \leq \frac{1}{2}$, and
4. $\alpha \geq 0$.

Since $m \leq 1$, it suffices to show $0 \leq \alpha \leq \frac{m}{2}$. To do so, we present the following lemma.

Lemma 9.1. *For γ as defined above, we have*

$$\gamma \geq m+1-2y \quad \textit{and} \quad \gamma \geq -(m+1-2y).$$

Proof. We prove the result by showing $\gamma \geq |m+1-2y|$.

$$\begin{array}{rcrcl} \gamma \geq |m+1-2y| & \iff & \gamma^2 & \geq & ((m+1)-2y)^2 \\ & \iff & (m+1)^2 - 4my & \geq & (m+1)^2 - 4y(m+1) + 4y^2 \\ & \iff & 4y & \geq & 4y^2. \end{array}$$

Thus, $\gamma \geq |m+1-2y|$ since $0 \leq y \leq 1$ implies $4y \geq 4y^2$. □

As a result of the lemma, we have that $\alpha = \frac{m}{4}\left(\frac{\gamma-(m+1-2y)}{\gamma}\right) \geq 0$. Also,

$$\gamma \geq -(m+1-2y) \quad \Longrightarrow \quad \frac{m}{2} \quad \geq \quad \alpha.$$

Therefore, $0 \leq \alpha \leq \frac{m}{2}$ which proves that ρ_1' and ρ_2' are positive semidefinite.

Note that

$$\mathrm{Tr}_{\mathcal{B}_1}(\rho_1') = \frac{1}{2}\begin{bmatrix} m & & & \sqrt{m} \\ & 1-m & & \\ & & 0 & \\ \sqrt{m} & & & 1 \end{bmatrix} = \mathrm{Tr}_{\mathcal{A}_1}|A\rangle\langle A|,$$

where $\mathrm{Tr}_{\mathcal{A}_1}|A\rangle\langle A|$ is given earlier. Also we have

$$\mathrm{Tr}_{\mathcal{A}_2}(\rho_1') = \begin{bmatrix} \alpha & & & \\ & \frac{1}{2}-\alpha & & \\ & & \frac{\alpha}{m} & \\ & & & \frac{1}{2}-\frac{\alpha}{m} \end{bmatrix} = \mathrm{Tr}_{\mathcal{B}_2}(\rho_2'),$$

so the first two constraints are satisfied, proving (ρ_1', ρ_2') is feasible.

Before evaluating the objective value, we need the following lemma.

Lemma 9.2. *Suppose we are given α as above. Then we have*

$$\sqrt{\alpha(m-2\alpha)} = \frac{m}{\sqrt{2}\gamma}\sqrt{y(1-y)}.$$

Proof. Suppose $\alpha = \frac{m}{4\gamma}(\gamma-(m+1-2y))$. Then we can write the following.

$$\begin{aligned}
\sqrt{\alpha(m-2\alpha)} &= \sqrt{\frac{m}{4\gamma}(\gamma-(m+1-2y))\left(m-\frac{m}{2\gamma}(\gamma-(m+1-2y))\right)} \\
&= \frac{m}{\sqrt{8}\gamma}\sqrt{(\gamma-(m+1-2y)(\gamma+(m+1-2y))} \\
&= \frac{m}{\sqrt{8}\gamma}\sqrt{(m+1)^2-4my-((m+1)-2y)^2} \\
&= \frac{m}{\sqrt{2}\gamma}\sqrt{y(1-y)}.
\end{aligned}$$

□

We have $\Pi_{A,1}$ explicitly as

$$\Pi_{A,1} = \begin{bmatrix} 0 & & & & & & & \\ & 0 & & & & & & \\ & & 1 & & & & & \\ & & & 0 & & & & \\ & & & & y & & & \sqrt{y(1-y)} \\ & & & & & 0 & 0 & \\ & & & & & 0 & 0 & \\ & & & & \sqrt{y(1-y)} & & & 1-y \end{bmatrix},$$

thus we can write the objective value below.

$$\begin{aligned} \langle \Pi_{A,1}, \rho_2 \rangle &= \left(\frac{1}{2} - \alpha\right) + y\left(\frac{\alpha}{m}\right) + (1-y)\left(\frac{1}{2} - \frac{\alpha}{m}\right) + 2\sqrt{y(1-y)}\sqrt{\frac{\alpha}{m}\left(\frac{1}{2} - \frac{\alpha}{m}\right)} \\ &= \frac{1}{2m}\left(2m + 4\alpha y - 2\alpha - 2m\alpha - ym + \sqrt{8}\sqrt{y(1-y)}\sqrt{\alpha(m-2\alpha)}\right) \\ &= \frac{1}{2m}\left(2m + 4\alpha y - 2\alpha - 2m\alpha - ym + \sqrt{8}\sqrt{y(1-y)}\frac{m}{\sqrt{2}\gamma}\sqrt{y(1-y)}\right) \\ &= \frac{1}{2m\gamma}\left(2m\gamma + \alpha(4y\gamma - 2\gamma - 2m\gamma) - ym\gamma + 2ym - 2y^2 m\right) \\ &= \frac{1}{4\gamma}\left(4\gamma - (\gamma - (m+1-2y))(m+1-2y) - 2y\gamma + 4y - 4y^2\right) \\ &= \frac{1}{4\gamma}\left(4\gamma - \gamma m - \gamma + 2y\gamma + (m+1)^2 - 4y(m+1) + 4y^2 - 2y\gamma + 4y - 4y^2\right) \\ &= \frac{1}{4\gamma}(3\gamma - \gamma m + \gamma^2) \\ &= \frac{3 - m + \gamma}{4}. \end{aligned}$$

To show the solution above is optimal for (P_B^1), consider the parametrized pair

$$X_2 = \begin{bmatrix} 0 & & & \\ & 1 & & \\ & & \frac{1}{2}(1+m+\gamma) & \\ & & & \frac{1}{2}(1-m+\gamma) \end{bmatrix},$$

and

$$X_1(t) = \begin{bmatrix} t & & & & -\sqrt{tm(t-1)} \\ & 1 & & & \\ & & \frac{1}{2}(1+m+\gamma) & & \\ -\sqrt{tm(t-1)} & & & & \frac{1}{2}(1-m+\gamma)+tm \end{bmatrix}.$$

To see this is feasible for (D_B^1), we can write $\mathrm{Tr}^*_{\mathcal{B}_1}(X_1(t)) \in \mathcal{A}_0 \otimes \mathcal{B}_1 \otimes \mathcal{A}_2$ as

$$\begin{bmatrix} t & & & & & & f_1(m,t) & & & \\ & 1 & & & & & & & & \\ & & t & & & & & & f_1(m,t) & \\ & & & 1 & & & & & & \\ & & & & f_2(m,y) & & & & & \\ f_1(m,t) & & & & & & f_3(m,\gamma,t) & & & \\ & & & & & & & f_2(m,y) & & \\ & & f_1(m,t) & & & & & & f_3(m,\gamma,t) & \end{bmatrix},$$

where $f_1(m,t) := -\sqrt{tm(t-1)}$, $f_2(m,y) := \frac{1}{2}(1+m+\gamma)$, and $f_3(m,\gamma,t) := \frac{1}{2}(1-m+\gamma)+tm$. (These functions are only defined to help fit the matrix on the page.) We can also write $\mathrm{Tr}^*_{\mathcal{A}_2}(X_2) \in \mathcal{A}_0 \otimes \mathcal{B}_1 \otimes \mathcal{A}_2$ as

$$\begin{bmatrix} 0 & & & & & & & & & \\ & 0 & & & & & & & & \\ & & 1 & & & & & & & \\ & & & 1 & & & & & & \\ & & & & \frac{1}{2}(1+m+\gamma) & & & & & \\ & & & & & \frac{1}{2}(1+m+\gamma) & & & & \\ & & & & & & \frac{1}{2}(1-m+\gamma) & & & \\ & & & & & & & \frac{1}{2}(1-m+\gamma) & & \end{bmatrix}.$$

To prove feasibility of the first constraint, let $X(t) := \mathrm{Tr}^*_{\mathcal{B}_1}(X_1(t)) - \mathrm{Tr}^*_{\mathcal{A}_2}(X_2)$, i.e.,

$$X(t) = \begin{bmatrix} t & & & & & -\sqrt{tm(t-1)} & & \\ & 1 & & & & & & \\ & & t-1 & & & & & -\sqrt{tm(t-1)} \\ & & & 0 & & & & \\ & & & & 0 & & & \\ -\sqrt{tm(t-1)} & & & & & m(t-1) & & \\ & & & & & & m & \\ & & -\sqrt{tm(t-1)} & & & & & m(t-1) \end{bmatrix}.$$

For $t \geq 1$, the above matrix is positive semidefinite by Lemma 8.10.

Since $\mathcal{B}_2 = \mathcal{A}_2 = \mathbb{C}^2$, we have $\mathrm{Tr}^*_{\mathcal{A}_2}(X_2) = \mathrm{Tr}^*_{\mathcal{B}_2}(X_2) \in \mathcal{A}_0 \otimes \mathcal{B}_1 \otimes \mathcal{B}_2$. To verify the second constraint, let $X' := \mathrm{Tr}^*_{\mathcal{B}_2}(X_2) - \Pi_{A,1}$, i.e.,

$$X' = \begin{bmatrix} 0 & & & & & & & \\ & 0 & & & & & & \\ & & 0 & & & & & \\ & & & 1 & & & & \\ & & & & s_1 - y & & & -\sqrt{y(y-1)} \\ & & & & & s_1 & & \\ & & & & & & s_2 & \\ & & & & -\sqrt{y(y-1)} & & & s_2 - (y-1) \end{bmatrix},$$

where $s_1 := \frac{1}{2}(1+m+\gamma)$ and $s_2 := \frac{1}{2}(1-m+\gamma)$. To show that X' is positive semidefinite, we first show it has nonnegative diagonal elements by showing the following.

1. $s_1 - y = \frac{1}{2}(1+m+\gamma) - y \geq 0$, and
2. $s_2 - (y+1) = \frac{1}{2}(1-m+\gamma) - (y-1) \geq 0$.

We see that $\frac{1}{2}(1-m+\gamma)-(y-1) \geq 0$ since $m, y \leq 1$ and $\gamma \geq 0$. Also, $\frac{1}{2}(1+m+\gamma)-y \geq 0$ holds by Lemma 9.1. We need to verify that

$$\det\left(\begin{bmatrix} 1+m+\gamma-y & -2\sqrt{y(y-1)} \\ -2\sqrt{y(y-1)} & 1-m+\gamma-y+1 \end{bmatrix}\right) \geq 0.$$

The determinant above is nonnegative since $(1+m+\gamma-y)(1-m+\gamma-y+1)-4y(y-1)$ can be written as $(1+m-y)(2-m-y)+(2-m-y)\gamma+(1+m-y)\gamma+(m+1)^2+4y(1-m-y)$, which is a sum of nonnegative terms. Thus, by Lemma 8.10, we have that $X' \succeq 0$, verifying the second constraint. This solution has objective value

$$\begin{aligned}
\lim_{t\to\infty} \langle X_1(t), \operatorname{Tr}_{\mathcal{A}_1} |A\rangle\langle A| \rangle &= \lim_{t\to\infty} \frac{1}{2}\left[(1-m) + \frac{1}{2}(1-m+\gamma) + 2tm - 2\sqrt{tm(t-1)}\sqrt{m}\right] \\
&= \frac{1}{2}\left[(1-m) + \frac{1}{2}(1-m+\gamma)\right] + m \cdot \lim_{t\to\infty}\left[t - \sqrt{t(t-1)}\right] \\
&= \frac{1}{4}(3-m+\gamma)\,.
\end{aligned}$$

This matches the primal objective value so the primal and dual feasible solutions given above are optimal. Now we prove $P^*_{B,1} < 1$ by showing

$$\begin{aligned}
& & P^*_{B,1} &< 1 \\
\iff & & \gamma^2 &< (1+m)^2 \\
\iff & (1+m)^2 - 4my &< (1+m)^2 \\
\iff & 0 &< my.
\end{aligned}$$

Since $my > 0$, we have that $P^*_{B,1} < 1$, as desired.

9.1.1 Optimal Cheating Strategy for Forcing Outcome 0

We begin by finding the optimal cheating strategy for Bob forcing outcome 0 with probability $P^*_{B,0} = 1$. Let C be an ancilla space used for finding purifications of the optimal primal solutions (which are density matrices). We can purify ρ_1 as $|\psi_1^{pure}\rangle$ and ρ_2 as $|\psi_2^{pure}\rangle$ below.

$$|\psi_1^{pure}\rangle = \sqrt{\frac{y}{2}}\,|001\rangle\,|0\rangle_C + \sqrt{\frac{m}{2}}\,|010\rangle\,|1\rangle_C + \sqrt{\frac{1}{2}}\,|111\rangle\,|1\rangle_C + \sqrt{\frac{1-y-m}{2}}\,|011\rangle\,|2\rangle_C$$

and

$$|\psi_2^{pure}\rangle = \sqrt{\frac{y}{2}}\,|000\rangle\,|0\rangle_C + \sqrt{\frac{1-y}{2}}\,|011\rangle\,|0\rangle_C + \sqrt{\frac{1}{2}}\,|110\rangle\,|1\rangle_C\,.$$

To find U_1, we append the state $|0\rangle_C$ to Alice's beginning state $|A\rangle$ as follows.

$$\begin{aligned} |A\rangle &= \sqrt{\frac{m}{2}}\,|000\rangle\,|0\rangle_C + \sqrt{\frac{1-m}{2}}\,|011\rangle\,|0\rangle_C + \sqrt{\frac{1}{2}}\,|101\rangle\,|0\rangle_C \\ &= \left(\sqrt{\frac{m}{2}}\,|00\rangle + \sqrt{\frac{1}{2}}\,|11\rangle\right) \otimes |00\rangle_{\mathcal{A}_1\otimes C} + \sqrt{\frac{1-m}{2}}\,|01\rangle \otimes |10\rangle_{\mathcal{A}_1\otimes C}\,. \end{aligned}$$

We also have

$$\begin{aligned} |\psi_1^{pure}\rangle &= \sqrt{\frac{y}{2}}\,|001\rangle\,|0\rangle_C + \sqrt{\frac{m}{2}}\,|010\rangle\,|1\rangle_C + \sqrt{\frac{1}{2}}\,|111\rangle\,|1\rangle_C + \sqrt{\frac{1-y-m}{2}}\,|011\rangle\,|2\rangle_C \\ &= \sqrt{\frac{y}{2}}\,|01\rangle \otimes |00\rangle_{\mathcal{B}_1\otimes C} + \left(\sqrt{\frac{m}{2}}\,|00\rangle + \sqrt{\frac{1}{2}}\,|11\rangle\right) \otimes |11\rangle_{\mathcal{B}_1\otimes C} \\ &\quad + \sqrt{\frac{1-y-m}{2}}\,|01\rangle\,|12\rangle_{\mathcal{B}_1\otimes C}\,. \end{aligned}$$

Thus, we define the unitary $U_1 : \mathcal{A}_1 \otimes C \to \mathcal{B}_1 \otimes C$ to act on the basis states as follows.

$$U_1 : \begin{array}{c} |00\rangle \to |11\rangle\,, \\ |10\rangle \to \sqrt{\dfrac{y}{1-m}}\,|00\rangle + \sqrt{\dfrac{1-y-m}{1-m}}\,|12\rangle\,. \end{array}$$

To find U_2, note that

$$\begin{aligned} |\psi_1^{pure}\rangle &= \sqrt{\frac{y}{2}}\,|001\rangle\,|0\rangle_C + \sqrt{\frac{m}{2}}\,|010\rangle\,|1\rangle_C + \sqrt{\frac{1}{2}}\,|111\rangle\,|1\rangle_C + \sqrt{\frac{1-y-m}{2}}\,|011\rangle\,|2\rangle_C \\ &= \sqrt{\frac{y}{2}}\,|00\rangle \otimes |10\rangle_{\mathcal{A}_2\otimes C} + \sqrt{\frac{m}{2}}\,|01\rangle \otimes |01\rangle_{\mathcal{A}_2\otimes C} + \sqrt{\frac{1}{2}}\,|11\rangle \otimes |11\rangle_{\mathcal{A}_2\otimes C} \\ &\quad + \sqrt{\frac{1-y-m}{2}}\,|01\rangle \otimes |12\rangle_{\mathcal{A}_2\otimes C} \end{aligned}$$

and

$$\begin{aligned} |\psi_2^{pure}\rangle &= \sqrt{\frac{y}{2}}\,|000\rangle\,|0\rangle_C + \sqrt{\frac{1-y}{2}}\,|011\rangle\,|0\rangle_C + \sqrt{\frac{1}{2}}\,|110\rangle\,|1\rangle_C \\ &= \sqrt{\frac{y}{2}}\,|00\rangle \otimes |00\rangle_{\mathcal{B}_2\otimes C} + \sqrt{\frac{1-y}{2}}\,|01\rangle \otimes |10\rangle_{\mathcal{B}_2\otimes C} + \sqrt{\frac{1}{2}}\,|11\rangle \otimes |01\rangle_{\mathcal{B}_2\otimes C}\,. \end{aligned}$$

Thus, we define the unitary $U_2 : \mathcal{A}_2 \otimes C \to \mathcal{B}_2 \otimes C$ to act on the basis states as follows.

$$U_2 : \begin{array}{c} |10\rangle \to |00\rangle\,, \\ |11\rangle \to |01\rangle\,, \\ \sqrt{\frac{m}{1-y}}\,|01\rangle + \sqrt{\frac{1-y-m}{1-y}}\,|12\rangle \to |10\rangle\,. \end{array}$$

Therefore, Bob's optimal cheating strategy is to create the ancilla space C, performing U_1 before sending $\mathcal{B}_1$ to Alice, and performing U_2 before sending $\mathcal{B}_2$ to Alice. Doing this, the optimal solutions of the SDP are Alice's density matrices from which she measures and accepts 0 with probability $P^*_{B,0} = 1$.

9.1.2 Optimal Cheating Strategy for Forcing Outcome 1

Now we find the optimal cheating strategy for Bob forcing outcome 1 with probability $P^*_{B,1}$. Again, let C be an ancilla space used for finding purifications of the optimal primal solutions. We can purify ρ'_1 and ρ'_2 as

$$\begin{aligned} |\psi_1^{pure}\rangle \;=\; & \sqrt{\alpha}\,|000\rangle\,|00\rangle_C + \sqrt{\frac{\alpha}{m}}\,|101\rangle\,|00\rangle_C \\ & + \sqrt{\frac{m}{2} - \alpha}\,|010\rangle\,|01\rangle_C + \sqrt{\frac{1}{2} - \frac{\alpha}{m}}\,|111\rangle\,|01\rangle_C \\ & + \sqrt{\frac{1-m}{2}}\,|011\rangle\,|11\rangle_C \end{aligned}$$

and

$$\begin{aligned} |\psi_2^{pure}\rangle \;=\; & \sqrt{\frac{\alpha}{2}}\,|000\rangle\,|00\rangle_C + \sqrt{\frac{\alpha}{2}}\,|001\rangle\,|01\rangle_C \\ & + \sqrt{\frac{1}{2} - \alpha}\,|010\rangle\,|10\rangle_C + \sqrt{\frac{\alpha}{m}}\,|100\rangle\,|11\rangle_C \\ & + \sqrt{\frac{1}{2} - \frac{\alpha}{m}}\,|111\rangle\,|11\rangle_C\,. \end{aligned}$$

To find V_1, we append $|00\rangle_C$ to the end of $|A\rangle$ and write it as

$$\begin{aligned}|A\rangle &= \sqrt{\frac{m}{2}}\,|000\rangle\,|00\rangle_C + \sqrt{\frac{1-m}{2}}\,|011\rangle\,|00\rangle_C + \sqrt{\frac{1}{2}}\,|101\rangle\,|00\rangle_C \\ &= \left(\sqrt{\frac{m}{2}}\,|00\rangle + \sqrt{\frac{1}{2}}\,|11\rangle\right)\otimes|0\rangle_{\mathcal{A}_1}\,|00\rangle_C + \sqrt{\frac{1-m}{2}}\,|01\rangle\otimes|1\rangle_{\mathcal{A}_1}\,|00\rangle_C\,.\end{aligned}$$

We can write $|\psi_1^{pure}\rangle$ as

$$\begin{aligned}|\psi_1^{pure}\rangle &= \left(\sqrt{\alpha}\,|00\rangle + \sqrt{\frac{\alpha}{m}}\,|11\rangle\right)\otimes|0\rangle_{\mathcal{B}_1}\,|00\rangle_C \\ &\quad + \left(\sqrt{\frac{m}{2}-\alpha}\,|00\rangle + \sqrt{\frac{1}{2}-\frac{\alpha}{m}}\,|11\rangle\right)\otimes|1\rangle_{\mathcal{B}_1}\,|01\rangle_C \\ &\quad + \sqrt{\frac{1}{2}-\frac{m}{2}}\,|01\rangle\otimes|1\rangle_{\mathcal{B}_1}\,|11\rangle_C\,.\end{aligned}$$

Therefore, we define the unitary $V_1 : \mathcal{A}_1 \otimes C \to \mathcal{B}_1 \otimes C$ to act on the basis states as follows.

$$\begin{aligned}V_1 :\ & |0\rangle\,|00\rangle \to \sqrt{\frac{2\alpha}{m}}\,|0\rangle\,|00\rangle + \sqrt{\frac{m-2\alpha}{m}}\,|1\rangle\,|01\rangle\,, \\ & |1\rangle\,|00\rangle \to |1\rangle\,|11\rangle\,.\end{aligned}$$

To find V_2, note that

$$\begin{aligned}|\psi_1^{pure}\rangle &= \sqrt{\alpha}\,|00\rangle\otimes|0\rangle_{\mathcal{A}_2}\,|00\rangle_C + \sqrt{\frac{\alpha}{m}}\,|10\rangle\otimes|1\rangle_{\mathcal{A}_2}\,|00\rangle_C \\ &\quad + \sqrt{\frac{m}{2}-\alpha}\,|01\rangle\otimes|0\rangle_{\mathcal{A}_2}\,|01\rangle_C + \sqrt{\frac{1}{2}-\frac{\alpha}{m}}\,|11\rangle\otimes|1\rangle_{\mathcal{A}_2}\,|01\rangle_C \\ &\quad + \sqrt{\frac{1}{2}-\frac{m}{2}}\,|01\rangle\otimes|1\rangle_{\mathcal{A}_2}\,|11\rangle_C\end{aligned}$$

and

$$\begin{aligned}|\psi_2^{pure}\rangle &= \sqrt{\frac{\alpha}{2}}\,|00\rangle\otimes|0\rangle_{\mathcal{B}_2}\,|00\rangle_C + \sqrt{\frac{\alpha}{2}}\,|00\rangle\otimes|1\rangle_{\mathcal{B}_2}\,|01\rangle_C \\ &\quad \sqrt{\frac{1}{2}-\alpha}\,|01\rangle\otimes|0\rangle_{\mathcal{B}_2}\,|10\rangle_C + \sqrt{\frac{\alpha}{m}}\,|10\rangle\otimes|0\rangle_{\mathcal{B}_2}\,|11\rangle_C \\ &\quad \sqrt{\frac{1}{2}-\frac{\alpha}{m}}\,|11\rangle\otimes|1\rangle_{\mathcal{B}_2}\,|11\rangle_C\,.\end{aligned}$$

Thus, we define the unitary $V_2 : \mathcal{A}_2 \otimes C \to \mathcal{B}_2 \otimes C$ to act on the basis states as follows.

$$V_2 : \begin{array}{c} |0\rangle\,|00\rangle \to \sqrt{\frac{1}{2}}\,|0\rangle\,|00\rangle + \sqrt{\frac{1}{2}}\,|1\rangle\,|01\rangle\,, \\ |1\rangle\,|00\rangle \to |0\rangle\,|11\rangle\,, \\ |1\rangle\,|01\rangle \to |1\rangle\,|11\rangle\,, \\ \sqrt{\frac{m-2\alpha}{1-2\alpha}}\,|0\rangle\,|01\rangle + \sqrt{\frac{1-m}{1-2\alpha}}\,|1\rangle\,|11\rangle \to |0\rangle\,|10\rangle\,. \end{array}$$

Bob's optimal cheating strategy for forcing outcome 1 is the same as his strategy for forcing outcome 0 except he applies V_1 instead of U_1 and V_2 instead of U_2. With this strategy Alice accepts outcome 1 with probability $P^*_{B,1} < 1$.

Chapter 10

Computational Results

10.1 The Search for 3 and 4-Round Protocols with Low Bias

In this chapter, we discuss how we approach finding BCCF protocols with low bias. Recall that the protocols in [2, 36, 18] achieve a bias of $\varepsilon = 1/4$. Also, recall from Section 8.3 that the bias is at least $1/4$ in any 3-round BCCF protocol [2]. This implies if we want to find protocols with a lower bias, we need more than 3 rounds. The first step is to search for 4-round BCCF protocols with low values for the bias.

We set up our search using the following pseudocode. Since there are an infinite number of starting states, we test all starting states to within a certain accuracy, which we describe later.

Pseudocode 10.1.

1. *Choose accuracy parameter. Initialize* $States_{best}$ *and* $Bias_{min}$ *being the best found starting states and bias, respectively.*
2. *For all starting states* $|\psi_0\rangle$, $|\psi_1\rangle$, $|\phi_0\rangle$, *and* $|\phi_1\rangle$, *within the given accuracy, such that* $|\psi_0\rangle \perp |\psi_1\rangle$, *and* $|\phi_0\rangle \perp |\phi_1\rangle$*:*
   (a) *Solve for* $P^*_{A,0}$, $P^*_{A,1}$, $P^*_{B,0}$, *and* $P^*_{B,1}$. *Set* $Bias = \max\{P^*_{A,0}, P^*_{A,1}, P^*_{B,0}, P^*_{B,1}\} - 1/2$.
   (b) *If* $Bias < Bias_{min}$ *then update* $States_{best}$ *and* $Bias_{min}$.

3. end For loop.

We later describe how we define the accuracy parameter and the way we choose the starting states, but first we discuss an equivalence on BCCF protocols which reduces redundant computation.

Theorem 10.2. *Suppose we have a BCCF protocol defined by the starting states $|\psi_0\rangle$, $|\psi_1\rangle$, $|\phi_0\rangle$, and $|\phi_1\rangle$. We define a new protocol by defining a set of new starting states, where we apply an arbitrary unitary of the form $U_{\mathcal{A}_j} \in \mathcal{L}(\mathcal{A}_j)$ to both of Alice's starting states*[1] *and correspondingly changing the projection operators $\Pi_{B,0}$ and $\Pi_{B,1}$. Then the optimal cheating probabilities of each outcome for each party are the same in both the original and new protocol.*

Proof. Recall the SDP for cheating Bob (corresponding to the original protocol) is as follows.

$$\begin{array}{lrcl} (P_B^x) \quad P_{B,x}^* = \sup & \left\langle \mathrm{Tr}^*_{\mathcal{A}_n}(\Pi_{A,x}), \rho_{n-1} \right\rangle & & \\ & \mathrm{Tr}_{\mathcal{B}_1}(\rho_1) & = & \mathrm{Tr}_{\mathcal{A}_1}|A\rangle\langle A|, \\ \text{subject to} & \mathrm{Tr}_{\mathcal{B}_i}(\rho_i) & = & \mathrm{Tr}_{\mathcal{A}_i}(\rho_{i-1}),\ \forall i \in \{2,\ldots,n-1\}, \\ & \rho_i & \in & \Sigma_+^{\mathcal{A}_0\otimes\mathcal{B}_1\otimes\cdots\otimes\mathcal{B}_i\otimes\mathcal{A}_{i+1}\otimes\cdots\otimes\mathcal{A}_n},\ \forall i \in \{1,\ldots,n-1\}. \end{array}$$

Let $(\tau_1,\ldots,\tau_{n-1})$ be a feasible solution of (P_B^x). Define a new protocol by applying a unitary of the form $U_{\mathcal{A}_j} \in \mathcal{L}(\mathcal{A}_j)$, for some $j \in \{1,\ldots,n\}$, to both of Alice's starting states and correspondingly changing the projectors $\Pi_{B,0}$ and $\Pi_{B,1}$. Note that Alice's projectors do not change. Let $(\widetilde{P}_B^x)$ be the cheating SDP for Bob corresponding to the new protocol. We see that if $j = 1$ or $j = n$, then $(\tau_1,\ldots,\tau_{n-1})$ is a feasible solution for $(\widetilde{P}_B^x)$ as well. Suppose $j \in \{2,\ldots,n-1\}$. Define

$$(\tau_1',\ldots,\tau_{n-1}') = (U_{\mathcal{A}_j}\tau_1 U_{\mathcal{A}_j}^*,\ldots,U_{\mathcal{A}_j}\tau_{j-1}U_{\mathcal{A}_j}^*,\tau_j,\ldots,\tau_{n-1}).$$

We show this is a feasible solution for $(\widetilde{P}_B^x)$. Note that $U_{\mathcal{A}_j}(\cdot)U_{\mathcal{A}_j}^* \in \mathrm{Aut}\,(\Sigma_+^{\mathcal{A}})$, thus $\tau_j' \succeq 0$, for all $j \in \{1,\ldots,n-1\}$. Also, unitary operations of the form $U_{\mathcal{A}_j}$ commute with the partial traces in the SDP's. This is either because they act on different tensor factors

[1] When we apply a unitary to a tensor factor, we assume it acts as the identity on the rest of the tensor factors.

or because the partial trace is invariant under changes of basis (Lemma 3.2). Thus, for $i \in \{2, \ldots, j-1\}$, we have

$$\begin{aligned} \mathrm{Tr}_{\mathcal{B}_i}(\tau_i') &= \mathrm{Tr}_{\mathcal{B}_i}(U_{\mathcal{A}_j}\tau_i U_{\mathcal{A}_j}^*) \\ &= U_{\mathcal{A}_j}(\mathrm{Tr}_{\mathcal{B}_i}(\tau_i))U_{\mathcal{A}_j}^* \\ &= U_{\mathcal{A}_j}(\mathrm{Tr}_{\mathcal{A}_i}(\tau_{i-1}))U_{\mathcal{A}_j}^* \\ &= \mathrm{Tr}_{\mathcal{A}_i}(U_{\mathcal{A}_j}\tau_{i-1}U_{\mathcal{A}_j}^*) \\ &= \mathrm{Tr}_{\mathcal{A}_i}(\tau_{i-1}'). \end{aligned}$$

For $i = j$, we have

$$\begin{aligned} \mathrm{Tr}_{\mathcal{B}_j}(\tau_j') &= \mathrm{Tr}_{\mathcal{B}_j}(\tau_j) \\ &= \mathrm{Tr}_{\mathcal{A}_j}(\tau_{j-1}) \\ &= \mathrm{Tr}_{\mathcal{A}_j}(U_{\mathcal{A}_j}\tau_{j-1}U_{\mathcal{A}_j}^*) \\ &= \mathrm{Tr}_{\mathcal{A}_j}(\tau_{j-1}'). \end{aligned}$$

The rest of the constraints can be similarly verified. Since $j \leq n-1$, we know that $\tau_{n-1}' = \tau_{n-1}$, thus,

$$\left\langle \mathrm{Tr}_{\mathcal{A}_n}^*(\Pi_{A,x}), \tau_{n-1}' \right\rangle = \left\langle \mathrm{Tr}_{\mathcal{A}_n}^*(\Pi_{A,x}), \tau_{n-1} \right\rangle.$$

Therefore, there is an objective value preserving bijection between feasible solutions of (P_B^x) and the feasible solutions of $(\widetilde{P}_B^x)$. This proves that the optimal values of the two SDP's are the same. Therefore, Bob's optimal cheating probabilities are the same in both the original and new protocol. By a similar argument, we can show Alice's cheating probabilities are also preserved. □

We can also apply a unitary of the form $U_{\mathcal{B}_j} \in \mathcal{L}(\mathcal{B}_j)$, for some $j \in \{1, \ldots, n-1\}$, to Bob's starting states and accordingly change Alice's projectors to define a new protocol. By a similar argument, this also preserves the optimal cheating probabilities.

This leads to the following definition.

Definition 10.3 (Equivalent Protocols)**.** *Suppose we are given two BCCF protocols defined by the starting states* $\{|\psi_a\rangle, |\phi_b\rangle : a, b, \in \{0,1\}\}$ *and* $\{|\psi_a'\rangle, |\phi_b'\rangle : a, b, \in \{0,1\}\}$*. We*

say they are equivalent *if there exist unitary operations of the form* $U_A := \bigotimes_i U_{\mathcal{A}_i}$ *and* $U_B := \bigotimes_j U_{\mathcal{B}_j}$ *such that*

$$|\psi'_a\rangle = U_A |\psi_a\rangle\,, \text{ for } a \in \{0,1\}, \quad \text{and} \quad |\phi'_b\rangle = U_B |\phi_b\rangle\,, \text{ for } b \in \{0,1\}.$$

From Theorem 10.2, we see that equivalent protocols have the same cheating probabilities and thus the same bias. Therefore, we need only test starting states from different equivalence classes. We achieve this through the following lemma.

Lemma 10.4. *Suppose we have a* 3 *or* 4*-round BCCF protocol with* $\mathcal{A}_1 = \mathcal{A}_2 = \mathbb{C}^d$*. We may assume*

$$|\psi_0\rangle = \sum_{i=1}^{d} \lambda_i |i\rangle |i\rangle\,,$$

where $\lambda_i \geq 0$ *satisfy* $\sum_{i=1}^{d} \lambda_i^2 = 1$*, called the Schmidt coefficients. Suppose we have a* 4*-round protocol with* $\mathcal{B}_1 = \mathcal{B}_2 = \mathbb{C}^{d'}$*. Then we may also assume*

$$|\phi_0\rangle = \sum_{i=1}^{d'} \mu_i |i\rangle |i\rangle\,,$$

where $\mu_i \geq 0$ *satisfy* $\sum_{i=1}^{d'} \mu_i^2 = 1$*.*

Proof. Consider the Schmidt decomposition[2] for Alice's starting states below.

$$|\psi_0\rangle = \sum_i \lambda_i |i_{\mathcal{A}_1}\rangle |i_{\mathcal{A}_2}\rangle \text{ and } |\psi_1\rangle = \sum_i \lambda'_i \left|i'_{\mathcal{A}_1}\right\rangle \left|i'_{\mathcal{A}_2}\right\rangle,$$

where $\{|i_{\mathcal{A}_1}\rangle\}$, $\{|i_{\mathcal{A}_2}\rangle\}$, $\{\left|i'_{\mathcal{A}_1}\right\rangle\}$, and$\{\left|i'_{\mathcal{A}_2}\right\rangle\}$ are orthonormal sets in their respective spaces. Using the equivalence above, we can apply the change of basis $|i_{\mathcal{A}_1}\rangle |i_{\mathcal{A}_2}\rangle \xrightarrow{U\otimes V} |i\rangle |i\rangle$ to both states to find an equivalent protocol. Alice now has the starting states $|\psi'_0\rangle, |\psi'_1\rangle$ with $|\psi'_0\rangle = \sum_{i=1}^{d} \lambda_i |i\rangle |i\rangle$. Similarly, we can do this for Bob's starting states in a 4-round protocol. □

Note the Schmidt decomposition does not hold true for tripartite states. I.e., for an arbitrary state $|\psi\rangle \in \mathcal{A}_1 \otimes \mathcal{A}_2 \otimes \mathcal{A}_3$, we can not always find $\lambda_i \geq 0$, $\sum_i \lambda_i^2 = 1$ such

[2]See, for example, Theorem 2.7 in [31]

that $|\psi\rangle = \sum_i \lambda_i |i_{\mathcal{A}_1}\rangle |i_{\mathcal{A}_2}\rangle |i_{\mathcal{A}_3}\rangle$, where $\{|i_{\mathcal{A}_1}\rangle\}$, $\{|i_{\mathcal{A}_2}\rangle\}$, and $\{|i_{\mathcal{A}_3}\rangle\}$ are orthonormal sets. Therefore, such a lemma does not exist for m-round BCCF protocols where $m \geq 5$.

A consequence of our equivalence is that we may assume $|\phi_0\rangle = |0\rangle$ and $|\phi_1\rangle = |1\rangle$ for 3-round BCCF protocols. Unfortunately, in 4 rounds we cannot assume much about the form of $|\psi_1\rangle$ and $|\phi_1\rangle$ except, by orthogonality, we can easily write down a basis for the orthogonal compliment of $\sum_i \lambda_i |i\rangle |i\rangle$.

We now describe the accuracy of our search for protocols with low bias. Suppose all of the messages are qubits, for ease of computation. To choose a starting state $|\psi_0\rangle$, we need only choose one Schmidt coefficient. This is because $\dim(\mathcal{A}_1) = \dim(\mathcal{A}_2) = 2$, so one Schmidt coefficient determines the other. For the first Schmidt coefficient, we loop through every value between 0 and 0.5 by increments of 10^{-q}. The reason we do not need any value larger than 0.5 is because then we could have chosen a different change of basis in the proof of Lemma 10.4, namely $|i_{\mathcal{A}_1}\rangle |i_{\mathcal{A}_2}\rangle \stackrel{U' \otimes V'}{\longrightarrow} |1 \oplus i\rangle |1 \oplus i\rangle$. For every choice of $|\psi_0\rangle$, we test different choices for $|\psi_1\rangle$. After fixing $|\psi_0\rangle$, we define an orthonormal basis for $|\psi_0\rangle^{\perp}$ as $\{|m_1\rangle, |m_1\rangle, |m_3\rangle\}$, where

$$|m_1\rangle := |01\rangle, \quad |m_2\rangle := |10\rangle, \quad \text{and} \quad |m_3\rangle := \lambda_1 |00\rangle - \lambda_0 |11\rangle.$$

We loop to find all the values of β_1 and β_2 by increments of 10^{-p}. For the values that satisfy $\beta_1^2 + \beta_2^2 \leq 1$, we define $|\psi_1\rangle$ to be

$$|\psi_1\rangle := \beta_1 |m_1\rangle + \beta_2 |m_2\rangle \pm \sqrt{1 - \beta_1^2 - \beta_2^2}\, |m_3\rangle.$$

We note here that we only test real values for β_1 and β_2 to keep the computation manageable. This is how we enumerate all the possible starting states for Alice. We do the same for Bob in the case of 4 rounds, but as mentioned above, in 3 rounds we can assume $|\phi_b\rangle = |b\rangle$.

10.1.1 Computational Platform

Programs were run on Matlab, Version 7, on a Sun V880 with 8 UltraSPARC III CPUs at 1.2 GHz, and 32 GB of memory, running Solaris 8. Semidefinite programs were solved using SeDuMi 1.1, a program for solving semidefinite optimization problems in Matlab.

10.2 Three-Round BCCF Search Results

We test the search for the case of 3-round BCCF protocols where all the messages are qubits. The best 3-round BCCF protocol where the first message is a qubit achieves a cheating probability of $\varepsilon = \frac{1}{4}(1+\sqrt{5}) \approx 0.8090$ with the starting states

$$|\psi_0\rangle := |11\rangle \text{ and } |\psi_1\rangle := \alpha_1 |01\rangle + \alpha_2 |10\rangle ,$$

where $\alpha_2 = \frac{1}{2}(\sqrt{5}-1) \approx 0.6180$ and $\alpha_1 = \sqrt{1-\alpha_2^2} \approx 0.7862$, see Question 2 in [28].

When we set our accuracy parameters to be $q := 1$ and $p := 1$, our search finds a protocol with cheating probability 0.8162 with the starting states

$$|\psi_0\rangle = 0.6325 |00\rangle + 0.7746 |11\rangle \text{ and } |\psi_1\rangle = |01\rangle .$$

When we set our accuracy parameters to be $q := 1$ and $p := 2$, our search finds a protocol with cheating probability 0.8100 with the starting states

$$|\psi_0\rangle = |11\rangle \text{ and } |\psi_1\rangle = -0.0849 |00\rangle + 0.7800 |01\rangle - 0.6200 |10\rangle .$$

When we set our accuracy parameters to be $q := 1$ and $p := 3$, our search finds a protocol with cheating probability 0.8091 with the starting states

$$|\psi_0\rangle = |11\rangle \text{ and } |\psi_1\rangle = -0.0167 |00\rangle + 0.7860 |01\rangle + 0.6180 |10\rangle .$$

We notice the search algorithm agrees with the known optimal states to within a certain level of accuracy. We now expand our search to find the optimal 4-round BCCF protocol.

10.3 Four-Round BCCF Search Results

In 4-round BCCF protocols we have to iterate through all possible starting states for Bob as well as Alice. The number of protocols is the square of the number of 3-round protocols since we have to test every choice of Alice's starting states with every choice of Bob's starting states. Furthermore, we have the added difficulty that we do not have a nice expression for the optimal cheating probabilities, thus we must solve for each of the four probabilities using the SDP formulations.

If we set the accuracy at $q := 1$ and $p := 1$ in the search, we have 3920400 protocols which means we solve 15681600 SDP's. Since we are only interested in finding the protocols which have a low bias, we can weed out certain bad starting states. Although we do not know the optimal cheating strategies, we do know some cheating strategies, namely the extensions of the optimal 3-round cheating strategies to 4 rounds. These extensions are basically the 3-round strategies where Alice and Bob ignore the last message. We check the (non-optimal) cheating probabilities using these strategies and, if any are larger than the lowest recorded cheating probability, move on and test the next choice of starting states. This skips the step of solving the SDP's which is not of any use in this case.

Setting $p := 1$ and $q := 1$, the best protocol has a cheating probability of 0.8128 with the starting states

$$\begin{aligned}
|\psi_0\rangle &= 0.6325\,|00\rangle + 0.7746\,|11\rangle\,, \\
|\psi_1\rangle &= |01\rangle\,, \\
|\phi_0\rangle &= |11\rangle\,, \text{ and} \\
|\phi_1\rangle &= -0.9798\,|00\rangle - 0.2000\,|10\rangle\,.
\end{aligned}$$

Setting $p := 1$ and $q := 2$, the best protocol has a cheating probability of 0.8099 with the starting states

$$\begin{aligned}
|\psi_0\rangle &= 0.7000\,|00\rangle + 0.7141\,|11\rangle\,, \\
|\psi_1\rangle &= -0.7141\,|00\rangle + 0.7000\,|11\rangle\,, \\
|\phi_0\rangle &= 0.6164\,|00\rangle + 0.7874\,|11\rangle\,, \text{ and} \\
|\phi_1\rangle &= |01\rangle\,.
\end{aligned}$$

We conjecture that the optimal bias achievable in 4 rounds is no better than what can be achieved in 3 rounds. This seems to be the case computationally.

Chapter 11

Conclusion and Future Research

We examine coin-flipping protocols based on quantum bit-commitment. For this class of protocols, we formulate cheating strategies as semidefinite programs and analyze the cheating strategies therein. We show that we can solve for the cheating probabilities by solving the corresponding dual SDP's. Within this class of protocols, we show a family of 4-round protocols with asymmetric cheating probabilities. Also, we show computational results which suggest that adding more rounds to BCCF protocols may not help in finding a smaller bias.

Future research includes solving the open problem of finding the optimal bias for strong coin-flipping. A special case of this problem is to find the optimal bias for BCCF protocols. Solving this special case could solve the general case if there exists a BCCF protocol with bias matching Kitaev's lower bound. If no such protocol exists, this could give insight for the analysis of the general case. An important step in analyzing BCCF protocols is understanding how the bias depends on the number of rounds. Also, understanding how to 'symmetrize' a BCCF protocol so that all four cheating probabilities are equal could characterize optimal protocols. The optimal 3-round BCCF protocol where the first message is a qubit has this property, as does the optimal 3-round qutrit protocol. We believe that all four cheating probabilities being equal is a necessary condition for an optimal strong coin-flipping protocol. Finding continuous methods for changing the four cheating probabilities is key to characterizing such a property. It is also important to have a better understanding of the duals of the cheating SDP's. If possible, having a quantum mechanical

interpretation of the duals could lead to an alternative method of examining coin-flipping protocols. This has been done for virtually every application of linear programming. A basic example is König's theorem in combinatorics. König's theorem states that the size of a maximum matching is equal to the size of a minimum vertex cover in any bipartite graph. A proof of this involves a linear programming formulation for finding a maximum matching and examining its dual program. (A linear program is a special case of a semidefinite program.) An open question is to determine whether a quantum mechanical interpretation of the duals exists and to see what this reveals about coin-flipping.

Bibliography

[1] D. Aharonov, A. Ta-Shma, U. Vazirani, and A. Yao. Quantum bit escrow. In *Proceedings of the 32nd annual ACM symposium on Theory of computing*, pages 705–714, 2000.

[2] A. Ambainis. A new protocol and lower bounds for quantum coin flipping. In *Proceedings of the 33rd Annual ACM Symposium on Theory of Computing*, pages 134 – 142, 2001.

[3] A. Ambainis. Lower bound for a class of weak quantum coin flipping protocols. quant-ph/0204063, 2002.

[4] A. Ambainis. Polynomial degree vs. quantum query complexity. *Journal of Computer and System Sciences*, 72: 220–238, 2006.

[5] A. Ambainis, H. Buhrman, Y. Dodis, and H. Röhrig. Multiparty quantum coin flipping. In *Proceedings of the 19th IEEE Annual Conference on Computational Complexity*, pages 250 – 259, 2004.

[6] T. Ando. Concavity of certain maps on positive definite matrices and applications to Hadamard products. *Linear Algebra and its Applications*, 26: 203 – 241, 1979.

[7] H. Barnum, M. Saks, and M. Szegedy. Quantum query complexity and semidefinite programming. In *Proceedings of the IEEE Conference on Computational Complexity*, 2003.

[8] C. 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.

[9] S. Boyd and L. Vandenberghe. Semidefinite programming. *SIAM Review*, 38(1): 49 – 95, 1996.

[10] R. Cleve, W. Slofstra, F. Unger, and S. Upadhyay. Perfect parallel repetition theorem for quantum XOR proof systems. To appear in *Proceedings of the 22nd IEEE Conference on Computational Complexity*, 2007.

[11] D. Dieks. Communication by EPR devices. *Physics Letters A*, 92(6): 271–272, 1982.

[12] C. Fuchs and J. van de Graaf. Cryptographic distinguishability measures for quantum mechanical states. *IEEE Transactions on Information Theory*, 45: 1216 – 1227, 1999.

[13] M. Goemans and D. Williamson. Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming. *Journal of the ACM*, 42: 1115 – 1145, 1995.

[14] G. Gutoski and J. Watrous. Toward a general theory of quantum games. To appear in *Proceedings of the 39th ACM Symposium on Theory of Computing*, 2007.

[15] R. Horn and C. Johnson. *Topics in Matrix Analysis*. Cambridge University Press, 1991.

[16] D. Hyland and E. Collins Jr. Block Kronecker products and block norm matrices in large-scale systems analysis. *SIAM Journal of Matrix Analysis and Applications*, 10: 18 – 29, 1989.

[17] R. Jozsa. Fidelity for mixed quantum states. *Journal of Modern Optics*, 41(12): 2315–2323, 1994.

[18] I. Kerenidis and A. Nayak. Weak coin flipping with small bias. *Information Processing Letters*, 89(3): 131 – 135, 2004.

[19] A. Kitaev. Results presented at *QIP 2003* (slides and video available from MSRI). http://www.msri.org/publications/ln/msri/2002/qip/kitaev/1/index.html.

[20] R. Koning, H. Neudecker, and T. Wansbeek. Block Kronecker products and vecb operator. *Linear Algebra and its Applications*, 149: 165 – 184, 1991.

[21] H. Lo and H. Chau. Is quantum bit commitment really possible? *Physical Review Letters*, 78: 3410 – 3413, 1997.

[22] H. Lo and H Chau. Why quantum bit commitment and ideal quantum coin tossing are impossible. *Physica D*, 120: 177 – 187, 1998.

[23] D. Mayers. Unconditionally secure quantum bit commitment is impossible. *Physical Review Letters*, 78(17): 3414 – 3417, 1997.

[24] E. Mendelson. *Introducing Game Theory and Its Applications*. Chapman & Hall/CRC, 2004.

[25] C. Mochon. Quantum weak coin-flipping with bias of 0.192. In *Proceedings of 45th Symposium on Foundations of Computer Science*, pages 2 – 11, 2004.

[26] C. Mochon. Serial composition of quantum coin-flipping, and bounds on cheat detection for bit-commitment. *Physical Review A*, 70: article no. 032312, 2004.

[27] C. Mochon. A large family of quantum weak coin-flipping protocols. *Physical Review A*, 72: article no. 022341, 2005.

[28] A. Nayak. Assignment 2, CO 781, Topics in Quantum Information, Spring 2004, University of Waterloo, June 2004.
http://www.math.uwaterloo.ca/∼anayak/co781-s04/homework/hw2.pdf.

[29] A. Nayak and P. Shor. Bit-commitment-based quantum coin flipping. *Physical Review A*, 67: article no. 012304, 2003.

[30] Y. Nesterov and M. Todd. Self-scaled barriers and interior-point methods for convex programming. *Mathematics of Operations Research*, 22: 1 – 42, 1997.

[31] M. Nielsen and I. Chuang. *Quantum Computation and Quantum Information*. Cambridge University Press, 2000.

[32] P. Shor. Algorithms for quantum computation: Discrete logarithms and factoring. In *Proceedings of the 35th Annual Symposium on Foundations of Computer Science*, pages 124 – 134. IEEE Computer Society Press, 1994.

[33] R. Singh. Some generalizations in matrix differentiation with applications in multivariate analysis. Ph.D. Dissertation, University of Windsor, 1972.

[34] R. Singh and D. Tracy. A new matrix product and its applications in partitioned matrix differentiation. *Statistica Neerlandica*, 26: 143 – 157, 1972.

[35] R. Spalek and M. Szegedy. All quantum adversary methods are equivalent. *Theory of Computing*, 2(1): 1 – 18, 2006.

[36] R. W. Spekkens and T. Rudolph. Degrees of concealment and bindingness in quantum bit commitment protocols. *Physical Review A*, 65: article no. 012310, 2002.

[37] R. W. Spekkens and T. Rudolph. Quantum protocol for cheat-sensitive weak coin flipping. *Physical Review Letters*, 89: article no. 227901, 2002.

[38] L. Tunçel. Polyhedral and semidefinite programming methods in combinatorial optimization. CO 671 Lecture Notes, University of Waterloo, Waterloo, Ontario, Canada, 2005.

[39] L. Tunçel and H. Wolkowicz. Strengthened existence and uniqueness conditions for search directions in semidefinite programming. *Linear Algebra and its Applications*, 400: 31 – 60, 2005.

[40] A. Uhlmann. The 'transition probability' in the state space of *-algebra. *Reports on Mathematical Physics*, 9: 273 – 279, 1976.

[41] H. Wolkowicz, R. Saigal, and L. Vandenberghe, editors. *Handbook of Semidefinite Programming*. Kluwer Academic Publishers, 2000.

[42] W. Wootters and W. Zurek. A single quantum cannot be cloned. *Nature*, 299: 802 – 803, 1982.