A Numerical Analysis of the NPA Semidefinite Programming Hierarchy for the Mod P Game

by

Shalom Abate

Submitted to the Department of Electrical Engineering and Computer Science
in partial fulfillment of the requirements for the degree of

Master of Engineering in Computer Science and Engineering

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2017

Author .....
Department of Electrical Engineering and Computer Science
May 19, 2017

Certified by .....
Peter Shor
Morss Professor of Applied Mathematics
Thesis Supervisor

Accepted by .....
Christopher J. Terman
Chairman, Master of Engineering Thesis Committee

A Numerical Analysis of the NPA Semidefinite Programming Hierarchy for the Mod P Game

by

Shalom Abate

Submitted to the Department of Electrical Engineering and Computer Science
on May 19, 2017, in partial fulfillment of the
requirements for the degree of
Master of Engineering in Computer Science and Engineering

Abstract

The Mod P game is a generalization of the famous CHSH game [6] to a field of order p. The CHSH game corresponds to the Mod P game for the value of $p = 2$. The CHSH game was one of the earliest and most important results in quantum mechanics because it predicted a clear and experimentally verifiable separation between classical and quantum physics in the form of a Bell's inequality violation. In this thesis, we study the maximum winning probability for the Mod P game over the set of quantum strategies. For $p = 2$, an early result by Tsirelson [15] showed that the maximum winning probability by a quantum strategy is ≈ 0.854. This result is also tight in that it is achievable. Here we are interested in studying the game for values of $p > 2$ which has seen little progress over the years. This research thesis serves two purposes. The first is to create a self contained reference for some of the most important results in the area. Among these results, a prominent work is the NPA hierarchy [13] of semidefinite programs for testing whether a given bipartite correlation corresponds to a valid quantum mechanical experiment. The second part of this thesis is an implementation of this hierarchy for the Mod P game. In the first level of the hierarchy, we obtain numerical results that match analytic upper bounds by Bavarian and Shor [2]. We also find that the Bavarian and Shor bound is tighter than the first level NPA hierarchy value for a prime power p. In a collaborative work with Matthew Coudron we also present an approach for a semidefinite relaxation of the Mod P game using unitary operators. This approach brings us closer to achieving an exact analytic solution for the winning probability of the Mod P game.

Thesis Supervisor: Peter Shor
Title: Morss Professor of Applied Mathematics

Acknowledgments

I am tremendously grateful to Matthew Coudron, not only for his close oversight and collaboration on this work, but also for proposing this research project to me in the first place. This thesis would not have been possible without him. I also want to thank my thesis supervisor Peter Shor for his guidance and helpful discussions about the project.

I want to thank the folks at CSAIL and the Lincoln Laboratory Supercomputing Center for giving me access to their systems which allowed me to run some of the most memory intensive optimization programs I have ever worked with.

I also want to thank the open source convex optimization community who have been greatly helpful in finding the right tools for efficiently solving semidefinite programs, as well as debugging several issues that came up during the implementation of the optimization programs in this thesis.

Dedicated to the memory of my beloved mother, Berhan Mekonnen.

Contents

List of Tables

Chapter 1

Introduction

1.1 Quantum Nonlocality

Until as recently as the early 20^{th} century, physicists strongly held two fundamental beliefs about the natural universe. These are the notions of *locality* and *realism*. Locality is the proposition that the outcomes of two sufficiently spatially separated physical events cannot possibly be correlated. That is, as long as two physical events are sufficiently far apart in space, they cannot affect one another. Realism is the proposition that the physical universe exists independent of observation.

With the discovery of quantum mechanics in the late 19^{th} and early 20^{th} century, physicists have had to abandon both of these notions. Quantum mechanics dictates that on a fundamental level in the natural universe, neither locality nor realism hold. Thus, two spatially separated events can indeed be correlated. In this sense quantum mechanics predicts that the universe is *nonlocal*. Furthermore, quantum mechanics shows that physical quantities may be left undetermined until the moment of observation. In this sense, the universe is non-real. In this thesis we explore the limits of quantum nonlocality.

Throughout the development of quantum mechanics, Einstein notably remained a critic of its physical implications. In 1935, Einstein along with Boris Podolsky and Nathan Rosen published a paper in which they derive a seemingly paradoxical prediction of quantum mechanics, which has come to be known as the EPR paradox.

This was the case of quantum nonlocality and nonrealism. Einstein was bothered that a theory of the universe predicts that the outcome of an event could instantaneously affect that of another which is spatially separated. This happens because of a phenomenon predicted by quantum mechanics known as *quantum entanglement*. He famously referred to this effect as "spooky action at a distance." His central objection to the the theory can be summarized as follows. *Instantaneous action at a distance cannot be possible because information cannot travel faster than the speed of light. Therefore, there must be some hidden local variables which we do not yet understand that can explain this phenomenon.*

1.2 Bell's Inequality and the CHSH Game

In 1964, John S. Bell proposed an experiment which would definitively decide whether or not certain physical effects of quantum entanglement could be reproduced by local hidden variables [3]. His result, known as Bell's inequality, is a condition that any physical experiment had to satisfy if nature could be faithfully described by a classical local hidden variable theory. But it would have to be violated if nature was really quantum mechanical. Subsequently, several experiments have been designed that have violated Bell's inequality, proving that no hidden variable theory can describe certain physical phenomena such as quantum entanglement.

1.2.1 The CHSH Game

The CHSH game is an experiment in the form of a two player, one round nonlocal game which imposes a type of Bell's inequality. It was proposed in 1969 by Clauser, Horne, Shimony and Holt [6].

The game is cooperatively played by two players, Alice and Bob, against a referee. The game proceeds as follows. First, the referee selects two "question" bits s and t uniformly at random from $\{0,1\}$ and sends s to Alice and t to Bob. Alice and Bob are physically separated from each other and cannot communicate during the game. After receiving her question s, Alice picks her answer $a \in \{0,1\}$ and sends it back to the verifier. Similarly, after receiving t, Bob responds with his answer $b \in \{0,1\}$. The verifier then compares the values (a XOR b) with (s AND t). Alice and Bob win the game if $a \oplus b = s \wedge t$ and lose otherwise.

Since Alice and Bob are not allowed to communicate during the game, they must agree on a strategy beforehand. We call a strategy *quantum* if Alice and Bob are allowed to share a quantum state $|\psi\rangle \in \mathbb{C}^d \otimes \mathbb{C}^d$ and can perform quantum measurements on their share of the state. Analogously, we call a strategy *classical* if they are not allowed to share a quantum state. Ultimately, we want to compare the best classical strategy with the best quantum strategy for the CHSH game. For any given

strategy that Alice and Bob choose, they have some probability of winning the game. Thus we are interested in the maximum probability with which Alice and Bob can win the game.

We define the *classical value* of the game as the maximum winning probability over all classical strategies. Similarly, the *quantum value* of the game is the maximum probability with which Alice and Bob can win using a quantum strategy.

In the following sections we show that the classical value of the CHSH game is $\frac{3}{4} = 0.25$, however the quantum value is $\frac{1}{2} + \frac{2\sqrt{2}}{8} \approx 0.85$. In other words, Alice and Bob can win the game with higher probability if they are allowed to use quantum mechanics. In this specific game, the classical value defines a Bell's inequality which any classical strategy is bounded by. Thus if experiments show that Alice and Bob win the game with probability $> \frac{3}{4}$, then this inequality is violated and nature must truly be quantum mechanical.

1.2.2 The Classical Value of the CHSH Game

Before giving an upper bound on the classical value of CHSH game, consider the following, somewhat obvious, classical strategy that achieves a winning probability of 3/4.

Lemma 1.2.1. *The fixed strategy where Alice always answers with $a = 0$ and Bob always answers with $b = 0$ achieves a winning probability of 3/4.*

Proof. Since the verifier picks s and t uniformly at random, $\Pr_{s,t}[s \wedge t = 1] = \frac{1}{4}$. Therefore, with probability 3/4, $s \wedge t = 0$. Because Alice and Bob always return 0 and $0 \oplus 0 = 0$, they win the game with probability 3/4. □

We now show that this simple strategy is actually optimal for any classical strategy. That is, we show that the classical value of the CHSH game is at most 3/4.

Theorem 1.2.2. *Any classical strategy for the CHSH game has a winning probability of at most 3/4.*

Proof. We can characterize any classical strategy for Alice as a pair of probability distribution functions $F_0 : \{0,1\} \to [0,1]$ and $F_1 : \{0,1\} \to [0,1]$ such that on question s, Alice responds with answer a with probability $F_s(a)$.

Similarly Bob has probability distribution functions G_0 and G_1 such that on question t, Bob returns answer b with probability $G_t(b)$.

$$\Pr[a|s] = F_s(a)$$
$$\Pr[b|t] = G_t(b)$$

Because each F_s and G_t define probabilities, we must have that for any s and t,

$$F_s(0) + F_s(1) = 1$$
$$G_t(0) + G_t(1) = 1$$

Note that our simple fixed strategy example in lemma 1.2.1 corresponds to setting $F_s(0) = 1$ and $G_t(0) = 1$ so that Alice and Bob always respond with 0.

We can now define the joint probability that Alice and Bob respond with answers (a, b) on questions (s, t) as follows.

$$\Pr_{a,b}[a, b|s, t] = F_s(a) \cdot G_t(b)$$

It is easy to verify that indeed $\sum_{a,b} \Pr[a, b|s, t] = 1$. Let $V(a, b|s, t)$ be the verification function such that $V(a, b|s, t) = 1$ if $a \oplus b = s \wedge t$ and 0 otherwise. Let $P(F_s, G_t)$ be the probability that Alice and Bob win the CHSH game, using strategies F_s for Alice and G_t for Alice and Bob. We can write the probability as follows.

$$\begin{aligned} P(F_s, G_t) &= \sum_{s,t} \Pr[(s,t)] \sum_{a,b} \Pr[a, b|s, t] V(a, b|s, t) \\ &= \frac{1}{4} \sum_{a,b,s,t} F_s(a) G_t(b) V(a, b|s, t) \end{aligned}$$

We can further simplify this equation by writing $V(a,b|s,t)$ more explicitly as follows.

$$V(a,b|s,t) = (-1)^{a\oplus b}\left(\frac{(-1)^{a\oplus b}+(-1)^{s\wedge t}}{2}\right)$$

It is easy to verify that $V(a,b|s,t)=1$ if and only if $a\oplus b = s\wedge t$. Plugging this into our expression for $P(F_s,G_t)$, we get the following.

$$\begin{aligned}
P(F_s,G_t) &= \frac{1}{4}\sum_{a,b,s,t} F_s(a)G_t(b)(-1)^{a\oplus b}\left(\frac{(-1)^{a\oplus b}+(-1)^{s\wedge t}}{2}\right)\\
&= \frac{1}{8}\sum_{a,b,s,t} F_s(a)G_t(b) + \frac{1}{8}\sum_{a,b,s,t} F_s(a)G_t(b)(-1)^{a\oplus b}(-1)^{s\wedge t}\\
&= \frac{1}{2} + \frac{1}{8}\sum_{s,t}(-1)^{s\wedge t}\sum_{a,b} F_s(a)G_t(b)(-1)^{a\oplus b}\\
&\leq \frac{1}{2} + \frac{1}{8}\sum_{s,t}(-1)^{s\wedge t}\\
&= \frac{1}{2}+\frac{1}{4} = \frac{3}{4}
\end{aligned}$$

Thus, for any pair of classical strategies F_s for Alice and G_t for Bob, the probability of winning the game, $P(F_s,G_t)$, is at most $\frac{3}{4}$. In fact, because we gave a simple strategy that achieves this upper bound, the classical value of the CHSH game is exactly 3/4. □

1.2.3 The Quantum Value of the CHSH Game

Recall that for a quantum strategy, Alice and Bob are allowed to share a quantum state $|\psi\rangle \in \mathbb{C}^d \otimes \mathbb{C}^d$ in any fixed dimension $d \geq 1$. Alice and Bob can then make local measurements on their share of the state as part of their strategy.

Before we give an upper bound on the quantum value of the CHSH game, consider the following quantum strategy. Alice and Bob share the maximally entangled state in 2 dimensions $|\psi\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$, where the first qubit corresponds to Alice's share and the second qubit is Bob's. Their strategy is then as follows.

Upon receiving question $s=0$, Alice measures her qubit in the standard basis

$\{|0\rangle, |1\rangle\}$ and returns her measurement result. And upon receiving $s = 1$, Alice measures her qubit in the basis $\left\{\cos(\frac{\pi}{4})(|0\rangle + |1\rangle), \cos(\frac{\pi}{4})(|0\rangle - |1\rangle)\right\}$.

Similarly, upon receiving question $t = 0$, Bob measures his qubit in the basis $\left\{\cos(\frac{\pi}{8})|0\rangle + \sin(\frac{\pi}{8})|1\rangle, \sin(\frac{\pi}{8})|0\rangle - \cos(\frac{\pi}{8})|1\rangle\right\}$ and returns his measurement result. And finally, upon receiving $t = 1$, Bob measures his qubit in the basis $\left\{\cos(\frac{\pi}{8})|0\rangle - \sin(\frac{\pi}{8})|1\rangle, -\sin(\frac{\pi}{8})|0\rangle - \cos(\frac{\pi}{8})|1\rangle\right\}$ and returns his measurement result.

Lemma 1.2.3. *With the quantum strategy described above, Alice and Bob win the CHSH game with probability* $\cos^2(\frac{\pi}{8}) = \frac{1}{2} + \frac{\sqrt{2}}{4} \approx 0.854$.

Proof. The proof is quite straightforward. For each pair of questions (s, t), we compute the probability that Alice and Bob win the game. We can then add these probabilities up to get the total probability of winning the game. In all cases, without loss of generality, we may assume Alice performs her measurement first, and then Bob performs his measurement. Indeed the order in which Alice and Bob make their measurements does not affect the probabilities of their final outcomes, since their measurement operators commute.

For clarity, we explicitly give the first case, which is when $s = 0$ and $t = 0$. In this case, $s \wedge t = 0$. We are therefore interested in the probability that Alice and Bob respond with the same value so that $a \oplus b = 0$. When alice performs her measurement in the $\{|0\rangle, |1\rangle\}$ basis, she gets 0 or 1 with equal probability and Bob is left with the state $|0\rangle$ or $|1\rangle$ respectively. In the first case, writing $|0\rangle$ in Bob's measurement basis for the $t = 0$ case gives a probability of $\cos^2(\frac{\pi}{8})$ of obtaining an outcome of 0. Similarly, in the second case, writing $|1\rangle$ in Bob's basis for $t = 0$ gives a probability of $\cos^2(\frac{\pi}{8})$ of obtaining an outcome of 1. Therfore, Alice and Bob win with probability $\cos^2(\frac{\pi}{8})$ in the first case.

The same result can be shown for cases 2, 3 and 4 where $(s = 0, t = 1)$, $(s = 1, t = 0)$ and $(s = 1, t = 1)$ respectively. In each case, Alice and Bob win with probability $\cos^2(\frac{\pi}{8})$. Therefore, because each case occurs with probability $\frac{1}{4}$, Alice and Bob win the game with total probability $\frac{4\cos^2(\frac{\pi}{8})}{4} = \cos^2(\frac{\pi}{8}) = \frac{1}{2} + \frac{\sqrt{2}}{4} \approx 0.854$. □

It turns out that the strategy given in lemma 1.2.3 is optimal. In the 1980s, Boris

Tsirelson showed a series of results in the theory of nonlocal games, one of which was the quantum value of the CHSH game[15]. Here we present a proof for the following theorem.

Theorem 1.2.4 (Tsirelson)**.** *The quantum value of the CHSH game is at most* $\frac{1}{2} + \frac{\sqrt{2}}{4} \approx 0.854$.

Proof. The following proof was presented by Thomas Vidick in his article on Tsirelson's result [16]. We can describe any quantum strategy in the following standard and generalized form. Recall that Alice and Bob share the state $\psi \in \mathbb{C}^d \otimes \mathbb{C}^d$ for some fixed $d \geq 1$. Upon receiving s Alice performs a measurement on her share of the state using d-dimensional orthogonal projection operators $\{A_s^0, A_s^1\}$ such that $A_s^0 + A_s^1 = \mathbb{1}_{d\times d}$ and returns her measurement result. Similarly, upon receiving t, Bob measures his qubit in orthogonal projections $\{B_t^0, B_t^1\}$ such that $B_t^0 + B_t^1 = \mathbb{1}_{d\times d}$. Note that orthogonality means $A_s^0 A_s^1 = 0$ and $B_t^0 B_t^1 = 0$ for all s, t.

Now, the probability that upon receiving pairs (s, t), Alice and Bob respond with (a, b) respectively is given by the following.

$$\Pr[a, b|s, t] = \langle\psi| A_s^a \otimes B_t^b |\psi\rangle$$

We first need to ensure that this formulation yields a valid probability distribution function. Since A_s^a and B_t^b are projection operators, i.e. $(A_s^a)^2 = A_s^a$ and $(B_t^b)^2 = B_t^b$, they have eigenvalues 0 or 1. Therefore, $\Pr[a, b|s, t] \geq 0$ for all s, t, a, b. Furthermore because $A_s^0 + A_s^1 = \mathbb{1}_{d\times d}$ and $B_t^0 + B_t^1 = \mathbb{1}_{d\times d}$ for all s, t, we have that $\sum_{a,b} P[a, b|s, t] = 1$ for all s, t.

Let us define $P(|\psi\rangle, \{A_s^a\}, \{B_t^b\})$ to be the total probability of Alice and Bob winning the CHSH game using the shared state $|\psi\rangle$, and measurement operators $\{A_s^a\}$ for Alice and $\{B_t^b\}$ for Bob. Throughout this thesis, will be using this notation to represent the winning probability of any given strategy. We can write an expression

for this probability using the rule of total probability as follows.

$$\begin{aligned}
P(|\psi\rangle, \{A_s^a\}, \{B_t^b\}) &= \sum_{s,t} \Pr[s,t] \sum_{a,b} \Pr[a,b|s,t] V(a,b|s,t) \\
&= \frac{1}{4} \sum_{a,b,s,t} \langle\psi| A_s^a \otimes B_t^b |\psi\rangle (-1)^{a\oplus b} \left(\frac{(-1)^{a\oplus b} + (-1)^{s\wedge t}}{2} \right) \\
&= \frac{1}{8} \sum_{a,b,s,t} \langle\psi| A_s^a \otimes B_t^b |\psi\rangle + \frac{1}{8} \sum_{s,t} (-1)^{s\wedge t} \sum_{a,b} (-1)^{a\oplus b} \langle\psi| A_s^a \otimes B_t^b |\psi\rangle \\
&= \frac{1}{2} + \frac{1}{8} \sum_{s,t} (-1)^{s\wedge t} \sum_{a,b} (-1)^{a\oplus b} \langle\psi| A_s^a \otimes B_t^b |\psi\rangle
\end{aligned}$$

Let us define the operators $A_s = A_s^0 - A_s^1$ and $B_t = B_t^0 - B_t^1$. Immediately, we can observe that $A_s^2 = B_t^2 = \mathbb{1}$ and have ± 1 eigenvalues. We can now simplify our expression further using A_s and B_t.

$$\begin{aligned}
P(|\psi\rangle, \{A_s^a\}, \{B_t^b\}) &= \frac{1}{2} + \frac{1}{8} \sum_{s,t} (-1)^{s\wedge t} \sum_{a,b} (-1)^{a\oplus b} \langle\psi| A_s^a \otimes B_t^b |\psi\rangle \\
&= \frac{1}{2} + \frac{1}{8} \sum_{s,t} (-1)^{s\wedge t} \langle\psi| A_s \otimes B_t |\psi\rangle
\end{aligned}$$

Finally, it only remains to show that $\sum_{s,t} (-1)^{s\wedge t} \langle\Psi| A_s \otimes B_t |\Psi\rangle \leq 2\sqrt{2}$. First, note that we can take the summation inside to obtain,

$$\begin{aligned}
\sum_{s,t} (-1)^{s\wedge t} \langle\Psi| A_s \otimes B_t |\Psi\rangle &= \langle\Psi| \left(\sum_{s,t} (-1)^{s\wedge t} A_s \otimes B_t \right) |\Psi\rangle \\
&= \langle\Psi| \left(A_0 \otimes B_0 + A_0 \otimes B_1 + A_1 \otimes B_0 - A_1 \otimes B_1 \right) |\Psi\rangle
\end{aligned}$$

We can then bound the norm of the expanded operator as follows.

$$
\begin{aligned}
&= (A_0 \otimes B_0 + A_0 \otimes B_1 + A_1 \otimes B_0 - A_1 \otimes B_1)^2 \\
&= (A_0 \otimes (B_0 + B_1) + A_1 \otimes (B_0 - B_1))^2 \\
&= \mathbb{1} \otimes (B_0 + B_1)^2 \\
&+ \mathbb{1} \otimes (B_0 - B_1)^2 \\
&+ A_0 A_1 \otimes (B_0 + B_1)(B_0 - B_1) \\
&+ A_1 A_0 \otimes (B_0 - B_1)(B_0 + B_1) \\
&= 4\mathbb{1} \otimes \mathbb{1} + (A_0 A_1 - A_1 A_0) \otimes (B_1 B_0 - B_0 B_1) \\
&\leq_{\text{OP}} 8\mathbb{1} \otimes \mathbb{1}
\end{aligned}
$$

Where the last inequality bounds the operator norm of the expression by the fact that all eigenvalues are ± 1. Therefore, this implies that the square root of the original expression is bounded by $\sqrt{8} = 2\sqrt{2}$. We therefore get the following bound.

$$
\begin{aligned}
P(|\psi\rangle, \{A_s^a\}, \{B_t^b\}) &= \frac{1}{2} + \frac{1}{8} \sum_{s,t} (-1)^{s \wedge t} \langle\psi| A_s \otimes B_t |\psi\rangle \\
&\leq \frac{1}{2} + \frac{\sqrt{2}}{4} \approx 0.854
\end{aligned}
$$

□

1.3 The Mod P Game

The Mod P game is a generalization of the CHSH game. It was first proposed by Buhrman and Massar [5] in 2005. In the Mod P game, the verifier sends Alice and Bob questions s, t chosen uniformly at random from $\{0, 1, 2, \ldots, p-1\}$. Alice and Bob then respond with answers $a, b \in \{0, 1, 2, \ldots, p-1\}$. Alice and Bob win the game if $a + b \equiv s \cdot t \mod p$. Note that for $p = 2$, this is exactly the CHSH game.

In this research, we are especially interested in the case $p = 3$ - the Mod 3 game. Ultimately, we are interested in finding a tight upper bound for the quantum value of the Mod 3 game. From here on out we shall simply refer to the quantum value of a game simply as the value of the game.

Chapter 2

Recent Results in Nonlocal Games

In this chapter we summarize some of the most important recent developments in the theory of nonlocal games. First we review the work of Julia Kempe, Oded Regev and Ben Toner [9] on finding an efficient algorithm using semidefinite programming for approximating the winning probability of a unique game and finding a strategy that achieves this approximate winning probability. We then review a result by Mohammad Bavarian and Peter Shor [2] which gives the first analytic upper bound for the winning probability of the Mod P game. They show a winning probability upper bound of $\frac{1}{p} + \frac{p-1}{p\sqrt{p}}$ for a prime or prime power p. For $p = 2$, this gives the famous Tsirelson bound. It is known however that this bound is not tight for $p > 2$. We finally present a summary of perhaps the most important result by Navascués, Pironio and Acín [13] which we will call the NPA hierarchy of semidefite programs.

2.1 Two Player One Round Games

The Mod P game is a specific instance of a more general class of problems in the form of two player one round games, simply known as nonlocal games. This class of games are connected to the complexity class MIP* of interactive proofs with multiple provers with entanglement[7]. Here we give the general formulation of the winning probability for a general nonlocal game.

A nonlocal game is played by two players (known formally as provers), Alice and Bob and a verifier. A game $G = (\pi, Q, k, V)$ is specified by a set Q and a number $k \geq 1$, a probability distribution $\pi : Q \times Q \to [0,1]$ and a predicate $V : [k] \times [k] \times Q \times Q \to \{0,1\}$. The verifier samples $(s,t) \in Q \times Q$ according to π and sends question s to Alice and t to Bob. Alice replies with an answer $a \in [k]$, and Bob with an answer $b \in [k]$. Alice and Bob are not allowed to communicate during the game, but are allowed to share a quantum state $|\psi\rangle \in \mathbb{C}^d \otimes \mathbb{C}^d$, and they are allowed to perform projective local measurements on their share of state as part of their strategy. Alice's measurement is described by a set of orthogonal projectors $\{A_s^a\}$ for each $s \in Q$ such that $(A_s^a)^2 = A_s^a$ and $\sum_a A_s^a = \mathbb{1}$. Similarly, Bob has measurement operators $\{B_t^b\}$ which are orthogonal projectors.

Alice and Bob win the game if $V(a,b|s,t) = 1$. We will typically write this as $V(a,b|s,t)$ to distinguish the questions from the answers. We denote by $\omega^*(G)$ the entangled value of the game G, which can be written as,

$$\omega^*(G) = \lim_{d\to\infty} \max_{|\psi\rangle \in \mathbb{C}^d \otimes \mathbb{C}^d} \max_{A_s^a, B_t^b} \sum_{a,b,s,t} \pi(s,t) V(a,b|s,t) \langle\psi| A_s^a \otimes B_t^b |\psi\rangle .$$

In general, we are interested in finding $\omega^*(G)$ for a nonlocal game G.

2.2 Semidefinite Approximation of Unique Games

A nonlocal game $G = (\pi, V)$ on an alphabet of size k is *unique* if the verification predicate $V(a, b|s, t)$ only accepts answers of the form $b = \sigma(a)$ for some permutation σ of $[1, 2, \ldots k]$. In other words, for every answer a that Alice sends to the verifier, there is a unique answer $b = \sigma(a)$ that Bob must send in order to win the game.

Lemma 2.2.1. *The Mod P game is a unique game.*

Proof. For a fixed s and t, and a given value of a, there is a unique solution for $a + b \equiv s \cdot t \mod p$. The solution is $b \equiv s \cdot t - a \mod p$. Similarly, for any given b, there is a unique solution for a as $a \equiv s \cdot t - b \mod p$. □

In [9], Kempe, Regev and Toner give an efficient approximation algorithm for the entangled value $\omega^*(G)$ of a unique game G. Their method defines a semidefinite program over the set of unit vectors in $\mathbb{C}^d$ and gives an algorithm to round the optimal solution into a basis for Alice and Bob to measure their half of the shared quantum state. We summarize their result here.

Theorem 2.2.2 (Kempe, Regev, Toner)**.** *Let G be a unique game with entangled value $\omega^*(G) = 1-\varepsilon$. There exists an efficient algorithm that outputs a value $\frac{\varepsilon}{6} \leq \varepsilon' \leq \varepsilon$ and a description of an entangled strategy for Alice and Bob which has winning probability at least $1 - 6\varepsilon'$.*

While we do not present the proof of this theorem here, we sketch the construction given in the paper. Recall that we are trying to approximate $\omega^*(G)$ which is given by,

$$\omega^*(G) = \lim_{d\to\infty} \max_{|\psi\rangle \in \mathbb{C}^d \otimes \mathbb{C}^d} \max_{A_s^a, B_t^b} \sum_{a,b,s,t} \pi(s,t) V(a,b|s,t) \langle\psi| A_s^a \otimes B_t^b |\psi\rangle .$$

The authors present the following SDP relaxation over the real vectors $\{u_s^a\}$ and $\{v_t^b\}$ in $n \geq 1$ dimensions.

SDP 2.1: (Kempe, Regev, Toner) Unique Games SDP Relaxation

$$\begin{aligned}
\text{Maximize: } & \sum_{a,b,s,t} \pi(s,t)V(a,b|s,t)\langle u_s^a, v_t^b\rangle \\
\text{Subject to: } & \forall s,t,\ \sum_{a,b}\langle u_s^a, v_t^b\rangle = 1 \text{ and } \sum_a \langle u_s^a, u_s^a\rangle = \sum_b \langle v_t^b, v_t^b\rangle = 1 \\
& \forall s,t,\ \forall a \neq b,\ \langle u_s^a, u_s^b\rangle = 0 \text{ and } \langle v_t^a, v_t^b\rangle = 0 \\
& \forall s,t,a,b\ \langle u_s^a, v_b^t\rangle \geq 0
\end{aligned}$$

The first and last constraints ensure that for any s and t, $\langle u_s^a, v_t^b\rangle$ describes a probability distribution over Alice's and Bob's answers a and b given questions s and t respectively. The second constraint requires that for any fixed s Alice's vectors $\{u_s^a\}$ form an orthogonal basis for $\mathbb{C}^k$. Similarly, for any fixed t, Bob's vectors $\{v_t^b\}$ are orthogonal and form a basis. This constraint is called completeness.

For a given game G, let $\omega_{SDP}(G)$ be the optimal value of SDP 2.1. It is not difficult to show that $\omega^*(G) \leq \omega_{SDP}(G)$. The proof is essentially the fact that any entangled strategy for Alice and Bob that achieves a value of $\omega^*(G)$ can be converted to a feasible set of $\{u_s^a\}$ and $\{v_t^b\}$ for SDP 2.1 which achieves the same objective value of $\omega^*(G)$.

The authors then give a clever algorithm for rounding an optimal solution for the SDP into an entangled strategy for Alice and Bob. Given an optimal solution to SDP 2.1 with value $1-\varepsilon$, the main idea behind their rounding scheme is to extend the set of orthonormal vectors for Alice and Bob into an orthonormal basis of $\mathbb{C}^n$ and perform measurements on their share of the maximally entangled state in $\mathbb{C}^d \otimes \mathbb{C}^d$, $|\Psi\rangle = \frac{1}{\sqrt{n}}\sum_{i=1}^n |i\rangle \otimes |i\rangle$, in this basis.

This result has many interesting implications in the field of complexity theory. Namely, it implies that the entangled version of the *unique games conjecture* [10, 11] is false. Furthermore, it provides a method of finding lower bounds on the entangled value of general unique games, which includes the Mod P game.

2.3 The Bavarian-Shor Bound

In [2] Mohammad Bavarian and Peter Shor give the first general analytic upper bound on the maximum entangled winning probability of the Mod P game. Prior to this work, the only upper bound known for $p > 2$ was the value of $\frac{1}{3} + \frac{2}{3}\frac{1}{\sqrt{3}}$ for $p = 3$ by Buhrman and Massar [5]. Bavarian and Shor generalize this bound and prove the following asymptotic upper bound.

Theorem 2.3.1 (Bavarian, Shor)**.** *Let G_p be the Mod P game for any prime or prime power p. Then,*

$$w^*(G_p) \leq \frac{1}{p} + \frac{p-1}{p}\frac{1}{\sqrt{p}}.$$

In the same paper the authors also give a bound on the classical value of the Mod P game, which we do not present here. The authors give two different proofs of theorem 2.3.1. The first uses a reduction to a problem in information theory known as Information Causality [14].

In the Information Causality problem, Alice and Bob share an entangled quantum state. Alice is given a data set $X = (X_1, X_2, \cdots, X_N)$ from a distribution π, and Bob is given an index $b \in [N]$. Alice then makes a measurement on her system and sends some value $\alpha \in \Sigma$ to Bob. Bob then makes a measurement on his system and outputs $Z \in \Lambda$, which is his guess for X_b. Alice and Bob want to maximize the quantity $IC(A, B) = \sum_{i=1}^{N} I(X_i; Z|b = i)$. This quantity can be thought of a measure of how well Bob's output can predict Alice's data set in terms of their correlation. In one extreme, the mutual information $I(A; B)$ of two systems A and B is 0 if A and B are completely independent (uncorrelated). On the other hand, if A and B are fully correlated, i.e. have the same distribution, then their mutual information $I(A; B) = H(A) = H(B)$ is simply the entropy of A.

The principle of Information Causality states the following.

Theorem 2.3.2 (Information Causality)**.** *The Information Causality $IC(A; B)$ be-*

tween Alice and Bob for the scenario described above is

$$\sum_{i=1}^{N} I(X_i; Z|b = i) = O_{|\Sigma|,|\Lambda|}(1)$$

Informally, the principle of Information Causality states that if Alice communicates m classical bits of information to Bob, then Bob cannot learn any more than m bits of information about Alice's data set, even with shared quantum entanglement and local measurements. For $m = 0$, this is simply the no-signaling principle which states that information cannot travel faster than the speed of light. In [14], Pawowski et al. generalize the no-signaling principle and argue for the principle of Information Causality as a fundamental law governing information in the physical universe.

In the original setting of Information Causality, each of the elements in Alice's data set $X = (X_1, X_2, \cdots, X_N)$ are drawn with full independence from some distribution π. In order to give a reduction from the Mod P game to the Information Causality game, Bavarian and Shor prove a stronger, more generalized result of Information Causality for only a pairwise independent data set for Alice.

Theorem 2.3.3 (Pairwise Independent Information Causality)**.** *In the scenario described for Information Causality, let Alice's data set $X = (X_1, X_2, \cdots, X_N)$ be drawn from a known distribution π with only pairwise independence. Then,*

$$\sum_{i=1}^{N} I(X_i; Z|b = i) = O_{|\Sigma|,|\Lambda|}(1)$$

Bavarian and Shor prove theorem 2.3.3 and give a reduction from the Mod P game to the pairwise independent Information Causality problem. They do this by choosing Alice's input uniformly at random from a pairwise independent subset of the generalized Hadamard code over $\mathbb{F}_p$ and ask Bob to guess one of Alice's coordinates uniformly at random. They show that in order to have theorem 2.3.3 hold for this instance of the Information Causality problem, theorem 2.3.1 must also hold.

In order to give their reduction, Bavarian and Shor prove an important result about the Mod P game which greatly simplified their analysis. This property is

called the regularization lemma. The regularization lemma states that it is safe to assume without loss of generality that any strategy for Alice and Bob makes all errors uniformly at random.

More formally, let us denote by $\mathcal{P}$ a strategy for Alice and Bob for the Mod P game as a collection of projector operators $\{A_s^a\}$ and $\{B_t^b\}$ and a shared quantum state $|\psi\rangle$. We say that a strategy $\mathcal{P}$ is *regular* if for all $1 \leq k \leq p-1$ and some fixed $0 \leq E \leq 1$ the following holds.

$$\Pr_{a,b \leftarrow \mathcal{P}}[a + b = st + k \mod p | s, t] = \frac{1}{p} - \frac{E}{p}$$

This would imply that $\omega(\mathcal{P})$, the winning probability of the strategy $\mathcal{P}$, can be written as follows.

$$\begin{aligned}
\omega(\mathcal{P}) &= \Pr_{a,b \leftarrow \mathcal{P}}[a + b = st \mod p | s, t] \\
&= 1 - \sum_{k=1}^{p-1} \Pr_{a,b \leftarrow \mathcal{P}}[a + b = st + k \mod p | s, t] \\
&= 1 - \sum_{k=1}^{p-1} \frac{1}{p} - \frac{E}{p} \\
&= 1 - \left(\frac{p-1}{p} - \frac{p-1}{p} E \right) \\
&= \frac{1}{p} - \frac{p-1}{p} E
\end{aligned}$$

In this formulation, E is known as the bias of the strategy. It is a measure of how much a strategy outperforms the trivial random guess which achieves $\frac{1}{p}$. So for $E = 0$, the strategy has the same winning probability as random guessing. And for $E = 1$, the strategy wins with the maximum probability of 1. In theorem 2.3.1, Bavarian and Shor essentially prove an upper bound of $E \leq \frac{1}{\sqrt{p}}$ for the optimal strategy of the Mod P game.

Bavarian and Shor prove the following regularization lemma.

Lemma 2.3.4 (Regularization Lemma)**.** *For any strategy $\mathcal{P}$ for the Mod P game with winning probability $\omega(\mathcal{P})$, there exists a generic method to obtain a regular protocol*

$\mathcal{P}^*$ from $\mathcal{P}$ such that $\omega(\mathcal{P}^*) = \omega(\mathcal{P})$.

Proof. The proof is constructive. We assume that Alice and Bob have a strategy $\mathcal{P}$ which achieves a winning probability of $\omega(\mathcal{P})$. We describe the following protocol which gives a regular strategy $\mathcal{P}^*$ for Alice and Bob. On inputs s and t, Alice and Bob first use shared randomness to agree on $1 \leq \alpha, \beta \leq p-1$ uniformly at random and $0 \leq \gamma, \delta \leq p-1$ also uniformly at random. They then use their strategy $\mathcal{P}$ on inputs $s' = \alpha s + \gamma \mod p$ and $t' = \beta t + \delta \mod p$. Let a' and b' be their outputs using their protocol $\mathcal{P}$ on s' and t'. Alice's final output will then be $a = \frac{1}{\alpha\beta}(a' - \delta\alpha s - \gamma\delta)$ and Bob's final output is $b = \frac{1}{\alpha\beta}(b' - \beta\gamma t)$. Note that $\frac{1}{\alpha\beta}$ in this notation refers to the multiplicative inverse of $\alpha\beta$ in $\mathbb{F}_p^*$. The following shows the overall process of strategy $\mathcal{P}^*$.

$$\xrightarrow[\text{input}]{} s, t \xrightarrow[\text{transform}]{} s', t' \xrightarrow[\text{use } \mathcal{P}]{} a', b' \xrightarrow[\text{transform}]{} a, b \xrightarrow[\text{output}]{}$$

We first compute the winning probability of $\mathcal{P}^*$ as follows.

$$\begin{aligned}
\omega(\mathcal{P}^*) &= \Pr_{a,b \leftarrow \mathcal{P}^*}[a + b = st \mod p | s, t] \\
&= \Pr_{a',b' \leftarrow \mathcal{P}}\left[\frac{1}{\alpha\beta}(a' - \delta\alpha s - \gamma\delta) + \frac{1}{\alpha\beta}(b' - \beta\gamma t) = st \mod p | s, t\right] \\
&= \Pr_{a',b' \leftarrow \mathcal{P}}[a' + b' = \alpha\beta st + \delta\alpha s + \gamma\delta + \beta\gamma t \mod p | s, t] \\
&= \Pr_{a',b' \leftarrow \mathcal{P}}[a' + b' = s't' \mod p | s', t'] \\
&= \omega(\mathcal{P})
\end{aligned}$$

The last inequality follows from the fact that by construction, $s't'$ is uniformly distributed in $\{0, 1, \cdots, p-1\}$. Finally, we show that $\mathcal{P}^*$ is regular. We can show that

by following the same expansion for any $1 \leq k, k' \leq p-1$, $k \neq k'$ as follows,

$$\begin{aligned}
\Pr_{a,b \leftarrow \mathcal{P}^*}[a+b = st+k \mod p | s,t] &= \Pr_{a',b' \leftarrow \mathcal{P}}[a'+b' = s't'+k\alpha\beta \mod p | s',t'] \\
&= \Pr_{a',b' \leftarrow \mathcal{P}}[a'+b' = s't'+k'\alpha\beta \mod p | s',t'] \\
&= \Pr_{a,b \leftarrow \mathcal{P}^*}[a+b = st+k \mod p | s,t]
\end{aligned}$$

Where the second step of replacing k with k' follows form the fact that $k\alpha\beta$ is uniform over $\{1, 2, \cdots, p-1\}$ and $s't'$ is uniform over $\{0, 1, \cdots, p-1\}$. Thus $s't'+k\alpha\beta$ has the same distribution as $s't'+k'\alpha\beta$. □

This regularization result is very important. In section 3.2 we use this regularization lemma to give a characterization of the Mod P game in terms of unitary operators.

2.4 Quantum Correlations and the NPA Hierarchy

A line of work by Navascués, Pironio and Acín gives a characterization of the set of quantum correlations as a hierarchy of SDPs [13]. This result lays down the foundations of all our work in this thesis.

In this scenario, we are looking at a more general set of quantum correlations than nonlocal games. Suppose Alice and Bob were spatially separated, and are not allowed to communicate. Each output a of Alice is uniquely associated to a single input $X(a)$, which can be seen as disjoint subsets of A, all possible outputs of Alice. Similarly, Bob's inputs are disjoint subsets of B. A measurement scenario is then a quadruple $(A, B, \mathcal{X}, \mathcal{Y})$ where $\mathcal{X}$ and $\mathcal{Y}$ are disjoint subsets of A and B respectively.

We define a behavior P as a finite set of probabilities over $A \times B$ as $P = \{P(a,b) : a \in A,\ b \in B\}$. In the quantum correlation problem, we are interested in knowing whether P describes a *quantum behavior*.

The authors define a behavior P to be a *quantum behavior* if there exists a pure state $|\psi\rangle$ in a Hilbert space $\mathcal{H}$, a set of measurement operators $\{E_a : a \in A\}$ for Alice, and a set of measurement operators $\{E_b : b \in B\}$ for Bob, such that for all $a \in A$ and $b \in B$,

$$P(a,b) = \langle\psi| E_a E_b |\psi\rangle ,$$

where the measurement operators satisfy the following properties.

1. $E_a^\dagger = E_a$ and $E_b^\dagger = E_b$ (hermiticity)
2. $E_a E_{a'} = \delta_{aa'} E_a$ if $X(a) = X(a')$ and $E_b E_{b'} = \delta_{bb'} E_b$ if $Y(b) = Y(b')$ (orthogonal projection)
3. $\forall X \in \mathcal{X}$, $\sum_{a \in X} E_a = \mathbb{1}$ and $\forall Y \in \mathcal{Y}$, $\sum_{b \in Y} E_b = \mathbb{1}$ (completeness)
4. $[E_a, E_b] = 0$ (commutativity)

Let Q be the set of all quantum behaviors, i.e. $Q = \{P \ : \ P \text{ is a quantum behavior}\}$. We are then interested in characterizing the set Q using semidefinite programming.

The general outline of the hierarchy is as follows. Let $\mathcal{E} = \{E_a \ : \ a \in A\} \cup \{E_b \ : \ b \in B\}$ be the set of all projectors satisfying the properties outlined in the above definition of a quantum behavior. Let $\mathcal{O} = \{O_1, O_2, \ldots O_n\}$ be a finite subset of the algebra generated by $\mathcal{E}$. Define $\mathcal{F}(\mathcal{O})$ as the set of all independent equalities of the form,

$$\sum_{i,j} (F_k)_{ij} \langle\psi| O_i^\dagger O_j |\psi\rangle = g_k(P) \qquad\qquad k = 1, 2, \ldots, m$$

which are satisfied by the operators O_i, where the coefficients $g_k(P)$ are linear functions of the probabilities $P(a, b)$:

$$g_k(P) = (g_k)_0 + \sum_{a,b} (g_k)_{ab} P(a, b)$$

and where $|\psi\rangle$ is such that it satisfies $P(a, b) = \langle\psi| E_a E_b |\psi\rangle$ as defined for a quantum behavior above. Note that all the constraints in the definition of the quantum behavior are contained in $\mathcal{F}(\mathcal{O})$ as they are a specific instance of the general form given above.

Now, let a *sequence* S be a finite product of elements from $\mathcal{E}$ and let the length $|S|$ be the minimum number of projectors needed to generate it. By convention, we define the length of the identity operator $|\mathbb{1}| = 0$. Define $\mathcal{S}_n$ to be the set of sequences of length smaller or equal to n. Thus we have,

$$\mathcal{S}_0 = \{\mathbb{1}\}$$
$$\mathcal{S}_1 = \mathcal{S}_0 \cup \{E_a : a \in A\} \cup \{E_b : b \in B\}$$
$$\mathcal{S}_2 = \mathcal{S}_0 \cup \mathcal{S}_1 \cup \{E_a E_{a'} : a, a' \in A\} \cup \{E_b E_{b'} : b, b' \in B\} \cup \{E_a E_b : a \in A, b \in B\}$$
$$\mathcal{S}_3 = \ldots$$

Note that in the original paper the authors define $\mathcal{O}$, and subsequently $\mathcal{S}$ over the set $\tilde{\mathcal{E}}$ that is an equivalent to $\mathcal{E}$ up to linear combinations. For the sake of keeping this summary simple, we define the sets over $\mathcal{E}$.

Clearly, we have that $\mathcal{S}_1 \subseteq \mathcal{S}_2 \subseteq \ldots$, and that any operator $O_i \in \mathcal{O}$ can be written

as linear combinations of operators in $\mathcal{S}_n$ for sufficiently large n.

The authors then prove the following lemma.

Lemma 2.4.1. *Let $\mathcal{O}$ be a set of operators and $\mathcal{F}(\mathcal{O})$ the set of equations of the form (38) satisfied by operators in $\mathcal{O}$. Then a necessary condition for a behavior P to be quantum is that there exists a complex hermitian $n \times n$ positive semidefinite matrix $\Gamma \succeq 0$ whose entries Γ_{ij} satisfy*

$$\sum_{i,j} (F_k)_{ij} \Gamma_{ij} = g_k(P) \qquad\qquad k = 1, 2, \ldots m$$

Moreover, if the coefficients F_k and g_k in (38) are real, we can take Γ to be real as well.

This lemma leads to a natural formulation of the problem of finding such a Γ for a given behavior P as a semidefinite program. The program can be formulated as follows.

$$\begin{aligned} &\max \ \ \lambda \\ &\text{Subject to: } \langle F_k \Gamma \rangle = g_k(P) \qquad\qquad k = 1, 2, \ldots m \\ &\qquad\qquad \Gamma - \lambda \mathbb{1} \succeq 0 \end{aligned}$$

A nonnegative solution $\lambda \geq 0$ to the above problem implies that there exists a positive semidefinite $\Gamma \succeq \lambda \mathbb{1} \succeq 0$ that satisfies (40). We call such a Γ a *certificate* for P in the sense that it is a proof that P is indeed a quantum behavior. A strictly negative solution $\lambda < 0$ for this SDP means that P does not represent a quantum behavior.

Another important observation of this SDP formulation of a quantum behavior is the invariance of a quantum behavior P under linear combinations of operators. The authors prove the following lemma.

Lemma 2.4.2. *Let $\mathcal{O}$ and $\mathcal{O}'$ be two sets of operators such that every operator in $\mathcal{O}'$ is a linear combination of operators in $\mathcal{O}$. Then, the existence of a certificate Γ*

associated to $\mathcal{O}$ (for a given P) implies the existence of a certificate Γ' associated to $\mathcal{O}'$.

Indeed this lemma is easy to see since all one needs to do is conjugate Γ with a matrix C so that $\Gamma' = C^\dagger \Gamma C$, where C is such that $O'_i = \sum_k C_{ik} O_k$ for all $O'_i \in \mathcal{O}'$.

We can now define a *certificate of order n*, denoted Γ^n to be a $|\mathcal{S}_n| \times |\mathcal{S}_n|$ matrix associated with the set of operators $\mathcal{S}_n$. Informally, Γ^n is a certificate of whether or not there exists a state $|\psi\rangle \in \mathbb{C}^n$ and operators in $\mathcal{S}_n$ such that P describes a quantum behavior which can be simulated using these operators and $|\psi\rangle$ as defined in (38).

Using this definition, and the fact that $\mathcal{S}_1 \subseteq \mathcal{S}_2 \subseteq \ldots \subseteq S_n \subseteq \ldots$, the family of certificate $\Gamma^1, \Gamma^2, \ldots, \Gamma^n \ldots$, is a hierarchy of conditions satisfied by quantum probabilities. Thus a strategy to verify whether P is a quantum behavior is to first check if there exists a certificate of order 1, Γ^1. If not, P is not a quantum behavior. If so, then we check for a certificate of order 2, Γ^2. Repeat the procedure in increasing order as long as P satisfies the previous tests.

It is easy to see that as long as P is a quantum behavior, there must exist a certificate Γ^n for any level of the hierarchy. Thus a natural question is whether for a given non-quantum behavior P, there exists some finite n' such that a certificate $\Gamma^{n'}$ necessarily does not exist. Furthermore, is there a bound on n' or the rate of convergence for such a certificate of non-quantum-ness. The authors give a characterization of the set Q and in fact prove that the hierarchy eventually converges, although the rate of convergence is not known.

In chapter 3, we essentially construct the NPA hierarchy for the Mod P game and obtain numerical values as upper bounds for the maximum winning probability. We also give a modified version of the hierarchy for unitary (non-Hermitian) operators based on the same principles as the NPA hierarchy.

Chapter 3

Semidefinite Relaxations for the Mod P Game

In this chapter we give two types of semidefinite relaxations for the Mod P game. The first is based on a direct formalism of projector operators for Alice and Bob acting on a shared quantum state. This SDP can be thought of as an implementation the NPA hierarchy [13] for the Mod P game. We use the form of the SDP as given by Coudron and Vidick in [7]. The second is a relaxation using an encoding of Alice's and Bob's measurement operators as unitary operators. The unitary relaxation presented in this chapter is a novel approach to this problem proposed by Matthew Coudron. The work presented in this section was a collaborative effort with Coudron.

The primary idea of a semidefinite relaxation is to capture a substantial subset of the constraints of a given optimization problem in a semidefinite program, in such a way that the optimal solution to the original problem is feasible for the relaxation. For the reader unfamiliar with semidefinite programming a brief overview is given in appendix A.

3.1 Projector Relaxations

Recall that in the Mod P game, we are looking for projector operators $\{A_s^a\}$ for Alice and $\{B_t^b\}$ for Bob, for $0 \leq a, b, s, t \leq p-1$ and a shared state $|\psi\rangle \in \mathbb{C}^d \otimes \mathbb{C}^d$ in some dimension $d \geq 1$ such that the following quantity, which is the winning probability $P(|\psi\rangle, \{A_s^a\}, \{B_t^b\})$, is maximized.

$$P(|\psi\rangle, \{A_s^a\}, \{B_t^b\}) = \frac{1}{p^2} \sum_{a,b,s,t} V(a,b|s,t) \langle\psi| A_s^a B_t^b |\psi\rangle$$

Here we assume that $A_s^a \equiv \hat{A}_s^a \otimes \mathbb{1}$ only acts on Alice's side of the shared state. And similarly $B_t^b \equiv \mathbb{1} \otimes \hat{B}_t^b$. Thus Alice's and Bob's measurement operators always commute.

In order to form our relaxation, for a given operator A_s^a, we define the vector $|A_s^a\rangle \equiv A_s^a |\psi\rangle$. Similarly define the vector $\left|B_t^b\right\rangle \equiv B_t^b |\psi\rangle$. Supposing we knew the dimension d of the Hilbert space where the optimal choices for Alice's and Bob's projectors and their shared quantum state $|\psi\rangle$ live, and we could enforce that the vectors correspond to images of projector operators acting on a quantum state in d dimensions, this relaxation would yield a tight solution by replacing our optimization over projectors $\{A_s^a\} \cup \{B_t^b\}$ by the vectors $\{|A_s^a\rangle\}$ and $\{\left|B_t^b\right\rangle\}$ in dimension d. This is because the inner product $\left\langle A_s^a \,\middle|\, B_t^b \right\rangle$ represents $\langle\psi| A_s^a B_t^b |\psi\rangle$.

In general, for any given set of such relaxed vectors, we can write our maximization objective problem over these vectors as follows.

$$P(\{|A_s^a\rangle\}, \{\left|B_t^b\right\rangle\}) = \frac{1}{p^2} \sum_{a,b,s,t} V(a,b|s,t) \left\langle A_s^a \,\middle|\, B_t^b \right\rangle$$

We observe that rather than the vectors themselves, we are interested in their inner products $\left\langle A_s^a \,\middle|\, B_t^b \right\rangle$. Consider the matrix M formed by the vectors $\{A_s^a\} \cup \{B_t^b\}$ as its

columns. The entries of the matrix $C = M^\dagger M$ correspond to all the inner products between these vectors. By construction, C is symmetric and positive semidefinite (PSD). That is, all its eigenvalues are non-negative.

Thus, consider a matrix C of variables indexed by these vectors such that $C_{A_s^a, B_t^b} \equiv \left\langle A_s^a \,\middle|\, B_t^b \right\rangle$. We can think of our optimization problem as a maximization over a linear function of the variables $C_{A_s^a, B_t^b}$ subject to the constraint that C is PSD, denoted by $C \succeq 0$. While this relaxation is valid, in that it would only give an upper bound on $P(\{|A_s^a\rangle\}, \{|B_t^b\rangle\})$, it is not necessarily tight, or even bounded for that matter. Nevertheless, we use this as a starting point.

To make this into a good relaxation, we can add more constraints on the inner products, which would force the relaxed vectors to satisfy more properties required of images of orthogonal projectors acting on a quantum state. For instance, we need to add constraints to ensure that the relaxed vectors satisfy that Alice's and Bob's measurement operators commute.

As a warm up, in the following section we demonstrate this idea by giving a semidefinite relaxation for the Mod 2 game.

3.1.1 Projector Relaxation for the Mod 2 Game

Recall that in the Mod 2 game, Alice and Bob are looking for 4 projectors each, $\{A_s^a\}$ and $\{B_t^b\}$ for $0 \leq a, b, s, t \leq 1$, such that they win the game with good probability. As shown in theorem 1.2.4, we know the maximum winning probability for the Mod 2 game is ≈ 0.854, and we also know projectors for Alice and Bob which achieve this value. However, for the sake of demonstration, we study the SDP relaxation for the Mod 2 game.

From the 8 projectors, we get us a total of 8 vectors $\{|A_s^a\rangle\} \cup \{|B_t^b\rangle\}$ in the projector relaxation. Let C be the 8×8 matrix of inner products of these vectors. Note that this includes inner products of the form $\left\langle A_s^a \,\middle|\, A_{s'}^{a'} \right\rangle$ and $\left\langle B_t^b \,\middle|\, B_{t'}^{b'} \right\rangle$. For $U, V \in \{A_s^a\} \cup \{B_t^b\}$, $C_{U,V}$ denotes the entry in the matrix corresponding to the inner product $\langle U \,|\, V \rangle$.

Using this formalism, consider the following semidefinite program, SDP 3.1, for

the Mod 2 game over the complex entries of C.

SDP 3.1: Projector SDP Relaxation for the Mod 2 Game

$$\begin{aligned}
\text{Maximize: } & \frac{1}{4}\sum_{a,b,s,t} V(a,b|s,t) C_{A_s^a, B_t^b} \\
\text{Subject to: } & C \succeq 0 \\
& (\forall s), C_{A_s^0, A_s^1} = 0 \\
& (\forall t), C_{B_t^0, B_t^1} = 0 \\
& (\forall s), \sum_{a=0}^{1} C_{A_s^a, A_s^a} = 1 \\
& (\forall t), \sum_{b=0}^{1} C_{B_t^b, B_t^b} = 1 \\
& (\forall s,t), \sum_{a,b} C_{A_s^a, B_t^b} = 1
\end{aligned}$$

In this SDP, the first constraint forces C to be postive semidefinite. This constraint implies that there exist a set of vectors $\{|A_s^a\rangle\} \cup \{|B_t^b\rangle\}$ whose inner products yield the matrix C. The vectors can be obtained from C by taking the Cholesky decomposition of the matrix. This decomposition is generally true for any positive semidefinite matrix.

The next two constraints enforce that for any given input, the set of measurement operators used by Alice and Bob are orthogonal. The next two constraints enforce that the marginal probability distributions of Alice and Bob are valid. That is, for any given input s, Alice's measurement operators A_s^0 and A_s^1, yield a valid probability distribution. And similarly for Bob.

The last constraint enforces that for any given pair of inputs (s,t), Alice's and Bob's measurements jointly form a valid probability distribution over their answers.

Note that this SDP is a valid relaxation for the Mod 2 game since any strategy for the Mod 2 game yields a set of vectors which are feasible for the SDP. And therefore

its optimal value is an upper bound on the actual maximum winning probability the Mod 2 game. In this case, it turns out that SDP 3.1 is actually tight. That is, the optimal value of the SDP is exactly equal to $\frac{1}{2} + \frac{\sqrt{2}}{4} \approx 0.854$, which is the maximum winning probability of the CHSH game. In general however, we only expect to get upper bounds on the actual winning probability using the method of SDP relaxation.

In the context of the NPA hierarchy [13], SDP 3.1 corresponds to the first level of the hierarchy for the Mod 2 game. Thus it is in fact known that this hierarchy converges to the actual winning probability in the first level [17].

3.1.2 Projector Relaxation for the Mod P Game

Here we present the general SDP for the N^{th} level of the NPA hierarchy for the Mod P game for any value of p. The SDP follows the form given in [7].

We will refer to the set $\Sigma = \{A_s^a \text{ for } 0 \leq s, a \leq p-1\} \cup \{B_t^b \text{ for } 0 \leq t, b \leq p-1\}$ as our alphabet for the program. Let $W_m \equiv \cup_{i=0}^{m} \Sigma^i$ be the set of all words of length at most m, where we define $\Sigma^0 = \{\mathbb{1}\}$ representing the identity operator. The N^{th} level NPA hierarchy is now an optimization problem over the set of complex positive semidefinite matrices $C \in \mathbb{C}^{|W_N| \times |W_N|}$. Similar to [7], for a string $U \in W_m$ we define $U^\dagger$ as the string U in reversed order. We then use the notation for strings $U, V \in W_N$, $C_{U^\dagger, V}$ corresponds to the entry in the matrix representing the inner product $\langle U \,|\, V \rangle = (|U\rangle)^\dagger \, |V\rangle$. We use this convention to intuitively manipulate the string entries of the matrix C as if we were working with operators.

The N^{th} level SDP can now be written as given in SDP 3.2. This SDP has the same form as given by Coudron and Vidick in [7].

SDP 3.2: Projector SDP Relaxation for the Mod P Game

Maximize: $\frac{1}{p^2} \sum_{a,b,s,t} C_{A_s^a, B_t^b} V(a,b|s,t)$

Subject to:

$$C \succeq 0 \qquad (0)$$

$$C_{\mathbb{1},\mathbb{1}} = 1 \qquad (1)$$

$$(\forall R \in \Sigma), (\forall U, V \in W_{N-1}), \quad C_{UR,V} = C_{U,RV} \qquad (2)$$

$$(\forall A_s^a, B_t^b \in \Sigma), (\forall U, V \in W_{N-1}), \quad C_{UA_s^a, B_t^b V} = C_{UB_t^b, A_s^a V} \qquad (3)$$

$$(\text{for } 0 \le s \le p-1), (\forall U \in W_{N-1}), (\forall V \in W_N), \quad \sum_{a=0}^{p-1} C_{UA_s^a, V} = C_{U,V} \qquad (4)$$

$$(\text{for } 0 \le t \le p-1), (\forall U \in W_{N-1}), (\forall V \in W_N), \quad \sum_{b=0}^{p-1} C_{UB_t^b, V} = C_{U,V} \qquad (4)$$

$$(\text{for } 0 \le s \le p-1), (\forall U, V \in W_{N-1}), (\forall a \neq a'), \quad C_{UA_s^a, A_s^{a'} V} = 0 \qquad (5)$$

$$(\text{for } 0 \le t \le p-1), (\forall U, V \in W_{N-1}), (\forall b \neq b'), \quad C_{UB_t^b, B_t^{b'} V} = 0 \qquad (5)$$

Recall that we are relaxing an operator U acting on $|\psi\rangle$ as $|U\rangle \equiv V|\psi\rangle$. Thus we represent an operator acting on Alice's and Bob's shared quantum state as a vector. For any two operators U, V our relaxation implies the following fundamental relationship. Note that we overload notation to use U as an operator when written without brackets, and corresponds to the string label of the operator when written as $|U\rangle$.

$$\begin{aligned} \langle\psi| UV |\psi\rangle &= (U^\dagger |\psi\rangle)^\dagger (V|\psi\rangle) \\ &\equiv (|U^\dagger\rangle)^\dagger |V\rangle && \text{apply relaxation} \\ &= \langle U | V \rangle && \text{by our string convention} \end{aligned}$$

Using this relationship between expectation values of operators and the inner products of vectors, we can make sense of our constraints from the SDP.

3.1.3 Understanding the Constraints

First note that a string entry $U \in W_m$ corresponds to a sequence of symbols of length at most m from the alphabet Σ. Hence U represents a sequence of operators constructed by putting together Alice's and Bob's measurement operators, and therefore corresponds to a valid quantum operator.

The relationship between operators $U, V \in W_m$ is therefore not arbitrary. Consider for example the length 2 operator $U = A_0^0 B_0^0$ and $V = B_0^1 A_0^1$. We have that $UV = A_0^0 B_0^0 B_0^1 A_0^1 = 0$ because $B_0^0 B_0^1 = 0$ by virtue of Bob's measurement operators for a fixed input being orthogonal. We therefore need to encode these constraints into the SDP and enforce the relationships required of the relaxed vectors.

Constraint (0) simply requires that C is positive semidefinite (PSD). This is required so that C is a valid inner product matrix of a set of $|W_N|$ vectors corresponding to the relaxations for each operator in W_N. As discussed before, for a given PSD matrix C, we can compute its Cholesky decomposition to obtain the $|W_N|$ vectors whose matrix of inner products produces C.

Constraint (1) requires that Alice and Bob share a valid quantum state $|\psi\rangle$. $C_{\mathbb{1},\mathbb{1}} = 1$ can be expanded using our relaxed vectors as $\langle \mathbb{1} \,|\, \mathbb{1} \rangle = 1$ where $\mathbb{1}$ corresponds to the identity operator. Working backwards, this relaxation corresponds to the requirement that $\langle \psi | \, \mathbb{1}\mathbb{1} \, | \psi \rangle = 1$. Which is of course required because $|\psi\rangle$ has to be a valid quantum state and therefore a unit vector.

Constraint (2) is an artifact of inner product notation. Expanding the constraint, we have that $\langle UR \,|\, V \rangle = \langle U \,|\, RV \rangle$ where $R \in \Sigma$ is a length 1 symbol. This constraint is necessary to ensure that the SDP treats each symbol as a sequence of individual operators.

Constraint (3) enforces that Alice's and Bob's measurement operators commute. Consider for instance the operator $UA_s^a B_t^b V$. We know that it must be equal to $UB_t^b A_s^a V$. Therefore we have that $\langle \psi | \, UA_s^a B_t^b V \, | \psi \rangle = \langle \psi | \, UB_t^b A_s^a V \, | \psi \rangle$, and in our relaxation, constraint (3) enforces this by setting $\left\langle UA_s^a \,\middle|\, B_t^b V \right\rangle = \left\langle UB_t^b \,\middle|\, A_s^a V \right\rangle$.

The two constraints labeled (4) enforce that Alice's and Bob's measurement op-

erators are complete. Recall that for a given in put s, in order to have a valid probability distribution over the outputs a, Alice's measurement operators have to satisfy $\sum_a A_s^a = \mathbb{1}$. This is called completeness. Extending this requirement to higher length operators, consider the operators UA_s^aV. We can sum over a to get $\sum_a UA_s^aV = U(\sum_a A_s^a)V = UV$. Thus in our relaxation language, we must have that $\sum_a \langle UA_s^a \,|\, V\rangle = \langle U \,|\, V\rangle$.

Finally constraints (5) enforce that Alice's and Bob's measurement operators are orthogonal projectors. That is, for any given input s, Alice's measurements satisfy $A_s^a A_s^{a'} = 0$ for $a \neq a'$. Similarly Bob's operators satisfy $B_t^b B_t^{b'} = 0$ for $b \neq b'$. Thus if such a pair of operators appears at any point within a sequence of operators, that entire operator must evaluate to 0. Thus applying this to our relaxations, we must enforce that $\left\langle UA_s^a \,\middle|\, A_s^{a'}V \right\rangle = 0$, and similarly $\left\langle UB_t^b \,\middle|\, B_t^{b'}V \right\rangle = 0$.

3.1.4 The Size of the SDP

The size of a semidefinite program is generally measured by the number of variables and constraints that it contains. Since the SDP is optimizing over the matrix C of size $|W_N| \times |W_N|$, the number of variables in the SDP is $O(|W_N|^2)$. Since C is symmetric, the actual number of variables is really the number of entries on or above the diagonal. Therefore more precisely, the number of variables is $\frac{1}{2}|W_N|(|W_N|+1)$. Thus we want to compute $|W_N|$ in terms of p and N.

Recall that $W_N = \cup_{i=0}^N \Sigma^i$. Since each Σ^i are disjoint, $|W_N| = \sum_{i=0}^N |\Sigma^i|$. Recall also that we have $\Sigma^0 = \{\mathbb{1}\}$ and therefore $|\Sigma^0| = 1$. Note that we also have $|\Sigma^i| = |\Sigma|^i$. Thus, we only need to compute $|\Sigma|$ for the Mod P game.

Since $\Sigma = \{A_s^a\} \cup \{B_t^b\}$ for $0 \leq s, t, a, b \leq p-1$, we have that $|\Sigma| = 2p^2$. Plugging

this back in, we obtain the size of W_N as follows.

$$\begin{aligned} |W_N| &= \sum_{i=0}^{N} |\Sigma|^i \\ &= \sum_{i=0}^{N} (2p^2)^i \\ &= \frac{(2p^2)^{N+1} - 1}{2p^2 - 1} \end{aligned}$$

Thus, the number of variables is $O(|W_N|^2) = O\left((2p^2)^{2N}\right)$.

To compute the number of constraint, we simply need to add up the number of constraints of each type. We ignore constraint (0) as it only enforces the domain of the optimization to the positive semidefinite matrices. Therefore for each constraint of type (1) - (5), we can count the total number of constraints of that type. Table 3.3 summarizes this count for each constraint.

Table 3.3: Number of Constraints for SDP 3.2

Constraint	Factor 1	Factor 2	Factor 3	Multiplicity	Total
(1)	1	1	1	1	1
(2)	$2p^2$	$\lvert W_{N-1}\rvert^2$	1	1	$2p^2\lvert W_{N-1}\rvert^2$
(3)	p^4	$\lvert W_{N-1}\rvert^2$	1	1	$p^4\lvert W_{N-1}\rvert^2$
(4)	p	$\lvert W_{N-1}\rvert$	$\lvert W_N\rvert$	2	$2p\lvert W_{N-1}\rvert\lvert W_N\rvert$
(5)	p	$\lvert W_{N-1}\rvert^2$	$p^2 - p$	2	$2p(p^2-p)\lvert W_{N-1}\rvert^2$

Plugging in for $|W_{N-1}| = \dfrac{(2p^2)^N - 1}{(2p^2) - 1}$ and $|W_N| = \dfrac{(2p^2)^{N+1} - 1}{(2p^2) - 1}$, we can add up all the different types of constraints and get the total number of constraints to be

$$(3p^2 + 2p^4 - 2p^3)\left(\frac{(2p^2)^N - 1}{2p^2 - 1}\right)^2 + 2p\left(\frac{(2p^2)^N - 1}{2p^2 - 1}\right)\left(\frac{(2p^2)^N - 1}{2p^2 - 1}\right).$$

With some simplifications, it is easy to see that the number of constraints is also $O(|W_N|^2) = O((2p^2)^{2N})$. Thus it is important to note that for a fixed p, the size of

the SDP grows exponentially in N, the level of the SDP hierarchy.

Table 3.4 gives numerical values for the exact number of constraints and variables for the first few values of p and N.

Table 3.4: Number of Variables and Constrains for the Projector Relaxation of the Mod P game.

p	N	#Variables	#Constraints
2	1	36	68
3	1	171	249
3	2	58,653	87,837
3	3	19,062,225	28,590,765
5	1	1,275	1,385
5	2	3,252,525	3,576,885
5	3	8,134,565,025	8,947,976,885

In chapter 4, we discuss in more detail the cost of running this semidefinite program. However, it is already clear that there is almost no hope of running this SDP beyond the 3^{rd} level for any value of p.

In the following section, we describe a second approach for building a semidefinite relaxation for the Mod 3 game, using unitary operators for Alice and Bob rather than projectors. While this unitary relaxation will also have the same exponential rate of growth in the level N, it would grow with a smaller base for the exponential.

3.2 Unitary Operators for the Mod P Game

The representation of the Mod P game's winning probability in terms of unitary operators was an idea developed by Matthew Coudron[1]. All the work in this section was done in collaboration and under the supervision of Coudron.

The unitary representation was motivated by the analysis of the original CHSH game in Tsirelson's theorem 1.2.4. In the proof of this theorem, we define unitary operators $A_s = A_s^0 - A_s^1$ and $B_t = B_t^0 - B_t^1$ for $s, t \in \{0, 1\}$. We then find a tight upper bound on the maximum winning probability by bounding the operator norm of a linear combination of these unitary operators. In this section we analogously define the unitary formulation for the Mod P game for a general value of p. In order to achieve this, we will invoke the regularization lemma from in [2] by Bavarian and Shor. In section 2.3, we restate this proposition in lemma 2.3.4 and give a proof.

We start by defining generalized unitary operators for the Mod P game and express the maximum winning probability as a linear combination of the expectation values of these operators. Recall that in the Mod P game, Alice has projector operators $\{A_s^a | 0 \leq s, a \leq p-1\}$ and Bob has projector operators $\{B_t^b | 0 \leq t, b \leq p-1\}$. For all $s, t \in \{0, \cdots, p-1\}$, consider the following operators created by taking complex linear combinations of their projector measurement operators.

$$A_s = \sum_{a=0}^{p-1} e^{\frac{2\pi i a}{p}} A_s^a$$
$$B_t = \sum_{b=0}^{p-1} e^{\frac{2\pi i b}{p}} B_t^b$$

Where $e^{\frac{2\pi i}{p}}$ is the principal p^{th} root of unity. Note that in the case where $p = 2$, this directly corresponds to the definitions from the proof of theorem 1.2.4. Therefore in a sense this definition is a generalization of this method.

[1] To our knowledge, this work was not part of any published results and is presented here for the first time.

We first observe that A_s and B_t are unitary operators. We can show this formally as follows.

$$\begin{aligned} A_s A_s^\dagger &= \left(\sum_{a=0}^{p-1} e^{\frac{2\pi i a}{p}} A_s^a\right)\left(\sum_{a=0}^{p-1} e^{\frac{2\pi i a}{p}} A_s^a\right)^\dagger \\ &= \sum_{a=0}^{p-1}\sum_{a'=0}^{p-1} e^{\frac{2\pi i(a-a')}{p}} A_s^a A_s^{a'} \\ &= \sum_{a=0}^{p-1} (A_s^a)^2 \\ &= \sum_{a=0}^{p-1} A_s^a \\ &= \mathbb{1} \end{aligned}$$

Now consider the following expectation value using the unitary operators for Alice and Bob. We can express it in terms of the original projector operators to get a relation with the winning probability of the game.

$$\begin{aligned} \langle\psi| A_s B_t |\psi\rangle &= \langle\psi| \sum_{a=0}^{p-1} e^{\frac{2\pi i a}{p}} A_s^a \sum_{b=0}^{p-1} e^{\frac{2\pi i b}{p}} B_t^b |\psi\rangle \\ &= \sum_{a=0}^{p-1}\sum_{b=0}^{p-1} e^{\frac{2\pi i(a+b)}{p}} \langle\psi| A_s^a B_t^b |\psi\rangle \end{aligned}$$

We can then perform a weighted sum over all s, t using the probability $\frac{1}{p^2}$ for a pair of questions (s, t) and the complementary coefficient $e^{\frac{-2\pi i s t}{p}}$ in order to get a coefficient of 1 for the values of a, b, s, t which evaluate $V(a, b|s, t) = 1$.

$$\begin{aligned} \frac{1}{p^2}\sum_{s=0}^{p-1}\sum_{t=0}^{p-1} e^{\frac{-2\pi i s t}{p}} \langle\psi| A_s B_t |\psi\rangle &= \frac{1}{p^2}\sum_{s=0}^{p-1}\sum_{t=0}^{p-1} e^{\frac{-2\pi i s t}{p}} \sum_{a=0}^{p-1}\sum_{b=0}^{p-1} e^{\frac{2\pi i(a+b)}{p}} \langle\psi| A_s^a B_t^b |\psi\rangle \\ &= \frac{1}{p^2}\sum_{a,b,s,t} e^{\frac{2\pi i(a+b-st)}{p}} \langle\psi| A_s^a B_t^b |\psi\rangle \end{aligned}$$

In this summation, note that for the values of a, b, s, t such that $a + b = st \mod p$, the complex coefficient becomes 1. Thus we can rewrite this summation to make the

component of the sum corresponding to the winning probability of the Mod P game more explicit.

$$\begin{aligned}\frac{1}{p^2}\sum_{s=0}^{p-1}\sum_{t=0}^{p-1} e^{\frac{-2\pi ist}{p}} \langle\psi| A_s B_t |\psi\rangle &= \frac{1}{p^2}\sum_{a,b,s,t} \langle\psi| A_s^a B_t^b |\psi\rangle V(a,b|s,t) \\ &+ \frac{1}{p^2}\sum_{k=1}^{p-1} e^{\frac{2\pi ik}{p}} \sum_{a+b-st=k \mod p} \langle\psi| A_s^a B_t^b |\psi\rangle\end{aligned} \tag{3.1}$$

To simplify notation, let us define the expectation on the left hand side of this expression as Z.

$$Z \coloneqq \frac{1}{p^2}\sum_{s=0}^{p-1}\sum_{t=0}^{p-1} e^{\frac{-2\pi ist}{p}} \langle\psi| A_s B_t |\psi\rangle$$

Let us further define V_k as the probability that Alice's and Bob's answers satisfy $a+b = st+k \mod p$ as follows.

$$\begin{aligned}V_k &\coloneqq \Pr_{|\psi\rangle,\{A_s^a\},\{B_t^b\}}[a+b = st+k \mod p] \\ &= \frac{1}{p^2}\sum_{a+b=st+k \mod p} \langle\psi| A_s^a B_t^b |\psi\rangle\end{aligned}$$

Note that V_0 is exactly the winning probability of the game given Alice's and Bob's strategy. i.e. $V_0 \equiv P(|\psi\rangle, \{A_s^a\}, \{B_t^b\})$. We can now invoke the regularization lemma 2.3.4 and assume without loss of generalization that $V_1 = V_2 = \cdots = V_{p-1}$[2]. Using these definitions, we can rewrite equation 3.1 more simply as follows.

$$\begin{aligned}Z &= V_0 + \sum_{k=1}^{p-1} e^{\frac{2\pi ik}{p}} V_k \\ &= V_0 + V_1 \sum_{k=1}^{p-1} e^{\frac{2\pi ik}{p}} \\ &= V_0 - V_1\end{aligned}$$

[2]In practice, this constraint must either be explicitly enforced in the SDP or we must use $Re(Z)$ as Z may be a complex number in general.

Note that since $\sum_{k=0}^{p-1} V_k = 1$, we have that $V_1 = \left(\frac{1-V_0}{p-1}\right)$. Substituting this in, we can solve for V_0 as follows.

$$V_0 = \frac{p-1}{p} Z + \frac{1}{p}$$

Note that since A_s and B_t are no longer Hermitian operators their expectation values will be complex. Therefore Z is generally complex value. Furthermore, Alice and Bob are no longer performing valid quantum measurements because only Hermitian operators are allowed as measurements in quantum mechanics. Nevertheless we can still think of a unitary strategy as if Alice and Bob were using the unitary operators to perform quantum measurements and obtaining complex expectation values as it is a convenient way to describe the relaxation. Thus, for any Mod P game, given a set of unitary operators $\{A_s\}$ for Alice, $\{B_t\}$ for Bob, where $0 \leq s, t, \leq p-1$, and a shared state $|\psi\rangle$, their winning probability $P(|\psi\rangle, \{A_s\}, \{B_t\}) \equiv V$ can be computed as follows.

$$\begin{aligned} P(|\psi\rangle, \{A_s\}, \{B_t\}) &\coloneqq \frac{p-1}{p} Re\left(Z\right) + \frac{1}{p} \\ &= \frac{p-1}{p^3} Re\left(\sum_{s,t} e^{\frac{-2\pi i s t}{p}} \langle\psi| A_s B_t |\psi\rangle\right) + \frac{1}{p} \end{aligned}$$

By using $Re(Z)$, we ensure that $P(|\psi\rangle, \{A_s\}, \{B_t\})$ is a real valued linear function of the expectation values. As we show in the next section, this is crucial because a semidefinite program requires a real valued linear objective function. Furthermore, we know by construction that there exists an optimal unitary strategy which achieves a value of Z such that $Im(Z) = 0$. We could in fact enforce this as a constraint in our semidefinite relaxatino in the next section.

Now that we know how to define the Mod P game in terms of unitary operators, we can use the same semidefinite relaxation principles to generate its NPA hierarchy [13].

3.2.1 Unitary Relaxation for the Mod P Game

Before defining our relaxation, we need to define our alphabet of symbols. Orthogonal projector operators were relatively simple to describe because any such operator O is Hermitian. i.e. $O^\dagger = O$. And therefore no distinction needs to be made between O and its adjoint $O^\dagger$.

Unitary operators, however, are more complex since they need not be Hermitian. We therefore need a formal symbol $O^\dagger$ for the adjoint of a unitary operator O. In our specific case however, if we really wanted to, there is one way to possibly avoid writing formal symbols for the adjoints of Alice's and Bob's unitary operators. For any value of p, we first make the observation that $(A_s)^p = \mathbb{1}$. This is easy to see, but for completeness we can prove it by expanding the definition for A_s as follows.

$$\begin{aligned}
(A_s)^p &= \left(\sum_{a=0}^{p-1} e^{\frac{2\pi i a}{p}} A_s^a\right)^p \\
&= \sum_{a_1,a_2,\cdots,a_p} e^{\frac{2\pi i(a_1+a_2+\cdots+a_p)}{p}} A_s^{a_1} A_s^{a_2} \cdots A_s^{a_p} \\
&= \sum_a (A_s^a)^p \\
&= \sum_a A_s^a \\
&= \mathbb{1}
\end{aligned}$$

Using this identity, we can see that $(A_s)^{p-1} A_s = \mathbb{1} = A_s^\dagger A_s$. Therefore we have that $A_s^\dagger = (A_s)^{p-1}$. Therefore, we can use $p-1$ copies of A_s to represent $A_s^\dagger$. However this would require a symbol of length $p-1$ to define the adjoint of a single operator. This would require running the NPA hierarchy for levels $N \geq p-1$ just to even capture the unitary-ness constraints. But we are interested in capturing as many constraints as possible in lower levels of the hierarchy as the size of the SDP grows exponentially in the level. We therefore resort instead to defining formal symbols $A_s^\dagger$ and $B_t^\dagger$ for the adjoint operators. This trade-off increases the base of the exponential for the size of the SDP, but results in more meaningful values in the lower levels of the hierarchy.

We can now begin by defining our alphabet as $\Sigma = \{A_s\} \cup \{A_s^\dagger\} \cup \{B_t\} \cup \{B_t^\dagger\}$ for $0 \leq s, t \leq p-1$. Similarly as before, we define $W_m \equiv U_{i=0}^m \Sigma^i$ to be the set of all words of length at most m on the alphabet Σ where we let $\Sigma^0 \equiv \{\mathbb{1}\}$, representing the identity operator.

For any symbol $U \in W_m$, our relaxation is once again to use a vector $|U\rangle$ as a relaxation for the operator described by U acting on Alice and Bob's shared quantum state. i.e. $|U\rangle \equiv U|\psi\rangle$ represents the operator U acting on the state $|\psi\rangle$. The N^{th} level of the hierarchy is then an optimization over the set of positive semidefinite matrices $C \in \mathbb{C}^{|W_N| \times |W_N|}$ where an entry $C_{U,V}$ for $U, V \in W_N$ corresponds to an inner product. However, unlike the projector relaxation, recall that the entries of $C_{U,V}$ are now complex and therefore C is a Hermitian PSD matrix (as opposed to symmetric). This requirement implies that $C_{U,V} = C_{V,U}^*$. In order to satisfy this, we must define $C_{U,V} \equiv \langle U^\dagger \,|\, V \rangle$, which corresponds to $\langle\psi| U^\dagger V |\psi\rangle$. This way, we have that $C_{V,U} \equiv \langle V^\dagger \,|\, U \rangle$ corresponding to $\langle\psi| V^\dagger U |\psi\rangle = (\langle\psi| U^\dagger V |\psi\rangle)^*$. Thus we have that $C_{U,V} = C_{V,U}^*$.

However, to make our notation intuitive, instead of refering to the entries of the matrix C itself, we define the map $\mathcal{C}$ over the entries of C as follows.

$$\mathcal{C}_{U,V} \coloneqq C_{U^\dagger,V}$$

Note that we can now think of $\mathcal{C}$ as representing the entries of the PSD matrix C, but satisfies $\mathcal{C}_{U,V} = C_{U^\dagger,V} \equiv \langle U \,|\, V \rangle \equiv \langle\psi| UV |\psi\rangle$. Furthermore, recall that since we have formal symbols for the adjoint, the string for $U^\dagger$ is no longer just the reverse string of U, but also where each symbol is converted to its adjoint, i.e. $(UA_s)^\dagger = A_s^\dagger U^\dagger$, where we define $(U^\dagger)^\dagger = U$ for any symbol $U \in W_N$ and $\mathbb{1}^\dagger = \mathbb{1}$.

We can now use our map $\mathcal{C}$ to define our relaxed SDP as given in SDP 3.5. In the following we assume p to be an odd prime or prime power, and we use the notation U^x to refer to the string $UUU\cdots U$ repeated x times.

SDP 3.5: Unitary SDP Relaxation for the Mod P Game

Maximize: $\frac{p-1}{p^3} Re\left(\sum_{s,t} e^{\frac{-2\pi ist}{p}} \mathcal{C}_{A_s,B_t}\right) + \frac{1}{p}$

Subject to:

$$C \succeq 0 \quad (0)$$

$$C_{\mathbb{1},\mathbb{1}} = 1 \quad (1)$$

$$(\forall R \in \Sigma), (\forall U, V \in W_{N-1}), \quad \mathcal{C}_{UR,V} = \mathcal{C}_{U,RV} \quad (2)$$

$$(\forall A \in \{A_s\} \cup \{A_s^\dagger\}, B \in \{B_t\} \cup \{B_t^\dagger\}), (\forall U, V \in W_{N-1}), \quad \mathcal{C}_{UA,BV} = \mathcal{C}_{UB,AV} \quad (3)$$

$$(\forall 0 \leq s \leq p-1), (\forall U, V \in W_{N-1}), \quad \mathcal{C}_{UA_s^\dagger,A_sV} = \mathcal{C}_{U,V} \quad (4)$$

$$\mathcal{C}_{UA_s,A_s^\dagger V} = \mathcal{C}_{U,V} \quad (4)$$

$$(\forall 0 \leq t \leq p-1), (\forall U, V \in W_{N-1}), \quad \mathcal{C}_{UB_t^\dagger,B_tV} = \mathcal{C}_{U,V} \quad (4)$$

$$\mathcal{C}_{UB_t,B_t^\dagger V} = \mathcal{C}_{U,V} \quad (4)$$

$$(\forall 0 \leq s \leq p-1), \left(\forall U, V \in W_{N-\left(\frac{p-1}{2}\right)}\right), \quad \mathcal{C}_{U(A_s)^{\frac{p-1}{2}},(A_s)^{\frac{p-1}{2}}V} = \mathcal{C}_{UA_s^\dagger,V} \quad (5)$$

$$\mathcal{C}_{U(A_s^\dagger)^{\frac{p-1}{2}},(A_s^\dagger)^{\frac{p-1}{2}}V} = \mathcal{C}_{UA_s,V} \quad (5)$$

$$(\forall 0 \leq t \leq p-1), \left(\forall U, V \in W_{N-\left(\frac{p-1}{2}\right)}\right), \quad \mathcal{C}_{U(B_t)^{\frac{p-1}{2}},(B_t)^{\frac{p-1}{2}}V} = \mathcal{C}_{UB_t^\dagger,V} \quad (5)$$

$$\mathcal{C}_{U(B_t^\dagger)^{\frac{p-1}{2}},(B_t^\dagger)^{\frac{p-1}{2}}V} = \mathcal{C}_{UB_t,V} \quad (5)$$

3.2.2 Understanding the Constraints

Recall that an entry $\mathcal{C}_{U,V} = C_{U^\dagger,V}$ is the value of $\langle U \,|\, V \rangle$ for our relaxed vectors $|U\rangle$ and $|V\rangle$ for $U, V \in W_N$ corresponding to the expectation value of the corresponding operators $\langle \psi | UV | \psi \rangle$. We write our constraints in SDP 3.5 using the map $\mathcal{C}$ to have our constraints resemble the more familiar notation of the inner products.

As a fundamental component of a semidefinite program, constraint (0) requires that C is a PSD matrix. And similarly to our projector relaxation, constraint (1) enforces that Alice and Bob share a valid quantum state.

Once again, similarly to our projector relaxation, constraint (2) encodes that there are multiple relaxations for the same expectation value. Expanding out this constraint we get that $\langle UR \,|\, V \rangle = \langle U \,|\, RV \rangle$. This is necessary to enforce since both values represent $\langle \psi | URV | \psi \rangle$ in our original problem.

So far, all the constraints were also present in the projector relaxation. Constraints (4) is where the unitary relaxation begins to deviate. These constraints enforce that Alice and Bob have unitary operators because we have that for a unitary operator A_s, $UA_s^\dagger A_s V = UV$. Thus we must have that an inner product of the form $\langle UA_s^\dagger \,|\, A_s V\rangle = \langle U \,|\, V\rangle$.

The final constraints, constraints (5), encode that Alice's and Bob's unitary p^{th} power to the identity property. Recall that by construction we must have $A_s^p = B_t^p = \mathbb{1}$. A more succinct way of representing this constraint is $A_s^{p-1} = A_s^\dagger$ and similarly $B_t^{p-1} = B_t^\dagger$. This is because that would imply $A_s^p = A_s A_s^{p-1} = A_s A_s^\dagger = \mathbb{1}$, where the last equality is enforced in the SDP by constraint (4). Conversely, by the same argument, we must also enforce that $(A_s^\dagger)^{p-1} = A_s$.

Therefore constraint (5) enforces that $\left\langle U(A_s)^{\frac{p-1}{2}} \,\middle|\, (A_s)^{\frac{p-1}{2}} V \right\rangle = \langle UA_s^\dagger \,|\, V\rangle$. By constraint (3) this would also imply that $\left\langle U(A_s)^{\frac{p-1}{2}} \,\middle|\, (A_s)^{\frac{p-1}{2}} V \right\rangle = \langle U \,|\, A_s^\dagger V\rangle$. In operator notation, this is enforcing that $\langle\psi| U(A_s)^{\frac{p-1}{2}} (A_s)^{\frac{p-1}{2}} V |\psi\rangle = \langle\psi| UA_s^\dagger V |\psi\rangle$.

Perhaps a most interesting observation about this unitary SDP relaxation is that the p^{th} power to unitary constraints, (constraints (5)) do not even exist until level $N \geq \frac{p-1}{2}$. For instance, once $p \geq 5$, we need the second level or more to get a tight SDP constraint. For $p = 3$, we can still get meaningful values for $N = 1$, and therefore we focus our attention on $p = 3$ for the remainder of this thesis.

3.2.3 The Size of the SDP

One of the primary motivations for the unitary relaxation is that it has a non-trivially smaller size than that of the projector relaxation. While its size still grows exponentially in the level N, it has a significantly smaller base due to the smaller sized alphabet. Here we compute the size for a generic p and level N.

Recall that the alphabet for SDP 3.5 is $\Sigma = \{A_s\} \cup \{A_s^\dagger\} \cup \{B_t\} \cup \{B_t^\dagger\}$ for $0 \leq s, t \leq p-1$. Hence we have that $|\Sigma| = 4p$. Using the same definition for $W_N = \cup_{i=0}^N \Sigma^i$, we have that $|W_N| = \sum_{i=0}^N 4p = \frac{(4p)^{N+1}-1}{4p-1}$.

Similarly as before, the number of variables is $O(|W_N|^2)$. Since the matrix C is Hermitian, it is still sufficient to keep track of the upper triangular entries, therefore

more precisely the number of variables is still $\frac{1}{2}|W_N|(|W_N|+1)$.

We can similarly compute the number of constraints by counting up the number of each type of constraints (1) - (5) and adding them all up. Table 3.6 gives the count for each type of constraints.

Table 3.6: Number of Constraints for SDP 3.5

Constraint	Factor 1	Factor 2	Multiplicity	Total
(1)	1	1	1	1
(2)	$4p$	$\|W_{N-1}\|^2$	1	$4p\|W_{N-1}\|^2$
(3)	$(2p)^2$	$\|W_{N-1}\|^2$	1	$4p^2\|W_{N-1}\|^2$
(4)	p	$\|W_{N-\frac{p-1}{2}}\|^2$	4	$4p\|W_{N-\frac{p-1}{2}}\|^2$
(5)	p	$\|W_{N-\frac{p-1}{2}}\|^2$	4	$4p\|W_{N-\frac{p-1}{2}}\|^2$

We can add up all the constraints to get the following for the total number of constraints.

$$\begin{aligned}\#\text{ Constraints} &= (4p+4p^2)|W_{N-1}|^2 + 8p|W_{N-\frac{p-1}{2}}|^2 \\ &= (4p+4p^2)\left(\frac{(4p)^N-1}{4p-1}\right)^2 + 8p\left(\frac{(4p)^{N-\frac{p-3}{2}}-1}{4p-1}\right)\end{aligned}$$

Table 3.7 gives numerical values for the size of this SDP for the first few values N for $p=3$ and compares these values to the size of the corresponding projector relaxations for $p=3$.

Table 3.7: Sizes of Unitary and Projector Relaxations for the Mod P Game

p	N	# Variables (U)	# Constraints (U)	# Variables (P)	# Constraints (P)
3	1	91	72	171	249
3	2	12,403	12,168	58,653	87,837
3	3	1,777,555	1,774,728	19,062,225	28,590,765

As shown in table 3.7, the unitary relaxation is significantly smaller than its respective projector relaxation. This is primarily due to its smaller alphabet size. Another interesting observation is that the unitary relaxation has less constraints than it has variables.

Chapter 4

Numerical Results for the SDP Relaxations

In this chapter we give our implementation and numerical results for the semidefinite relaxation programs described in chapter 3. Our numerical work in this chapter makes use of several optimization packages for solving SDPs including Mosek [1] and CVXPY [8]. We also use several high level SDP modeling frameworks including primarily YALMIP [12] for MATLAB and PICOS (available at picos.zib.de) for Python. All of our code for this section is written in Python, but we also make use of the MATLAB Engine API for Python to execute MATLAB code from Python.

4.1 Implementation

In the practice of writing semidefinite programs, the process of systematically laying out the variables and constraints is called modeling. In the following sections we describe our approach for constructing an efficient and intuitive model for our SDP relaxations. In the first part, we present a function which takes as input a value for p and N and outputs a model for the projector SDP relaxation of the Mod P game for the specified value of p and level N.

4.1.1 Mapping Symbols to Index Numbers

In order to make our SDP model readable and correspond more directly to the way it is given in SDP 3.2, we begin by creating a function $f : W_N \to [|W_N|]$ which maps a symbol $U \in W_N$ to a numerical index into our SDP matrix variable. We use the symbol X to denote our $|W_N| \times |W_N|$ matrix variable of inner products for our SDP model, where $X[i, j]$ is the entry at the i^{th} row and j^{th} column. Thus, the entry $C_{U,V}$ from SDP 3.2 corresponds to the actual entry $X[f(U), f(V)]$. Note that since Python is 0-indexed, the entries of the matrix range from $0 \leq i, j \leq |W_N| - 1$. MATLAB on the other hand is 1-indexed, and simply requires a shift $f'(U) \leftarrow f(U) + 1$.

We create the function f by defining an intuitive strict lexicographic ordering of the symbols in W_N according to the following rules, where for a string $U \in W_N$, we let $|U|$ to be the length of the string (recall that $|U| \leq N$ for any $U \in W_N$).

1. $\mathbb{1}$ is the first element.

2. For strings of length 1, $A^*_* < B^*_*$, $A^*_s < A^*_{s'}$ for $s < s'$, $B^*_t < B^*_{t'}$ for $t < t'$, and finally $A^a_s < A^{a'}_s$ for $a < a'$ and similarly $B^b_t < B^{b'}_t$ for $b < b'$.

3. For strings $U, V \in W_N$ where $|U| < |V|$, $U < V$.

4. For strings $U, V \in W_N$ where $U \neq V$ and $|U| = |V|$, then $U < V$ if $U[i] < V[i]$ where i is the first index from left to right where $U[i] \neq V[i]$.

As a simple example, these rules enforce that $A_0^0 B_0^1 < A_0^0 B_1^0$. It is not difficult to see that the above four rules define a strict ordering on all the entries of W_N. We then let $f(U)$ be the rank of U in this ordering. Therefore we have that $f(\mathbb{1}) = 0$, $f(A_0^0) = 1$, $f(A_0^1) = 2$ etc. For a fixed value of p, note that the value of $f(U)$ is independent of the level N since it only depends on $|U|$.

In our modeling program, f can be defined as a simple recursive function.Note that since f is a one to one mapping from symbol to index, it is an invertible function. Hence, given an index j, there exists a unique U such that $f(U) = j$. We implement the inverse map $g = f^{-1}$ in our work, however the pseudocode for g is omitted. The inverse map g allows to iterate through all symbols in W_N by simply iterating through the numbers $0 \leq j \leq |W_N|$ and calling $g(i)$ to get the corresponding symbol.

High level convex optimization modeling frameworks such as PICOS for Python and YALMIP for MATLAB have convenient and intuitive ways of defining SDP variables and constraints. Using our symbol to index mapping function f and its inverse g, we can directly encode the SDP constraints in 3.2. An entry $C_{U,V}$ corresponds to $X[f(U), f(V)]$ in our model. And iterating through $U \in W_N$ can be done efficiently by iterating through the numbers from 0 to $|W_N|$ and using our inverse map g. We can therefore encode a constraint such as $C_{U,V} = C_{U',V'}$ by directly adding $X[f(U), f(V)] = X[f(U'), f(V')]$ in our model.

4.1.2 The Unitary Relaxation

The modeling process for the unitary relaxation was not too different from the projector version. The first major difference is that the symbol to index mapping function is assigned based on a different ordering of the symbols. The basic idea for the f and $g = f^{-1}$ remains the same. We assign a total ordering on the symbols in W_m using the same rules as before with the addition that for any symbol A_s, we let $A_s < A_s^\dagger$.

The constraints for the unitary relaxation can also be modeled in a straightforward fashion by simply iterating through all symbols in W_m using g and adding each specific type of constraint.

4.2 Results

For the projector relaxation, we were able to run only the first level of the hierarchy, obtaining values that match the Bavarian-Shor bound from theorem 2.3.1. Table 4.1 shows some numerical values of the implementation of the projector relaxation, SDP 3.2. The numerical value shown is the best result chosen among numerical values obtained using five different modeling and solver settings. The five settings used to solve the SDP are (1) PICOS (with solver Mosek), (2) Fusion (with solver Mosek), (3) YALMIP (with solver Mosek), (4) CVXPY (with solver SCS) and (5) CVXPY (with solver CVXOPT). We label the chosen best value in the table by the relevant number which indicates the solver setting which achieves this value. The analytic Bavarian-Shor bound given in the table has been truncated to match the precision level of the numerical value.

Table 4.1: Level 1 Numerical Results for the Projector Relaxation of the Mod P Game

	Mod 2	Mod 3	Mod 4
Bavarian-Shor	0.85355339059	0.718233512793	0.625
SDP 3.2	0.85355339121 (1)	0.718233512793 (2,3)	0.6767766952 (1,2,3)
	Mod 5	Mod 7	Mod 8
Bavarian-Shor	0.557770876399	0.46682669115	0.434359216769114
SDP 3.2	0.557770876393 (1)	0.46682669115 (1,2,3)	0.51516504293 (2,3)
	Mod 9		
Bavarian-Shor	0.407407407407		
SDP 3.2	0.461633393153 (1,2,3)		

Perhaps a most interesting observation from the results in table 4.1 is that for values of p of 4, 8 and 9, the Bavarian-Shor bound gives a noticeably tighter bound

than the first level NPA hierarchy. This suggests that for a prime power value of p, the Bavarian-Shor bound is in fact better than the NPA first level hierarchy, whereas for a prime p, the Bavarian-Shor bound yields the same upper value as the first level NPA hierarchy.

So far, we have been unable to run higher levels ($N > 1$) of the projector relaxation. Even running on rather large powerful machines, the solvers have not been successful in finding the optimal value for higher levels. The most crucial constraint for the SDP is memory. Simply writing out the variables and constraints could take hundreds of Gigabytes of memory. We leave as future work running higher levels of this SDP on more powerful and memory optimized machines.

For the unitary relaxation, we obtain results for the first and second levels of SDP 3.5. These values were obtianed using only YALMIP (with solver Mosek). Table 4.2 gives the values for both the first and second levels.

Table 4.2: Mod 3 Level 1 and 2 Numerical Results for the Unitary Relaxation of The Mod P Game

	Mod 3 Level 1	Mod 3 Level 2
Bavarian-Shor	0.718233512793	0.718233512793
SDP 3.5	0.718233512532	0.718231009506

As would be expected, the first level SDP value essentially matches the Bavarian-Shor bound. We believe the mismatch in the last few decimal places of the value comes from numerical precision issues with the solver. The second level value is non-trivially lower than the first.

For similar resource limitation reasons, we have been unable to run the unitary SDP beyond $N = 2$. For $p = 3$, we conjecture that obtaining the value for the third level would yield a method for guessing the closed form tight upper bound for the winning probability. We leave as future work to obtain the third level value as well as the positive semidefinite matrix which achieves this value.

Appendix A

Semidefinite Programming

Semidefinite Programming is a method of convex optimization of a linear objective function over the cone of positive semidefinite matrices.

The variables of a semidefinite program are the entries of a matrix. Let X be an $n \times n$ matrix and $x_{i,j}$ be the (i,j)-th entry of X. The variables of a semidefinite program are then the $x_{i,j}$. The objective function can therefore be any linear function of the $x_{i,j}$. The constraints of the optimization problem are inequalities and equalities involving linear functions of the $x_{i,j}$ as well. A required constraint is that the matrix X formed by the variables be positive semidefinite, denoted by $X \succeq 0$. This requirement is equivalent to having all eigenvalues of X be non-negative. Thus, the general form of a semidefinite program is as follows.

$$\begin{aligned} \text{Maximize: } & \langle C, X \rangle \\ \text{Subject To: } & X \succeq 0 \\ & \langle A_k, X \rangle \leq b_k \qquad \text{for } 1 \leq k \leq m \end{aligned}$$

where C and A_K are $n \times n$ matrices of coefficients and $\langle \cdot, \cdot \rangle$ corresponds to the Frobenius inner product. Thus this gives a general maximization semidefinite program in n dimensions and m constraints. Of course the objective can easily be formed as a minimization as well.

The intuitive significance of semidefinite programs can be seen by thinking about

the Cholesky decomposition of positive semidefinite matrices.

Lemma A.0.1 (Cholesky decomposition)**.** *Given a real (or complex) $n \times n$ positive semidefinite matrix X, there is an efficient way to get n vectors $y_1, y_2, \cdots, y_n \in \mathbb{R}^n$ (or $\mathbb{C}^n$) such that $X_{i,j} = \langle y_i, y_j \rangle$.*

In the above lemma, $\langle \cdot, \cdot \rangle$ refers to the inner product of the vectors. Using this interpretation, a semidefinite program can equivalently be phrased as an optimization over vectors $y_1, y_2, \cdots, y_n$ as follows.

$$\begin{aligned} \text{Maximize: } & \sum_{i,j} c_{i,j} \langle y_i, y_j \rangle \\ \text{Subject To: } & \sum_{i,j} a^k_{i,j} \langle y_i, y_j \rangle \leq b_k \qquad \text{for } 1 \leq k \leq m \end{aligned}$$

where $c_{i,j}$, $a^k_{i,j}$ and b_k are real (or complex) coefficients. Note that in the complex case, it is necessary that the objective function be a real function of the inner products. Similarly, a constraint only makes sense if both sides of the inequality are real valued functions - except in the case of an equality constraint.

In general, there are efficient, polynomial time methods for solving a semidefinite program. The most popular methods use a variation of the interior point method, much like solving linear programs.

For a more detailed approach to semidefinite programming as well as general optimization theory, we recommend "Convex Optimization" by Boyd and Vandenberghe [4]. This book has been instrumental in helping design and understand the semidefinite programs used in this thesis.

Bibliography

[1] MOSEK ApS. *The MOSEK optimization toolbox for MATLAB manual. Version 7.1 (Revision 28).*, 2015.

[2] Mohammad Bavarian and Peter W. Shor. Information causality, szemerdi-trotter and algebraic variants of chsh. *Proceedings of the 2015 Conference on Innovations in Theoretical Computer Science - ITCS '15*, 2015.

[3] John S Bell. On the einstein podolsky rosen paradox, 1964.

[4] Stephen Boyd and Lieven Vandenberghe. *Convex optimization*. Cambridge University Press, 2004.

[5] H. Buhrman and S. Massar. Causality and tsirelson's bounds. *Physical Review A*, 72(5), Mar 2005.

[6] John F Clauser, Michael A Horne, Abner Shimony, and Richard A Holt. Proposed experiment to test local hidden-variable theories. *Physical review letters*, 23(15):880, 1969.

[7] Matthew Coudron and Thomas Vidick. Interactive proofs with approximately commuting provers. *Automata, Languages, and Programming Lecture Notes in Computer Science*, page 355366, 2015.

[8] Steven Diamond and Stephen Boyd. CVXPY: A Python-embedded modeling language for convex optimization. *Journal of Machine Learning Research*, 17(83):1–5, 2016.

[9] Julia Kempe, Oded Regev, and Ben Toner. Unique games with entangled provers are easy. *SIAM Journal on Computing*, 39(7):32073229, 2010.

[10] Subhash Khot. On the power of unique 2-prover 1-round games. In *Proceedings of the thiry-fourth annual ACM symposium on Theory of computing*, pages 767–775. ACM, 2002.

[11] Subhash Khot and Nisheeth K Vishnoi. On the unique games conjecture. In *Annual Symposium on Foundations of Computer Science*, volume 46, page 3. IEEE COMPUTER SOCIETY PRESS, 2005.

[12] J. Löfberg. Yalmip : A toolbox for modeling and optimization in matlab. In *In Proceedings of the CACSD Conference*, Taipei, Taiwan, 2004.

[13] Miguel Navascus, Stefano Pironio, and Antonio Acn. A convergent hierarchy of semidefinite programs characterizing the set of quantum correlations. *New Journal of Physics*, 10(7):073013, Aug 2008.

[14] Marcin Pawowski, Tomasz Paterek, Dagomir Kaszlikowski, Valerio Scarani, Andreas Winter, and Marek ukowski. Information causality as a physical principle. *Nature*, 461(7267):11011104, 2009.

[15] Boris S Tsirel'son. Quantum analogues of the bell inequalities. the case of two spatially separated domains. *Journal of Soviet Mathematics*, 36(4):557–570, 1987.

[16] Thomas Vidick. Tsirelson's bound, Oct 2012.

[17] Stephanie Wehner. Tsirelson bounds for generalized clauser-horne-shimony-holt inequalities. *Physical Review A*, 73(2), 2006.