QUANTIFYING QUANTUM NONLOCALITY

Thesis by

Benjamin Francis Toner

In Partial Fulfillment of the Requirements
for the Degree of
Doctor of Philosophy

California Institute of Technology
Pasadena, California

2007
(Defended June 19, 2006)

Copyright notice:

Chapter 2 is taken from [1], Chapter 4 is taken from [2], and Chapter 5 is taken from [3]. These works are copyrighted by the American Physical Society (APS), used with permission.

The remaining material is

to my parents,

Chris and Mark Toner

Acknowledgements

First and foremost, I thank my advisor, John Preskill, for his guidance and support. The breadth of his knowledge has resulted in many insights, often in innocuous questions at the end of talks. I also owe thanks to Dave Bacon, who acted as a secondary advisor during my first years here at Caltech. I have been a teaching assistant for Alexei Kitaev and Jeff Kimble, and I have learned from them both.

John Preskill and Alexei Kitaev served on both my candidacy and thesis committees. I thank Jeff Kimble and Mark Wise, the other members of my candidacy committee, and Chris Umans and Alan Weinstein, the other members of my thesis committee.

Many people have been part of the Institute for Quantum Information during my time here. I thank Anura Abeyesinghe, Michael Adams, Charlene Ahn, Panos Aliferis, Robin Blume-Kohout, Sougato Bose, Sergey Bravyi, Andrew Childs, John Cortese, Sumit Daftuar, Andrew Doherty, Kovid Goyal, Sean Hallgren, Jim Harrington, Patrick Hayden, Hui Khoon Ng, Andrew Landahl, Debbie Leung, Hideo Mabuchi, Carlos Mochon, Ashwin Nayak, Michael Nielsen, Stefano Pironio, Ben Rahn, Robert Raussendorf, Leonard Schulman, Yaoyun Shi, Federico Spedalieri, Graeme Smith, Steven van Enk, Greg Ver Steeg, Frank Verstraete, Guifre Vidal, Pawel Wocjan, Jon Yard, and Michael Zwolak.

For administrative support I thank Donna Driscoll, Ann Harvey, and Carol Silberstein.

My parents, Mark and Chris, my siblings, Tim and Emily, and my good friend, Quynh-Nhu Nguyen, have been fantastic in encouraging me while overseas and I thank them for all they have done.

More specifically, I thank Antonio Acín and Nicolas Gisin for encouraging Chapter 3 and for many helpful suggestions; John Preskill, Andrew Doherty, Patrick Hayden, Andre Methot, Carlos Mochon, and Michael Steiner for useful suggestions about Chapter 5; Richard Cleve, Michael Nielsen John Preskill, Graeme Smith, Frank Verstraete, and John Watrous for useful suggestions concerning Chapter 7; Richard Cleve and Wim van Dam for useful discussions about Chapter 8; and Steven Finch for providing me with Refs. [4, 5].

My thesis work was supported in part by the National Science Foundation under grants EIA-0086038 and PHY-0456720, and by the Army Research Office under Grant No. W911NF-05-1-0294, through the Institute for Quantum Information.

Joint work

Chapter 2 is joint with Antonio Acín and Nicolas Gisin. Chapters 4 and 5 are joint with Dave Bacon. Chapter 6 is joint with Oded Regev. Chapter 8 is joint with Frank Verstraete.

Abstract

Quantum mechanics is nonlocal, meaning it cannot be described by any classical local hidden variable model. In this thesis we study two aspects of quantum nonlocality.

Part I addresses the question of what classical resources are required to simulate nonlocal quantum correlations. We start by constructing new local models for noisy entangled quantum states. These constructions exploit the connection between nonlocality and Grothendieck's inequality, first noticed by Tsirelson. Next, we consider local models augmented by a limited amount of classical communication. After generalizing Bell inequalities to this setting, we show that (i) one bit of communication is sufficient to simulate the correlations of projective measurements on a maximally entangled state of two qubits, and (ii) five bits of communication are sufficient to simulate the joint correlation of two-outcome measurements on any bipartite quantum state. The latter result can be interpreted as a stronger (constrained) version of Grothendieck's inequality.

In part II, we investigate the monogamy of nonlocal correlations. In a setting where three parties, A, B, and C, share an entangled quantum state of arbitrary dimension, we: (i) bound the trade-off between AB's and AC's violation of the CHSH inequality, obtaining an intriguing generalization of Tsirelson's bound, and (ii) demonstrate that forcing B and C to be classically correlated prevents A and B from violating certain Bell inequalities. We study not only correlations that arise within quantum theory, but also arbitrary correlations that do not allow signaling between separate groups of parties. These results are based on new techniques for obtaining Tsirelson bounds, or bounds on the quantum value of a Bell inequality, and have applications to interactive proof systems and cryptography.

Contents

Chapter 1

Introduction

1.1 Motivation

In 1935, Einstein, Podolsky, and Rosen suggested that quantum theory might emerge from a deterministic local theory, by averaging over the values of some *hidden variables*, or properties of the system inaccessible to experiment [6]. Some thirty years later, Bell proved that this is impossible: Any hidden variable model for quantum theory must be nonlocal, in a manner I shall make precise below [7]. Bell's conclusion has since been validated by a large number of experiments [8, 9, 10, 11, 12].

More recently, theoretical research into quantum algorithms [13], quantum communication complexity [14], and quantum cryptography [15] has shown that quantum devices are more powerful than their classical counterparts. Indeed, the goal of the burgeoning field of quantum information theory [16] is to obtain an information-theoretic understanding of the power of quantum resources. Here I apply this perspective to one quantum resource: nonlocal correlations. The aim is to go beyond the conclusion that quantum theory is nonlocal, instead answering the question of just *how* nonlocal it is.

1.2 Nonlocal correlations

1.2.1 An example

We start with an example. Alice and Bob, who will star in this thesis, each have a machine. Each machine has a switch, which can be in one of two positions (labeled 0 and 1), together with two lights, a red one and a green one. Every second, exactly one of the lights flashes.

After running some experiments, Alice deduces the following:

1. No matter which position her switch is in, the red light will flash with probability 1/2, otherwise the green light will flash.

In other words, Alice's machine behaves as a random number generator. Bob notices the same thing:

2. No matter which position his switch is in, the red light will flash with probability 1/2, otherwise the green light will flash.

But when they look at each other's machines, they observe that

3. If Alice's switch and Bob's switch are both in position 1, then the color of Alice's light is always different from the color of Bob's light; otherwise their lights are always the same color.

Furthermore, this is true no matter how far apart the machines are. So, although the colors of the lights on Alice and Bob's machines are locally random, they are correlated in a way that depends on how their switches are set. After providing the relevant definitions, we'll prove below that these correlations are nonlocal, meaning they are incompatible with any classical local theory. In fact, the machines just described are together known as a nonlocal box [17, 18]. We make three observations:

1. *The switches are essential.* Suppose the switches on Alice and Bob's machines are stuck in position 1. Then Alice and Bob's machines are just a correlated random source, a classical resource.

2. *The machines can be realized with instantaneous communication and a correlated random source.* Suppose that, by some means, the position of Alice's switch is (instantaneously) communicated to Bob's machine. Then this is sufficient to reproduce their behavior. Of course, instantaneous communication is unphysical.

3. *The machines cannot be used to communicate.* Suppose Alice is in Amsterdam and Bob in Melbourne. Then there is no way for Alice to send a message to Bob using the machines. No matter how she sets her switch, the data Bob obtains from his machine is just a sequence of random bits. It is only when Alice and Bob meet up and compare their data that they notice something nonclassical is going on. We term such correlations *no-signaling.* No-signaling correlations are a weaker information-processing resource than communication, but a stronger resource than a classical correlated random source.

The operation of the machines can be summarized by specifying the *conditional joint probability distribution* of Alice and Bob's results. Assume Alice sets her switch in position $i \in \{0, 1\}$ and Bob sets his in position $j \in \{0, 1\}$. Label the output of Alice's machine (the color of the light) $a \in \{0, 1\}$, and the output of Bob's $b \in \{0, 1\}$. Then the probability that Alice observes outcome a and Bob observes outcome b is

$$\Pr(a, b|i, j) = \frac{1}{2}\,[a \oplus b = i \wedge j], \tag{1.1}$$

where $[t] = 1$ if the clause t is true and 0 otherwise.

We shall be concerned with the following question: Which conditional joint probability distributions are accessible in classical, quantum, and no-signaling theories? For example, we shall observe by the end of this chapter that the machine described above is unphysical: The distribution Eq. (1.1) is not realizable, even with quantum resources.

1.2.2 Notation and definitions

We restrict attention to scenarios with two parties, Alice and Bob. In a *measurement scenario*, each party selects one of M measurements (labeled $0, 1, \dots, M-1$) and then outputs one of K different outcomes (labeled $0, 1, \dots, K-1$). Mostly we shall be interested in the case $K = 2$. As above, we label Alice's measurement i, and Bob's j. Alice's output is labeled a; Bob's b.

A local hidden variable (LHV) model for a measurement scenario is defined as follows:

Definition 1.2.1 (LHV model)**.** An LHV model for a (bipartite) measurement scenario is defined by (i) a set Λ and a probability distribution q over Λ, (ii) a function $A : \Lambda \times \mathbb{Z}_M \to \mathbb{Z}_K$, and (iii) a function $B : \Lambda \times \mathbb{Z}_M \to \mathbb{Z}_K$. We write the LHV model as a protocol:

Protocol 1.2.2. (Random Variables) Alice and Bob share a variable $\lambda \in \Lambda$, chosen according to the distribution q.

(Alice) Alice outputs $a = A(\lambda, i)$.

(Bob) Bob outputs $b = B(\lambda, j)$.

We note that this definition is completely general: Any unshared randomness can be replaced by shared randomness on which the one party does not act. To calculate the resulting conditional probability distribution, we average over λ:

$$\Pr(a, b|i, j) = \int d\lambda\, q(\lambda)[a = A(\lambda, i)][b = B(\lambda, j)]. \tag{1.2}$$

If there exists an LHV model reproducing some correlations $p(a, b|i, j)$, we say that the correlations are *local.* Otherwise they are *nonlocal.*

Definition 1.2.3 (No-signaling conditional probability distribution)**.** A conditional probability distribution is no-signaling if each party's marginal distribution is independent of the other's choice of input. More formally, $p(a, b|i, j)$ is no-signaling if

$$\Pr(a|i, j) = \sum_b p(a, b|i, j) \tag{1.3}$$

is independent of j for all a and i and

$$\Pr(b|i, j) = \sum_a p(a, b|i, j) \tag{1.4}$$

is independent of i for all b and j.

Nonlocal correlations in quantum theory arise from making local measurements on entangled quantum states. We assume that Alice and Bob share a mixed quantum state ρ with support on $\mathcal{H}_{\mathcal{A}} \otimes \mathcal{H}_{\mathcal{B}}$, where $\mathcal{H}_{\mathcal{A}}$ ($\mathcal{H}_{\mathcal{B}}$) is the local Hilbert space of Alice's (Bob's) system. The dimensions of the local spaces are denoted $d_A = \dim \mathcal{H}_{\mathcal{A}}$, $d_B = \dim \mathcal{H}_{\mathcal{B}}$. By extending the smaller of $\mathcal{H}_{\mathcal{A}}$ and $\mathcal{H}_{\mathcal{B}}$ we can assume that the local spaces have the same dimension $d = \max(d_A, d_B)$. The operator $\mathbb{1}_d$ is the identity operator operating on a space of dimension d. Where it is clear from the context, we omit the subscript indicating the dimension. The most general measurement possible in quantum theory is termed a postive operator-valued measure (POVM). A K-outcome POVM $\mathcal{M}$ on $\mathcal{H}_A$ is defined by K positive Hermitian operators $\mathbf{A}_0$, $\mathbf{A}_1$, ..., $\mathbf{A}_{K-1}$ such that $\sum_k \mathbf{A}_k = \mathbb{1}$. The elements $\mathbf{A}_k$ are termed *effects*.

Definition 1.2.4 (quantum model)**.** A quantum model for a (bipartite) measurement scenario is defined by (i) a state ρ on $\mathcal{H}_A \otimes \mathcal{H}_B$; (ii) a set of M POVMs $\mathcal{M}_i$ on $\mathcal{H}_A$ with effects $\mathbf{A}_i^a$; and (iii) a set of M POVMs $\mathcal{N}_j$ on $\mathcal{H}_B$ with effects $\mathbf{B}_j^b$. We write the quantum model as a protocol:

Protocol 1.2.5. (Preparation) Alice and Bob share the state ρ.

(Alice) Alice measures the POVM $\mathcal{M}_i$ on ρ, outputting her result a.

(Bob) Bob measures the POVM $\mathcal{N}_j$ on ρ, outputting his result b.

This results in the conditional probability distribution

$$\Pr(a, b|i, j) = \operatorname{tr}\left(\mathbf{A}_i^a \otimes \mathbf{B}_j^b \rho\right). \tag{1.5}$$

We say that a conditional probability distribution $p(a, b|i, j)$ is *realizable with quantum resources* if there is a quantum model for $p(a, b|i, j)$.

Specializing to two-outcome measurements, we define the *joint correlation*

$$\langle \alpha_i \beta_j \rangle = \sum_{a,b=0}^{1} (-1)^{a+b} \Pr(a, b|i, j). \tag{1.6}$$

We define the *observable* corresponding to a two-outcome measurement $\mathcal{M}$ as $\mathbf{A} = \mathbf{A}^0 - \mathbf{A}^1$. When Alice measures an observable $\mathbf{A}$ on ρ, we label the outcome $\alpha \in \{-1, +1\}$. We similarly label Bob's outcome β. Denote the set of possible observables by $\mathcal{O}_A$ for Alice and $\mathcal{O}_B$ for Bob. Then the joint correlation of Alice and Bob's measurement results is given by

$$\langle \alpha\beta \rangle_{\mathrm{QM}} = \operatorname{tr}\left(\mathbf{A} \otimes \mathbf{B}\, \rho\right). \tag{1.7}$$

Note that it is conventional to suppress the dependence of the left-hand side on $\mathbf{A}$ and $\mathbf{B}$. The *marginal probabilities* are given by

$$\langle\alpha\rangle_{\mathrm{QM}} = \mathrm{tr}\,(\mathbf{A}\otimes\mathbb{1}_B\,\rho)\,, \tag{1.8}$$

$$\langle\beta\rangle_{\mathrm{QM}} = \mathrm{tr}\,(\mathbb{1}_A\otimes\mathbf{B}\,\rho)\,. \tag{1.9}$$

Together, these three equations define the full probability distribution for two-outcome measurements on ρ.

Often the conditional probability distributions we wish to study arise from making measurements on quantum states, in which case the existence of a quantum model is trivial. A *quantum measurement scenario* for two-outcome measurements is a measurement scenario where we label the measurements by the corresponding quantum observables. For completeness, we state the following:

Definition 1.2.6 (LHV model for two-outcome observables)**.** An LHV model for a (bipartite) quantum measurement scenario is defined by (i) a set Λ and a probability distribution q over Λ, (ii) a function $A : \Lambda\times\mathcal{O}_A \to \{-1,+1\}$, and (iii) a function $B : \Lambda\times\mathcal{O}_B \to \{-1,+1\}$. We write the LHV model as a protocol:

Protocol 1.2.7. (Random Variables) Alice and Bob share a variable $\lambda\in\Lambda$, chosen according to the distribution q.

(Alice) Alice outputs $a = A(\lambda,\mathbf{A})$.

(Bob) Bob outputs $b = B(\lambda,\mathbf{B})$.

We average over λ to calculate the resulting correlations:

$$\langle\alpha\rangle_{\mathrm{LHV}} = \int d\lambda\, q(\lambda)A(\lambda,\mathbf{A}), \tag{1.10}$$

$$\langle\beta\rangle_{\mathrm{LHV}} = \int d\lambda\, q(\lambda)B(\lambda,\mathbf{B}), \tag{1.11}$$

$$\langle\alpha\beta\rangle_{\mathrm{LHV}} = \int d\lambda\, q(\lambda)A(\lambda,\mathbf{A})B(\lambda,\mathbf{B}). \tag{1.12}$$

We say that the LHV model reproduces the joint correlation on a state ρ when $\langle\alpha\beta\rangle_{\mathrm{LHV}} = \langle\alpha\beta\rangle_{\mathrm{QM}}$ for all observables $\mathbf{A}$ and $\mathbf{B}$. We say that the LHV model reproduces the full probability distribution when, in addition, $\langle\alpha\rangle_{\mathrm{LHV}} = \langle\alpha\rangle_{\mathrm{QM}}$ and $\langle\beta\rangle_{\mathrm{LHV}} = \langle\beta\rangle_{\mathrm{QM}}$.

1.2.3 Bell polytopes

Bell inequalities [7] describe necessary conditions on the probabilities $\Pr(a,b|i,j)$, which must be satisfied if these probabilities are to be produced by an LHV theory. When a set of these conditions is also sufficient, we say that we have a complete set of Bell inequalities. The construction of complete sets of Bell inequalities is an exercise in convex geometry, which we briefly sketch. See Ref. [19] for more details.

Consider a deterministic protocol, i.e., one in which no randomness, shared or otherwise, is used. Each party's output can only depend on their local measurement. Such a protocol is characterized by the two functions A and B in Def. 1.2.1, which describe the outcomes of the two parties' measurements: If A selects measurement i, she outputs $A(i)$ and if B selects measurement j, he outputs $B(j)$. The probabilities for the scenario are then $p(a,b|i,j) = [a = A(i)][b = B(j)]$.

By listing the components, we may view the probabilities $p(a,b|i,j)$ as vectors $\vec{p}$ in $\mathbb{R}^D$ with $D = M^2(K^2 - 1)$ (there is a normalization constraint $\sum_{a,b} p(a,b|i,j) = 1$). To each pair of functions $\{A, B\}$, there corresponds a deterministic protocol, so the set of all deterministic protocols is a finite collection of such vectors $\{\vec{d}_\zeta | \zeta = 1, ..., K^{2M}\}$.

Now consider the effect of allowing randomness. For any fixed choice of the random variables $\lambda \in \Lambda$, the functions $A(\lambda, \cdot)$ and $B(\lambda, \cdot)$ are deterministic, so that the set of all possible protocols that use randomness is described by a convex sum of the deterministic protocols

$$\vec{p} = \sum_\zeta \lambda_\zeta \vec{d}_\zeta, \quad \sum_\zeta \lambda_\zeta = 1, \quad \lambda_\zeta \geq 0. \tag{1.13}$$

The set of all protocols therefore corresponds to a region Ω_{MK} in $\mathbb{R}^D$, which is a polytope because there is a finite number of extreme vectors $\vec{d}_\zeta$ [20]. This permits an alternative description: Instead of describing the polytope Ω_{MK} as the convex combination of a finite set of extreme points, we can describe it by specifying a complete (finite) set of facet inequalities. A facet inequality is a pair $\{\vec{f}, c\}$ that defines a half-space of $\mathbb{R}^D$ via the inequality $\vec{f} \cdot \vec{p} \leq c$. Complete sets of facet inequalities $\vec{f}_\eta, c_\eta$ are satisfied if and only if $\vec{p}$ is in Ω_{MK}:

$$\vec{p} \in \Omega_{MK} \text{ iff } \vec{f}_\eta \cdot \vec{p} \leq c_\eta, \ \forall \eta. \tag{1.14}$$

Each facet is therefore a Bell inequality and complete sets of facet inequalities are complete sets of Bell inequalities. Complete sets are known in the two-party case when $M = 2, K = 2$ [21]; $M = 3, K = 2$; $M = 2, K = 3$ [22, 23]; and also when extra symmetry constraints are imposed [24].

1.2.4 The CHSH inequality

In the simplest nontrivial case $M = 2$, $K = 2$, there is (up to symmetries) one nontrivial Bell inequality, the Clauser-Horne-Shimony-Holt (CHSH) inequality [25]. Given a conditional probability distribution $p(a,b|i,j)$, define

$$\langle \mathcal{B}_{\text{CHSH}}(p) \rangle = \langle \alpha_0 \beta_0 \rangle + \langle \alpha_0 \beta_1 \rangle + \langle \alpha_1 \beta_0 \rangle - \langle \alpha_1 \beta_1 \rangle, \tag{1.15}$$

where $\langle \alpha_i \beta_j \rangle$ are the joint correlations defined by Eq. (1.6).

Theorem 1.2.8 ([21, 25])**.** *Suppose there is an LHV model for a distribution $p(a,b|i,j)$. Then*

$$\langle \mathcal{B}_{CHSH}(p)\rangle \leq 2. \tag{1.16}$$

Furthermore, if a distribution $p(a,b|i,j)$ satisfies Eq. (1.16), as well as equivalent inequalities obtained by permuting parties, measurements, and outputs, then it is obtained from some LHV model.

Proof. The proof of completeness is via facet enumeration, as described in the previous section, and we omit it. It is, however, simple to see that if there is an LHV model for $p(a,b|i,j)$, then it satisfies the inequality. Suppose the LHV model is defined by a set Λ, a probability distribution q on Λ, and functions $A, B : \Lambda \times \{0,1\} \to \{-1,+1\}$, where

$$\langle \alpha_i \beta_j \rangle = \int d\lambda q(\lambda) A(\lambda, i) B(\lambda, j). \tag{1.17}$$

Then

$$\begin{aligned}
\langle \mathcal{B}_{\mathrm{CHSH}}(p)\rangle &= \int d\lambda q(\lambda) \left[A(\lambda,0)\left(B(\lambda,0)+B(\lambda,1)\right) + A(\lambda,1)\left(B(\lambda,0)-B(\lambda,1)\right)\right] && (1.18)\\
&\leq \int d\lambda q(\lambda) \left[\left|B(\lambda,0)+B(\lambda,1)\right| + \left|B(\lambda,0)-B(\lambda,1)\right|\right]. && (1.19)
\end{aligned}$$

Now, $|B(\lambda,0)+B(\lambda,1)|$ is either 0 or 2. If it is 0, we're done; if it is 2, then $|B(\lambda,0)-B(\lambda,1)| = 0$. In either case $\langle \mathcal{B}_{\mathrm{CHSH}}(p)\rangle \leq 2$. ■

At this point we can return to the nonlocal box, the example of nonlocal correlations given at the start of this introduction. By construction, the correlations defined in Eq. (1.1),

$$\Pr(a,b|i,j) = \frac{1}{2}\left[a \oplus b = i \wedge j\right], \tag{1.20}$$

yield $\langle \alpha_0\beta_0\rangle = \langle \alpha_0\beta_1\rangle = \langle \alpha_1\beta_0\rangle = -\langle \alpha_1\beta_1\rangle = 1$, which gives $\langle \mathcal{B}_{\mathrm{CHSH}}\rangle = 4$. Since 4 is larger than 2, these correlations violate the CHSH inequality and are nonlocal. Furthermore, they are maximally nonlocal, within this measurement scenario.

The nonlocal box is a powerful resource. Consider some Boolean function $f : \{0,1\}^n \times \{0,1\}^n \to \{0,1\}$ with two n-bit strings as inputs, labeled a and b. Alice knows string a, while Bob knows string b. They want to compute $f(a,b)$ in a distributed manner, by exchanging some messages. Given a supply of nonlocal boxes (such as the machine described above), van Dam has shown that Alice and Bob can compute the function f with just one bit of communication, independent of n [26]. Thus nonlocal boxes trivialize communication complexity. In fact this conclusion is true even in the presence of some noise [27].

The nonlocal box is not realizable with quantum resources. This follows from the following, due to Tsirelson:

Theorem 1.2.9 (Tsirelson [28])**.** *Suppose Alice and Bob make local measurements on an entangled quantum state, yielding a distribution* $p(a,b|i,j)$*. Then*

$$\langle \mathcal{B}_{CHSH}(p)\rangle \leq 2\sqrt{2}. \tag{1.21}$$

Proof. Assume there is a quantum model for $p(a,b|i,j)$. By taking a purification of the shared state ρ and conditioning on the randomness, it is sufficient to consider the case where Alice and Bob share a pure state $|\psi\rangle$. Suppose Alice measures $\mathbf{A}_i$ on input i, and Bob $\mathbf{B}_j$ on input j. Then

$$\langle \alpha_i \beta_j \rangle = \langle \psi | \mathbf{A}_i \otimes \mathbf{B}_j |\psi\rangle. \tag{1.22}$$

Define $|a_i\rangle = \mathbf{A}_i \otimes \mathbb{1}_B |\psi\rangle$ and $|b_j\rangle = \mathbb{1}_A \otimes \mathbf{B}_j|\psi\rangle$. Then $\langle a_i|a_i\rangle = \langle b_j|b_j\rangle \leq 1$, since the eigenvalues of $\mathbf{A}_i$ and $\mathbf{B}_j$ are in $[-1,+1]$. This implies

$$\begin{aligned}
\langle \mathcal{B}_{\mathrm{CHSH}}(p)\rangle &= \langle a_0|\left(|b_0\rangle + |b_1\rangle\right) + \langle a_1|\left(|b_0\rangle - |b_1\rangle\right) && (1.23)\\
&\leq \||b_0\rangle + |b_1\rangle\| + \||b_0\rangle - |b_1\rangle\| && (1.24)\\
&= 2\left(\cos\theta/2 + \sin\theta/2\right) && (1.25)\\
&\leq 2\sqrt{2}, && (1.26)
\end{aligned}$$

where $\cos\theta = |\langle b_0|b_1\rangle|$. ■

1.2.5 Measures of nonlocality

One benefit of narrowing in on one aspect of quantum theory—nonlocality—is that it allows us to compare quantum theory to *stronger* theories, and not just to classical mechanics. But first we need some way of measuring the nonlocality of some distribution $p(a,b|i,j)$. We have seen one way already: the extent of violation of a Bell inequality. This measure is the right one in a number of contexts, such as the extent to which sharing entanglement allows the provers in a multi-prover interactive proof system to "cheat," an application I have explored in Ref. [29]. But other measures are more relevant in physical applications, such as experimental tests of Bell inequality violation. We explore some of these here.

Resistance to noise—Given a distribution $p(a,b|i,j)$, define the noisy distribution

$$p_\mu(a,b|i,j) = \mu\, p(a,b|i,j) + (1-\mu)q(a,b|i,j), \tag{1.27}$$

where $0 \leq \mu \leq 1$ and $q(a, b|i, j)$ is a local distribution. What is the largest value of μ such that $p_\mu(a, b|i, j)$ is local? Here the local distribution $q(a, b|i, j)$ can be some fixed distribution, such as the uniform distribution, or can be chosen adversarily. We shall explore this measure in Chapters 2 and 3.

Communication cost of simulation—What is the communication complexity of generating a distribution $p(a, b|i, j)$? In other words, suppose we augment an LHV model with a limited amount of communication, after the parties decide on which setting to measure and before they output results. How many bits of communication are required to reproduce the correlations exactly or approximately? We formalize this notion for one-way communication (which is what will be relevant for our results) as follows:

Definition 1.2.10 (LHV model with c bits of one-way classical communication)**.** An LHV model augmented by c bits of one-way classical communication for a (bipartite) measurement scenario is defined by (i) a set Λ and a probability distribution q over Λ, (ii) a function $A : \Lambda \times \mathbb{Z}_M \to \mathbb{Z}_K$, (iii) a function $m : \Lambda \times \mathbb{Z}_M \to \mathbb{Z}_c$, and (iv) a function $B : \Lambda \times \mathbb{Z}_M \times \mathbb{Z}_c \to \mathbb{Z}_K$. We write the model as a protocol:

Protocol 1.2.11. (Random Variables) Alice and Bob share a variable $\lambda \in \Lambda$, chosen according to the distribution q.

(Alice) Alice outputs $a = A(\lambda, i)$. Alice sends a message $m = m(\lambda, i)$ to Bob.

(Bob) Bob outputs $b = B(\lambda, j, m(\lambda, i))$.

We explore this measure in Chapters 4, 5, and 6.

Dectector efficiency required for loophole-free Bell inequality violation—There are a number of "loopholes" in experimental tests of Bell inequality violation that a local theory might exploit in order to simulate quantum correlations. For example, if there is time for a signal to travel from Alice's apparatus to Bob's apparatus after she chooses her measurement but before he outputs his results, then it is possible to simulate what appear to be quantum correlations with classical resources. Known as the locality loophole, this can be closed by ensuring Alice and Bob's measurement events are spacelike separated. This has been done experimentally in Refs. [9, 10, 11].

Another loophole arises if the detectors used are not 100% efficient. We should compare correlations measured with inefficient detectors to LHV models with an extra outcome, viz., "the detector failed." For a fixed value of the shared randomness, if the probability of detector failure depends on the measurement setting, then it is possible to reproduce quantum correlations via a local model, for sufficiently inefficient detectors. For a given set of correlations, what detector efficiency is required for loophole-free Bell inequality violation? This loophole has been closed experimentally in [12]. We note that the locality and detector inefficiency loopholes have not yet been closed in the same experiment.

Statistical distance to local theories—How many trials of a Bell inequality experiment should we perform to observe a contradiction with LHV models at some level of statistical significance? In this case, the relevant parameter is the relative entropy between the nonlocal distribution and the best local model. See Ref. [30].

1.3 Overview of the thesis

1.3.1 Classical models for the quantum joint correlation

In the first part of the thesis, we study classical models for the joint correlation $\langle \alpha\beta \rangle_{\mathrm{QM}} = \mathrm{tr}\,(\mathbf{A} \otimes \mathbf{B}\rho)$, resulting from performing two-outcome measurements on a quantum state ρ.

For a quantum state ρ, define

$$\rho_\mu = \mu\,\rho + (1-\mu)\frac{\mathbb{1}}{d_A d_B}. \tag{1.28}$$

In Chapter 2 we give bounds on the amount of noise required to make the correlations on ρ_μ local. These build on work of Tsirelson [31, 32], who connected Bell inequality violation with Grothendieck's inequality. For two-qubit Werner states $\rho_\mu^W = \mu\,|\psi^-\rangle\langle\psi^-| + (1-\mu)\mathbb{1}/4$, we show that there is an LHV model for projective measurements if and only if $\mu \leq 1/K_G(3)$. If we restrict the projective measurements to a plane, then there is a local model for projective meausurements on ρ_μ if and only if $\mu \leq 1/\sqrt{2}$.

In Chapter 3, we exploit this connection to construct explicit LHV models, based on (the proofs of) Krivine's upper bounds on $K_G(n)$ [33]. Among the constructions are local hidden variables models for (i) projective measurements on the qubit-qubit Werner state $\rho_\mu^W = \mu\,|\psi^-\rangle\langle\psi^-| + (1-\mu)\mathbb{1}/4$, for $\mu \leq 0.6595$, (ii) the joint correlation of projective measurements on $\rho_\mu = \mu\,\rho + (1-\mu)\mathbb{1}/4$, where ρ is an arbitrary qubit-qubit quantum state, for $\mu \leq 0.6009$; and (iii) traceless two-outcome observables on $\rho_\mu^{\max} = \mu\,|\psi_d^+\rangle\langle\psi_d^+| + (1-\mu)\mathbb{1}/d^2$, where $|\psi_d^+\rangle$ is a maximally entangled state in d dimensions, for $\mu \leq 0.5611$.

Next we turn to simulation with LHV models augmented by classical communication. In Chapter 4, we show how to generalize Bell inequalities to this setting. Suppose Alice and Bob each choose one of M two-outcome measurements and exchange one bit of information. We present the complete set of inequalities for $M = 2$, and the complete set of inequalities for the joint correlation observable for $M = 3$. The correlations produced by quantum theory satisfy both of these sets of inequalities. One bit of communication is therefore sufficient to simulate quantum correlations in both of these scenarios.

In Chapter 5, we show that one bit of communication is sufficient to simulate the correlations of projective measurements on a maximally entangled state of two qubits. We apply this result to show that certain quantum teleportation experiments, which teleport a single qubit, admit an LHV model.

Part I culminates in Chapter 6. Here we put the ideas of Chapters 2 and 3 together with those of Chapter 5 to show that five bits of communication are sufficient to simulate the joint correlation of two-outcome measurements on any bipartite quantum state. This result can be interpreted as a stronger (constrained) version of Grothendieck's inequality.

1.3.2 Monogamy of nonlocal correlations

In Part II, we describe new techniques for obtaining Tsirelson bounds, or upper bounds on the quantum value of a Bell inequality. Since quantum correlations do not allow signaling, we obtain a Tsirelson bound by maximizing over all no-signaling probability distributions. In Chapter 7 we show this maximization can be cast as a linear program.

We apply these techniques in a setting where three parties, A, B, and C, share an entangled quantum state of arbitrary dimension. We: (i) bound the trade-off between AB's and AC's violation of the CHSH inequality, and (ii) demonstrate that forcing B and C to be classically correlated prevents A and B from violating certain Bell inequalities, relevant for interactive proof systems and cryptography.

Part I

Classical Models for the Quantum Joint Correlation

Chapter 2

Local models for noisy entangled quantum states: existence

2.1 Introduction

In this chapter, we relate the nonlocal properties of noisy entangled states to Grothendieck's constant, a mathematical constant appearing in Banach space theory. For two-qubit Werner states $\rho_p^W = p\,|\psi^-\rangle\langle\psi^-| + (1-p)\mathbb{1}/4$, we show that there is a local model for projective measurements if and only if $p \leq 1/K_G(3)$, where $K_G(3)$ is Grothendieck's constant of order 3. Known bounds on $K_G(3)$ prove the existence of this model at least for $p \lesssim 0.66$, quite close to the current region of Bell violation, $p \sim 0.71$. We generalize this result to two-outcome measurements on arbitrary quantum states. This chapter is joint work with Antonio Acín and Nicolas Gisin.

From an operational point of view it is not difficult to define when a quantum state exhibits nonclassical correlations. Suppose that two parties, Alice (A) and Bob (B), share a mixed quantum state ρ with support on $\mathcal{H}_A \otimes \mathcal{H}_B$, where $\mathcal{H}_A$ ($\mathcal{H}_B$) is the local Hilbert space of A's (B's) system. Then ρ contains quantum correlations when its preparation requires a nonlocal quantum resource. Conversely, a quantum state is classically correlated, or separable, when it can be prepared using only local quantum operations and classical communication (LOCC). From this definition, due to Werner [34], it follows that a quantum state ρ is separable if it can be expressed as a mixture of product states, $\rho = \sum_{i=1}^N p_i |\psi_A^i\rangle\langle\psi_A^i| \otimes |\psi_B^i\rangle\langle\psi_B^i|$. A state that cannot be written in this form has quantum correlations and is termed entangled. But the above definition, in spite of its clear physical meaning, is somewhat impractical. Tests to distinguish separable from entangled states are complicated [35], except when $d_A = 2$ and $d_B \leq 3$ [36, 37], d_A and d_B denoting the dimensions of the local subsystems.

Violation of a Bell inequality by a quantum state is, in many situations, a witness of useful correlations [38]. In particular, Bell inequality violation is a witness of a quantum state's entanglement. Now, the question is: Are all entangled states nonlocal? For the case of pure states, the answer is

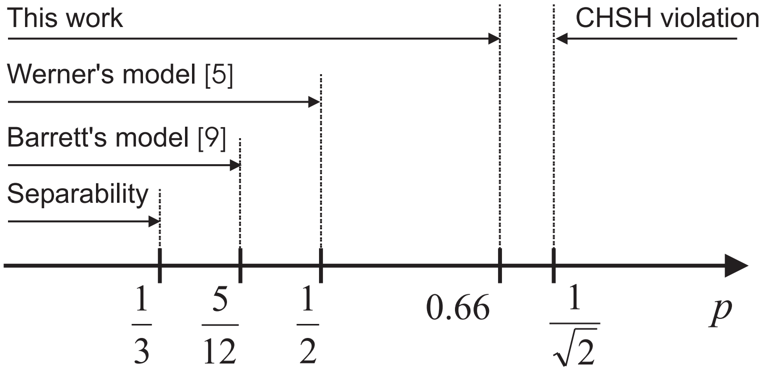


Figure 2.1: Nonlocal properties of two-qubit Werner states, ρ_p^W. Werner's local model works up to $p = 1/2$, while the CHSH inequality is violated when $p > 2^{-1/2} \sim 0.71$. Here, we prove the existence of a local model for projective measurements when $p \lesssim 0.66$.

yes [39]: All entangled pure states violate the CHSH inequality. In 1989, Werner showed that the previous result cannot be generalized to mixed states. He introduced what are now called Werner states, and gave a local hidden variable (LHV) model for measurement outcomes for some entangled states in this family [34]. Although the construction only worked for projective measurements, his result has since been extended to general measurements [40].

In spite of these partial results, it is generally extremely difficult to determine whether an entangled state has a local model or not, since (i) finding all Bell inequalities is a computationally hard problem [22, 41] and (ii) the number of possible measurement is unbounded (see however [42] for recent progress). This question remains unanswered even in the simplest case of Werner states of two qubits. These are mixtures of the singlet $|\psi^-\rangle = (|01\rangle - |10\rangle)/\sqrt{2}$ with white noise of the form

$$\rho_p^W = p\,|\psi^-\rangle\langle\psi^-| + (1-p)\frac{\mathbb{1}}{4}. \tag{2.1}$$

It is known that Werner states are separable iff $p \leq 1/3$, admit an LHV model for all measurements for $p \leq 5/12$ [40], admit an LHV model for projective measurements for $p \leq 1/2$ [34] and violate the CHSH inequality for $p > 1/\sqrt{2}$ (see Fig. 2.1). However, the critical value of p, denoted p_c^W, at which two-qubit Werner states cease to be nonlocal under projective measurements is unknown. This question is particularly relevant from an experimental point of view, since p_c^W specifies the amount of noise the singlet tolerates before losing its nonlocal properties.

In this chapter, we exploit the connection between correlation Bell inequalities and Grothendieck's constant [43], first noticed by Tsirelson [31], to prove the existence of a local model for several noisy entangled states. We first demonstrate that p_c^W is related to a generalization of this constant, namely, $p_c^W = 1/K_G(3)$, where $K_G(3)$ is Grothendieck's constant of order 3 [33]. The exact value of $K_G(3)$ is unknown, but known bounds establish that $0.6595 \leq p_c^W \leq 1/\sqrt{2}$. Thus, we close more than three-quarters of the gap between Werner's result and the known region of Bell inequality violation (see Fig. 2.1). Next, we show that if Alice (or Bob) is restricted to make measurements in a plane of the Poincaré sphere, then there is an explicit LHV model for all $p \leq 1/K_G(2) = 1/\sqrt{2}$. This improves on the bound of Larsson, who constructed an LHV model for planar measurements for $p \leq 2/\pi$ [44]. Thus, in the case of planar projective measurements, violation of the CHSH inequality completely characterizes the nonlocality of two-qubit Werner states.

In the case of *traceless two-outcome* observables, we can extend our results to mixtures of an arbitrary state ρ on $\mathbb{C}^d \otimes \mathbb{C}^d$ with the identity, of the form

$$\rho_p = p\,\rho + (1-p\,)\frac{\mathbb{1}}{d^2}. \tag{2.2}$$

Denote by $p_c(\rho)$ the maximum value of p for which there exists an LHV model for the joint correlation of traceless two-outcome observables on ρ_p, and define

$$p_c^d = \min_{\rho} p_c(\rho) \qquad p_c = \lim_{d\to\infty} p_c^d. \tag{2.3}$$

Then $p_c = 1/K_G$ where K_G is Grothendieck's constant. Again, the exact value of K_G is unknown, but known bounds imply $0.5611 \leq p_c \leq 0.5963$.

Finally, we discuss the opposite question of finding Bell inequalities better than the CHSH inequality at detecting the nonlocality of ρ_p^W, or, more generally, of Bell diagonal states.

Before proving our results, we require some notation. We write a two-outcome measurement by Alice as $\{A^+, A^-\}$, where the projectors $A^\pm$ correspond to measurement outcomes ± 1. Similarly, a two-outcome measurement made by Bob is denoted $\{B^+, B^-\}$. We define the *observable* corresponding to Alice's (Bob's) measurement as $A = A^+ - A^-$ ($B = B^+ - B^-$). An observable A is *traceless* if $\operatorname{tr} A = 0$, or equivalently $\operatorname{tr} A^- = \operatorname{tr} A^+$. The *joint correlation* of Alice and Bob's measurement results, denoted α and β, respectively, is

$$\langle \alpha\beta \rangle = \operatorname{tr}\left(A \otimes B\,\rho\right). \tag{2.4}$$

Alice's *local marginal* is specified by $\langle \alpha \rangle = \operatorname{tr}\left(A \otimes \mathbb{1}\,\rho\right)$, and Bob's by $\langle \beta \rangle = \operatorname{tr}\left(\mathbb{1} \otimes B\,\rho\right)$. Together, $\langle \alpha\beta \rangle$, $\langle \alpha \rangle$ and $\langle \beta \rangle$ define the full probability distribution for two-outcome measurements on ρ. An LHV model for the full probability distribution is one that gives the same values $\langle \alpha\beta \rangle$, $\langle \alpha \rangle$ and $\langle \beta \rangle$ as quantum theory. An LHV model for the joint correlation is one that gives the same joint correlation $\langle \alpha\beta \rangle$, but not necessarily the correct marginals. In the qubit case, the projective measurements applied by the parties are specified by the direction of their Stern-Gerlach apparatuses, given by normalized three-dimensional real vectors $\vec{a}$ and $\vec{b}$: $A = \vec{a} \cdot \vec{\sigma}$ and $B = \vec{b} \cdot \vec{\sigma}$.

2.2 Werner states

Let us first consider the case of Werner states (2.1). For projective measurements on ρ_p^W, LHV simulation of the joint correlation is sufficient to reproduce the full probability distribution. This follows from:

Lemma 2.2.1. *Suppose that there is an LHV model L that gives joint correlation $\langle \alpha\beta\rangle_L$. Then there is an LHV model L' with the same joint correlation and uniform marginals: $\langle \alpha\beta\rangle_{L'} = \langle \alpha\beta\rangle_L$, $\langle\alpha\rangle_{L'} = \langle\beta\rangle_{L'} = 0$.*

Proof. Let α and β be the outputs generated by the LHV L (dependent on the hidden variables and measurement choices). Define a new LHV L' by augmenting the hidden variables of L with an additional random bit $c \in \{-1, 1\}$. In L', Alice outputs $c\alpha$ and Bob $c\beta$. ■

Therefore, the analysis of the nonlocal properties of Werner states under projective measurements can be restricted to Bell inequalities involving only the joint correlation. Actually, this holds for any Bell diagonal state, under projective measurements, since $\mathrm{tr}_A\,\rho = \mathrm{tr}_B\,\rho = \mathbb{1}/2$ for all these states, so all projective measurements give uniform marginals. In the Bell scenarios we consider, Alice and Bob each choose from m observables, specified by $\{A_1, \ldots, A_m\}$ and $\{B_1, \ldots, B_m\}$. We can write a generic correlation Bell inequality as

$$|\sum_{i,j=1}^{m} M_{ij}\,\langle\alpha_i\beta_j\rangle| \leq 1, \tag{2.5}$$

where $M = (M_{ij})$ is an $m \times m$ matrix of real coefficients defining the Bell inequality. The matrix M is normalized such that the local bound is achieved by a deterministic local model, i.e.,

$$\max_{a_i=\pm1,\, b_j=\pm1} |\sum_{i,j=1}^{m} M_{ij}\,a_i b_j| = 1. \tag{2.6}$$

For the singlet state $\langle\alpha_i\beta_j\rangle_{\Psi^-} = -\vec{a}_i \cdot \vec{b}_j$. We obtain the maximum ratio of Bell inequality violation for the singlet state, denoted Q, by maximizing over normalized Bell inequalities, and taking the limit as the number of settings goes to infinity:

$$Q = \lim_{m\to\infty} \sup_{M_{ij}} \max_{\vec{a}_i,\,\vec{b}_j} |\sum_{i,j=1}^{m} M_{ij}\,\vec{a}_i \cdot \vec{b}_j|. \tag{2.7}$$

Since all joint correlations vanish for the maximally mixed state, it follows that the critical point at which two-qubit Werner states do not violate any Bell inequality is $p_c^W = 1/Q$.

As first noticed by Tsirelson, the previous formulation of the Bell inequality problem is closely related to the definition of Grothendieck's inequality and Grothendieck's constant, K_G (see [31] for details). Grothendieck's inequality first arose in Banach space theory, particularly in the theory of p-summing operators [45]. We shall need a refinement of his constant, which can be defined as follows [43]:

Definition 2.2.2 (Grothendieck's constant of order n)**.** For any integer $n \geq 2$, Grothendieck's constant of order n, denoted $K_G(n)$, is the smallest number with the following property: Let M be any $m \times m$ matrix for which

$$|\sum_{i,j=1}^{m} M_{ij}\, a_i b_j| \leq 1, \tag{2.8}$$

for all real numbers $a_1, \ldots, a_m, b_1, \ldots, b_m \in [-1, +1]$. Then

$$|\sum_{i,j=1}^{m} M_{ij}\, \vec{a}_i \cdot \vec{b}_j| \leq K_G(n), \tag{2.9}$$

for all unit vectors $\vec{a}_1, \ldots, \vec{a}_m, \vec{b}_1, \ldots, \vec{b}_m$ in $\mathbb{R}^n$.

Definition 2.2.3 (Grothendieck's constant)**.** Grothendieck's constant is defined as

$$K_G = \lim_{n \to \infty} K_G(n). \tag{2.10}$$

The best bounds currently known for K_G are $1.6770 \leq K_G \leq \pi/(2\log(1+\sqrt{2})) = 1.7822$ [46]. The lower bound is due to Reeds and, independently, Davies [4, 5], while the upper bound is due to Krivine [33].

It follows immediately from the first definition that the maximal Bell violation for the singlet state (2.7) is $K_G(3)$. We have therefore proved

Theorem 2.2.4. *There is an LHV model for projective measurements on the Werner state ρ_p^W if and only if $p \leq p_c^W = 1/K_G(3)$.*

It is known that $\sqrt{2} \leq K_G(3) \leq 1.5163$. The lower bound follows from the CHSH inequality; the upper bound is again due to Krivine [33]. He shows that $K_G(3) \leq \pi/(2c_3)$ where c_3 is the unique solution of

$$\frac{\sqrt{c_3}}{2} \int_0^{c_3} t^{-3/2} \sin t \, dt = 1 \tag{2.11}$$

in the interval $[0, \pi/2]$. Numerically we find that $c_3 \approx 1.0360$. This implies $K_G(3) \leq 1.5163$ and $p_c^W \geq 0.6595$. Furthermore, it turns out that an explicit LHV model emerges from Krivine's upper bound on $K_G(3)$, as we shall see in the following chapter.

Another result follows from Krivine's work:

Theorem 2.2.5. *If Alice's projective measurements are restricted to a plane in the Poincaré sphere, then there is an LHV model for ρ_p^W if and only if $p \leq 1/\sqrt{2}$.*

Proof. In this case, the vectors $\vec{a}_i$ in (2.7) are two-dimensional. Since the quantum correlation depends only on the projection of $\vec{b}_j$ onto $\vec{a}_i$, we can assume that the vectors $\vec{b}_j$ lie in the same plane. It follows that $p_c^W = 1/K_G(2)$ for planar measurements, and Krivine has shown that $K_G(2)$ is equal to $\sqrt{2}$ [33]. ■

Again Krivine's proof can be adapted to give an explicit LHV model for planar measurements, valid for $p \leq 1/\sqrt{2}$ and we shall do this in the next chapter.

2.3 Generalization to higher dimensions

It is possible to extend these results to general states of the form (2.2), if we restrict our analysis to correlation Bell inequalities of *traceless* two-outcome observables. Admittedly, this analysis is far from sufficient. Indeed, it does not allow us to determine whether the full probability distribution admits an LHV model even in the case of two-outcome measurements, since the most general Bell inequalities have terms that depend on marginal probabilities [23]. Mindful of this caveat, we now prove the existence of LHV models for the joint correlation of the states (2.2). To make the connection with Grothendieck's constant, we start with a representation of quantum correlations as dot products, first noted by Tsirelson [31]. It is sufficient to restrict to the case of pure states, since we can obtain an LHV model for a mixed state ρ by decomposing it into a convex sum of pure states, and taking a convex combination of the LHVs for those pure states.

Lemma 2.3.1. *Suppose Alice and Bob measure observables A and B on a pure quantum state $|\psi\rangle \in \mathbb{C}^d \otimes \mathbb{C}^d$. Then we can associate a real unit vector $\vec{a} \in \mathbb{R}^{2d^2}$ with A (independent of B), and a real unit vector $\vec{b} \in \mathbb{R}^{2d^2}$ with B (independent of A) such that $\langle\alpha\beta\rangle_\psi = \vec{a}\cdot\vec{b}$. Moreover, if $|\psi\rangle$ is maximally entangled, then we can assume the vectors $\vec{a}$ and $\vec{b}$ lie in $\mathbb{R}^{d^2-1}$.*

Proof. Let $|a\rangle = A \otimes \mathbb{1}_B|\psi\rangle$ and $|b\rangle = \mathbb{1}_A \otimes B|\psi\rangle$. Then $\langle\alpha\beta\rangle = \langle a|b\rangle$, $\langle a|a\rangle = \langle b|b\rangle = 1$. Denote the components of $|a\rangle$ as a_i where $i = 1, 2, \ldots, d^2$, and similarly for $|b\rangle$. We now define a $2d^2$–dimensional real vector $\vec{a} = (\operatorname{Re} a_1, \operatorname{Im} a_1, \operatorname{Re} a_2, \operatorname{Im} a_2, \ldots, \operatorname{Re} a_{d^2}, \operatorname{Im} a_{d^2})$, and similarly $\vec{b} = (\operatorname{Re} b_1, \operatorname{Im} b_1, \operatorname{Re} b_2, \operatorname{Im} b_2, \ldots, \operatorname{Re} b_{d^2}, \operatorname{Im} b_{d^2})$. Then $\vec{a}\cdot\vec{a} = \vec{b}\cdot\vec{b} = 1$ and $\langle\alpha\beta\rangle = \vec{a}\cdot\vec{b}$ (because $\langle a|b\rangle$ is real).

If $|\psi\rangle$ is maximally entangled, we can assume $|\psi\rangle = |\psi^+\rangle = 1/\sqrt{d}\sum_{i=1}^d |ii\rangle$. We calculate $\langle\alpha\beta\rangle_{\psi^+} = \operatorname{tr}_A(AB^t)/d$ where B^t is the transpose of B. Introduce a (d^2-1)–dimensional basis g_i for traceless operators on $\mathcal{H}_A$, normalized such that $\operatorname{tr}(g_i g_j) = d\delta_{ij}$. Let $A = \sum_i a_i g_i$, $B^t = \sum_i b_i g_i$, which define the vectors $\vec{a}$ and $\vec{b}$. Squaring these definitions and taking the trace gives $\sum_i a_i^2 = \sum_i b_i^2 = 1$. Finally, $\operatorname{tr}(AB^t) = d\sum_i a_i b_j$, which implies that $\langle\alpha\beta\rangle = \sum_i a_i b_i = \vec{a}\cdot\vec{b}$. ■

The converse of Lemma 2.3.1 is also true: All dot products of normalized vectors, $\vec{a}, \vec{b} \in \mathbb{R}^n$, are realized as observables on $|\psi^+\rangle$, where $n = 2\lfloor \log_2 d\rfloor + 1$. This result was derived by Tsirelson in

Ref. [31]. For the sake of completeness, we state it here without proof (see [31] for the details).

Theorem 2.3.2 (Tsirelson [31])**.** *Let $\{\hat{a}_i\}_{i=1}^m$ and $\{\hat{b}_j\}_{j=1}^m$ be sets of unit vectors in $\mathbb{R}^n$. Let $d = 2^{\lfloor n/2 \rfloor}$ and $|\Phi\rangle$ be a maximally entangled state on $\mathbb{C}^d \otimes \mathbb{C}^d$. Then there are observables $A_1 \ldots, A_m$ and $B_1 \ldots, B_m$ on $\mathbb{C}^d$ such that*

$$\begin{aligned} \langle \alpha_i \rangle &= \langle \Phi | A_i \otimes \mathbb{1} | \Phi \rangle = 0, && (2.12) \\ \langle \beta_j \rangle &= \langle \Phi | \mathbb{1} \otimes B_j | \Phi \rangle = 0, && (2.13) \\ \langle \alpha_i \beta_j \rangle &= \langle \Phi | A_i \otimes B_j | \Phi \rangle = \hat{a}_i \cdot \hat{b}_j, && (2.14) \end{aligned}$$

for all $1 \leq i, j \leq m$.

Note that in our case, the stipulation that the observables be traceless ensures that their outcomes are random on the maximally mixed state. Theorem 2.3.3 follows from Lemma 2.3.1 and Theorem 2.3.2:

Theorem 2.3.3. *Let ρ be a state on $\mathbb{C}^d \otimes \mathbb{C}^d$ and define ρ_p and p_c^d as in Eqs. (2.2,2.3). Then*

$$\frac{1}{K_G(2d^2)} \leq p_c^d \leq \frac{1}{K_G(2\lfloor \log_2 d \rfloor + 1)}. \tag{2.15}$$

In other words, there is always an LHV model for the joint correlation of traceless two-outcome observables on ρ_p for $p \leq 1/K_G(2d^2)$ and there is a state (in fact, the maximally entangled state on $\lfloor \log_2 d \rfloor$ qubits) such that the joint correlation is nonlocal for $p > 1/K_G(2\lfloor \log_2 d \rfloor + 1)$.

Corollary 2.3.4. *The threshold noise for the joint correlation of two-outcome traceless observables is $p_c = 1/K_G$.*

This follows from the previous theorem, taking the limit $d \to \infty$. The known bounds imply $0.5611 \leq p_c \leq 0.5963$. Compare this to p_s, the threshold noise at which the state ρ_p is guaranteed separable: While p_s decreases with dimension at least as $1/(1+d)$ [47], p_c approaches a constant. In the case of two-qubit systems, we can be more specific, because projective measurements are traceless and have two outcomes:

Corollary 2.3.5. *Suppose ρ is an arbitrary state on $\mathbb{C}^2 \otimes \mathbb{C}^2$. Then there is an LHV model for the joint correlation on $\rho_p = p\,\rho + (1-p)\mathbb{1}/4$ for $p \leq 1/K_G(8)$. In particular, $K_G(8) \leq 1.6641$ [33], which implies there is an LHV model for $p \leq 0.6009$.*

For maximally entangled states, marginals of traceless observables are uniform, so Lemmas 2.2.1 and 2.3.1 imply:

Theorem 2.3.6. *Let $\rho_p = p\,|\psi^+\rangle\langle\psi^+| + (1-p)\mathbb{1}/d^2$ where $|\psi^+\rangle$ is a maximally entangled state in*

$\mathbb{C}^d \otimes \mathbb{C}^d$. *Then there is an LHV for the full probability distribution arising from traceless observables for* $p \leq 1/K_G(d^2-1)$.

2.4 Bell inequalities for Werner states

Just as upper bounds on $K_G(n)$ yield LHV models, lower bounds yield Bell inequalities. The case of Werner states appears of particular interest: At present, there is no Bell inequality better than CHSH at detecting the nonlocality of ρ_p^W. This and other approaches to construct new Bell inequalities will be presented in [48]. Unfortunately, none of these inequalities could be proven to be better than CHSH. It is remarkable how difficult it is to enlarge this region of Bell violation or, equivalently, to show that $K_G(3) > K_G(2) = \sqrt{2}$. Actually, in the case of random marginal probabilities, as for Bell diagonal states under projective measurements, no improvement over the CHSH inequality can be obtained using $3 \times n$ measurements [49].

This result, however, would imply that $K_G(3) = K_G(2) = \sqrt{2}$, which seems unlikely. Actually, one can find in [46] an explicit construction with 20 settings showing that $K_G(5) \geq 10/7 > \sqrt{2}$. More recently, one of us has shown that $K_G(4) > \sqrt{2}$ as well [48].

2.5 Conclusions

In this work, we have exploited the connection between Bell correlation inequalities and Grothendieck's constants to prove the existence of LHV models for several noisy entangled states. In the case of Werner states, one can demonstrate the existence of a local model for projective measurements up to $p \sim 0.66$, close to the known region of Bell violation. Although we only proved here the existence of the LHV models, the correspondence between noise thresholds and Grothendieck's constants can also be exploited to construct the *explicit* models. Indeed, these can be extracted from (the proofs of) Krivine's upper bounds on $K_G(n)$. The details are presented in the following chapter.

Chapter 3

Local models for noisy entangled quantum states: constructions

3.1 Introduction

In the preceding chapter, I, with Acín and Gisin, related the nonlocal properties of noisy entangled states to $K_G(n)$, Grothendieck's constant of order n. In this chapter, I exploit this connection to construct explicit local hidden variables models, based on Krivine's proofs of upper bounds on $K_G(n)$.

As before, we consider the one-parameter family of states obtained by mixing an arbitrary quantum state ρ with white noise

$$\rho_p = p\,\rho + (1-p)\mathbb{1}_d/d^2, \tag{3.1}$$

where $0 \leq p \leq 1$. We list our results:

1. Qubit-qubit Werner states are mixtures of the singlet $|\psi^-\rangle = (|01\rangle - |10\rangle)/\sqrt{2}$ with white noise of the form

   $$\rho_p^W = p\,|\psi^-\rangle\langle\psi^-| + (1-p\,)\frac{\mathbb{1}}{4}. \tag{3.2}$$

   It is known that Werner states are separable iff $p \leq 1/3$, admit an LHV model for all measurements for $p \leq 5/12$ [40], admit an LHV model for projective measurements for $p \leq 1/2$ [34] and violate the CHSH inequality for $p > 1/\sqrt{2}$ (see Fig. 2.1). In the previous chapter, we showed that there is an LHV model for projective measurements iff $p \leq 1/K_G(3)$. Although $K_G(3)$ is not known exactly, it is known that $\sqrt{2} \leq K_G(3) \leq 1.5163$. The lower bound comes from the CHSH inequality; the upper bound is due to Krivine [33]. He shows that $K_G(3) \leq \pi/(2c_3)$ where c_3 is the unique solution of

   $$\frac{\sqrt{c_3}}{2}\int_0^{c_3} t^{-3/2}\sin t\,dt = 1 \tag{3.3}$$

in the interval $c_3 \in [0, \pi/2]$. Numerically we find that $c_3 \approx 1.0360$. This implies $K_G(3) \leq 1.5163$. It turns out that we can extract an explicit LHV model from Krivine's proof, valid for $p \leq 0.6595$. We present it in Section 3.5.

2. For projective measurements on ρ_p^W, but where Alice's measurements are restricted to a plane of the Poincaré sphere, there is an LHV model iff $p \leq 1/K_G(2) = 1/\sqrt{2}$. We construct this model in Section 3.7. Previously, it was known that there is an LHV model for planar measurements for $p \leq 2/\pi = 0.6366$ [44].

3. For the joint correlation of projective meausurements on $p\rho + (1-p)\mathbb{1}/4$, where ρ is an arbitrary state on $\mathbb{C}^2 \otimes \mathbb{C}^2$, there is an LHV model for $p \leq 1/K_G(8)$. Known bounds give an explicit model for $p \leq 0.6009$. We give this construction in Section 3.6.

4. For traceless observables on $p\,|\psi_d^+\rangle\langle\psi_d^+| + (1-p)\mathbb{1}/d^2$, where $|\psi_d^+\rangle$ is a maximally entangled state in $\mathbb{C}^d \otimes \mathbb{C}^d$, there is an LHV model for the full probability distribution if $p \leq 1/K_G(d^2-1)$. This generalizes result 1 (for qubit-qubit Werner states) to higher dimensional systems. In Section 3.4, we give an explicit construction that works for all d, provided $p \leq 0.5611$. For any particular value of d, we can do better: We describe this construction in Section 3.6.

5. For the joint correlation arising from traceless observables on $p\rho + (1-p)\mathbb{1}/d^2$, where ρ is an arbitrary state on $\mathbb{C}^d \otimes \mathbb{C}^d$, there is an LHV model for $p \leq 1/K_G(2d^2)$. In Section 3.4, we give an explicit construction that works for all d, provided $p \leq 0.5611$. For any particular value of d, we can do better: We describe this construction in Section 3.6.

In each case we state the LHV model first, before proving its correctness. This means we shall have to pull some constants and functions from thin air, without much motivation. We have chosen this format so that a reader who is only interested in the protocols need not wade through the proofs. At first glance, the LHV models may appear rather mysterious, and so we attempt to give some intuition where they come from. For further intuition, we refer the reader to Krivine's original paper. As these results owe much to Krivine's work, we have attempted, where possible, to keep the notation consistent with his paper.

The remaining material in the chapter is organized as follows: In Section 3.2, we present a simple LHV model that does not actually reproduce any quantum correlations, but is used as a primitive in LHV models that do. In Sections 3.3 and 3.4 we present two LHV models for the correlation of traceless observables on mixtures of the identity with an arbitrary quantum state. The model in Section 3.3 is simpler than that in Section 3.4; it is, however, valid only for $p \leq 0.4345$, whereas the model in Section 3.4 works for $p \leq 0.5611$. Next, in Section 3.5, we give an LHV for observables on qubit-qubit Werner states. In Section 3.6, we present an LHV model for traceless observables on mixtures of the identity with an arbitrary state on $\mathbb{C}^d \otimes \mathbb{C}^d$. The most important application of this result is the LHV for observables on noisy qubit-qubit states described above. The LHV model for Werner states of Section 3.5 is actually just a special case of the model in Section 3.6: We present it separately, however, because it is the case of most interest and requires only mathematical machinery that is likely to be familiar to quantum theorists. Finally, in Section 3.7 we present an LHV model for planar measurements on a Werner state.

Section 3.2 is required to understand what follows, but Section 3.3 through Section 3.7 can then be read independently of one another. We have, however, ordered the sections so that the LHV models increase in sophistication, and we do recommend they be read in order.

We write S_n for the unit sphere in $\mathbb{R}^n$. We write the complex conjugate of a vector $\vec{v} \in \mathbb{C}^n$ as $\vec{c}^*$. The function "sgn" is the sign function: $\operatorname{sgn} x = x/|x|$.

3.2 A primitive for LHV models

In this section, we present a simple protocol that doesn't itself give correct correlations for the quantum states of interest, but is used as a primitive in the models that do. Originally due to Grothendieck [45], it was presented independently by Bell [7] and others. The LHV model is as follows:

Protocol 3.2.1. (Random Variables) Alice and Bob share a unit vector $\hat{\lambda} \in \mathbb{R}^n$ chosen uniformly at random from the unit sphere.

(Alice) Alice outputs $\alpha = \operatorname{sgn}(\hat{a} \cdot \hat{\lambda})$.

(Bob) Bob outputs $\beta = \operatorname{sgn}(\hat{b} \cdot \hat{\lambda})$.

We claim:

Lemma 3.2.2. *Protocol 3.2.1 results in correlations*

$$\langle \alpha\beta \rangle = \frac{2}{\pi} \sin^{-1} \hat{a} \cdot \hat{b}. \tag{3.4}$$

Proof. Let us calculate $\Pr(\alpha = \beta)$. Introduce an azimuthal coordinate ϕ for $\hat{\lambda}$ in the plane spanned by $\hat{a}$ and $\hat{b}$, such that $\hat{a}$ has coordinate $\phi = 0$ and $\hat{b}$ has coordinate $r = \cos^{-1} x_s \cdot y_t \in [0, \pi]$. Then $\operatorname{sgn}(\hat{a} \cdot \hat{\lambda}) = 1$ for $\phi \in [-\pi/2, \pi/2]$ and -1 otherwise, while $\operatorname{sgn}(\hat{b} \cdot \hat{\lambda}) = 1$ for $\phi \in [r - \pi/2, r + \pi/2]$ and -1 otherwise. Because $\hat{\lambda}$ is distributed uniformly in $\mathbb{R}^N$, ϕ is distributed uniformly in $[0, 2\pi)$. Thus $\Pr(\alpha = \beta) = (\pi - r)/\pi$, the fraction of the interval $[0, 2\pi)$ on which $\alpha = \beta$. It follows that

$$\langle \alpha\beta \rangle = \Pr(\alpha = \beta) - \Pr(\alpha \neq \beta) = \frac{2}{\pi} \left(\frac{\pi}{2} - r \right) = \frac{2}{\pi} \sin^{-1} \hat{a} \cdot \hat{b}, \tag{3.5}$$

as required. ■

Protocol 3.2.1 works in any dimension n. Notice that we didn't really need the shared vector $\vec{\lambda}$ to be a unit vector: For any $r \in (0, \infty)$, replacing $\hat{\lambda}$ by $\vec{\lambda} = r\hat{\lambda}$ does not change the sign of $\vec{v} \cdot \vec{\lambda}$, so we can draw our shared random vectors from any set with the property that the projection of $\vec{\lambda}$ onto the unit sphere is uniform. A particularly convenient choice is to sample each coordinate of $\vec{\lambda}$ at random from a Gaussian distribution with mean 0 and standard deviation 1. This results in the following model, which is equivalent to Protocol 3.2.1:

Protocol 3.2.3. (Random Variables) Alice and Bob share a sequence $\lambda_1, \lambda_2, \ldots, \lambda_n$ of numbers where each λ_i is drawn from a normal distribution with mean 0 and standard deviation 1. We write $\vec{\lambda} = (\lambda_1, \lambda_2, \ldots, \lambda_n) \in \mathbb{R}^n$.

(Alice) Alice outputs $\alpha = \operatorname{sgn}(\hat{a} \cdot \hat{\lambda})$.

(Bob) Bob outputs $\beta = \operatorname{sgn}(\hat{b} \cdot \hat{\lambda})$.

Lemma 3.2.4. *Protocol 3.2.3 results in correlations*

$$\langle \alpha\beta \rangle = \frac{2}{\pi} \sin^{-1} \hat{a} \cdot \hat{b}. \tag{3.6}$$

We omit the proof.

3.3 LHV model based on Grothendieck's upper bound on K_G

In this section we present an LHV model for traceless observables on mixtures of an arbitrary quantum state with the identity. It is based on Grothendieck's original upper bound on what came to be known as his constant K_G and is valid for $p < 0.4345$. To proceed, let $\hat{a} \in \mathbb{R}^n$ be the vector associated with Alice's observable A by Lemma 2.3.1, and let $\hat{b} \in \mathbb{R}^n$ be the vector associated with Bob's observable B by the same lemma.

Protocol 3.3.1. (Random Variables) Alice and Bob share an integer k drawn from the probability distribution $\Pr(k) = \frac{(\pi/2)^{2k+1}}{(2k+1)!\sinh(\pi/2)}$. They also share $2k+1$ unit vectors $\hat{\lambda}_1, \hat{\lambda}_2, \ldots, \hat{\lambda}_{2k+1}$ drawn uniformly at random from the unit sphere in $\mathbb{R}^n$.

(Alice) Alice outputs $\alpha = (-1)^k \operatorname{sgn}\left[(\hat{a} \cdot \hat{\lambda}_1)(\hat{a} \cdot \hat{\lambda}_2) \cdots (\hat{a} \cdot \hat{\lambda}_{2k+1})\right]$

(Bob) Bob outputs $\beta = \operatorname{sgn}\left[(\hat{b} \cdot \hat{\lambda}_1)(\hat{b} \cdot \hat{\lambda}_2) \cdots (\hat{b} \cdot \hat{\lambda}_{2k+1})\right]$.

Theorem 3.3.2. *Protocol 3.3.1 results in correlations*

$$\langle \alpha\beta \rangle = \frac{1}{\sinh(\pi/2)} \hat{a} \cdot \hat{b}. \tag{3.7}$$

Proof. We had better start by making sure that our protocol is well defined, i.e., check that

$\sum_k \Pr(k) = 1$. This is straightforward:

$$\sum_k \Pr(k) = \frac{1}{\sinh(\pi/2)} \sum_k \frac{1}{(2k+1)!} (\pi/2)^{2k+1} = 1, \tag{3.8}$$

since the Taylor series for $\sinh x$ is $\sum_k \frac{1}{(2k+1)!} x^{2k+1}$. Thus the protocol is well defined.

We calculate

$$\begin{aligned} \langle \alpha\beta \rangle &= \sum_k (-1)^k \frac{(\pi/2)^{2k+1}}{(2k+1)! \sinh(\pi/2)} \int d\hat{\lambda}_1 \cdots \int d\hat{\lambda}_{2k+1} & (3.9) \\ & \qquad \operatorname{sgn}\left[(\hat{a}\cdot\hat{\lambda}_1)\cdots(\hat{a}\cdot\hat{\lambda}_{2k+1})\right] \operatorname{sgn}\left[(\hat{b}\cdot\hat{\lambda}_1)\cdots(\hat{b}\cdot\hat{\lambda}_{2k+1})\right]. & (3.10) \end{aligned}$$

by averaging over the shared randomness. Since the vectors $\hat{\lambda}_1, \ldots, \hat{\lambda}_n$ are chosen independently, we may break up the integral in the k'th term of the sum into the product of $2k+1$ terms:

$$\langle \alpha\beta \rangle = \sum_k (-1)^k \frac{(\pi/2)^{2k+1}}{(2k+1)! \sinh(\pi/2)} \left[\int d\hat{\lambda}_1 \operatorname{sgn}(\hat{a}\cdot\hat{\lambda}_1) \operatorname{sgn}(\hat{b}\cdot\hat{\lambda}_1)\right]^{2k+1}. \tag{3.11}$$

Fortunately, we have already evaluated this integral in the preceding section. Indeed, it the just the expression for the correlations arising from Protocol 3.2.1. Therefore, by Lemma 3.2.2 we have

$$\int d\hat{\lambda}_1 \operatorname{sgn}(\hat{a}\cdot\hat{\lambda}_1) \operatorname{sgn}(\hat{b}\cdot\hat{\lambda}_1) = \frac{2}{\pi} \sin^{-1} \hat{a}\cdot\hat{b}. \tag{3.12}$$

Substituting this into Eq. (3.11), we obtain

$$\begin{aligned} \langle \alpha\beta \rangle &= \frac{1}{\sinh(\pi/2)} \sum_k (-1)^k \frac{1}{(2k+1)!} \left[\sin^{-1} \hat{a}\cdot\hat{b}\right]^{2k+1} & (3.13) \\ &= \frac{1}{\sinh(\pi/2)} \sin(\sin^{-1} \hat{a}\cdot\hat{b}) = \frac{1}{\sinh(\pi/2)} \hat{a}\cdot\hat{b}. & (3.14) \end{aligned}$$

This completes the proof. ■

3.4 LHV model based on Krivine's upper bound on K_G

In this section we give a better LHV model for correlations of traceless observables on mixtures an arbitrary quantum state with the identity. The model in the previous section was valid for $p \leq 0.4345$; the model in this section works for $p \leq 0.5611$. It is adapted from a proof of Krivine, later simplified by Alon and Naor [41]. Let $\hat{a} \in \mathbb{R}^n$ be the vector associated with Alice's measurement in accordance with Lemma 2.3.1, and let $\hat{b} \in \mathbb{R}^n$ be the vector associated with Bob's measurement.

Let $c = \sinh^{-1} 1 = \ln(1 + \sqrt{2})$. We present a protocol that gives correlations

$$\langle \alpha\beta \rangle = \frac{2c}{\pi} \hat{a} \cdot \hat{b}. \tag{3.15}$$

To simulate these correlations, we map $\hat{a}$ and $\hat{b}$ to new vectors, $A(\hat{a})$ and $B(\hat{b})$, respectively, which live in a much larger space. We then run Protocol 3.2.1 on $A(\hat{a})$ and $B(\hat{b})$. The trick is in choosing appropriate functions A and B.

To this end, let $A, B : \mathbb{R}^n \to \bigoplus_{k=0}^{\infty} (\mathbb{R}^n)^{\otimes(2k+1)}$. The range of A and B is a direct sum of tensor products of $\mathbb{R}^n$. Of course $\bigoplus_{k=0}^{\infty} (\mathbb{R}^n)^{\otimes(2k+1)}$ is just $\mathbb{R}^\infty$ (usually denoted l_∞), the space of infinite sequences, but this particular parameterization is convenient. We write $A(\vec{v}) = \bigoplus_{k=0}^{\infty} A^{2k+1}(\vec{v})$, and term the functions $A^{2k+1}(\vec{v})$ "coordinates" of $A(\vec{v})$. Similarly $B(\vec{v}) = \bigoplus_{k=0}^{\infty} B^{2k+1}(\vec{v})$

Define

$$A^{2k+1}(\vec{v}) = (-1)^k \sqrt{\frac{c^{2k+1}}{(2k+1)!}}\, \vec{v}^{\otimes(2k+1)}; \qquad B^{2k+1}(\vec{v}) = \sqrt{\frac{c^{2k+1}}{(2k+1)!}}\, \vec{v}^{\otimes(2k+1)}, \tag{3.16}$$

where $\vec{v}^{\otimes(2k+1)}$ denotes the vector $\vec{v} \otimes \vec{v} \otimes \cdots \otimes \vec{v}$ with $2k+1$ tensor factors.

The LHV model is as follows:

Protocol 3.4.1. (Random Variables) Alice and Bob share an infinite sequence $\lambda_1, \lambda_2, \ldots$ of real numbers, where each λ_i is drawn from a normal distribution with mean 0 and standard deviation 1. We write $\vec{\lambda} = (\lambda_1, \lambda_2, \ldots) \in l_\infty$.

(Alice) Alice outputs $\alpha = \text{sgn}(A(\hat{a}) \cdot \vec{\lambda})$.

(Bob) Bob outputs $\beta = \text{sgn}(B(\hat{b}) \cdot \vec{\lambda})$.

Theorem 3.4.2. *Protocol 3.4.1 results in correlations*

$$\langle \alpha\beta \rangle = \frac{2c}{\pi} \hat{a} \cdot \hat{b} = \frac{2 \sinh^{-1} 1}{\pi} \hat{a} \cdot \hat{b}. \tag{3.17}$$

Proof. In order to apply Lemma 3.2.2, we have to check that that $A(\hat{a})$ is a unit vector whenever $\hat{a}$ is, and do the same for $B(\hat{b})$. We'll defer verification of this fact until after we calculate the correlations.

Provided $A(\hat{a})$ and $B(\hat{b})$ are unit vectors, Lemma 3.2.2 implies that Protocol 3.4.1 results in correlations

$$\langle \alpha\beta \rangle = \frac{2}{\pi} \arcsin A(\hat{a}) \cdot B(\hat{b}). \tag{3.18}$$

Now,

$$A(\hat{a}) \cdot B(\hat{b}) = \sum_{k=0}^{\infty} A^{2k+1}(\hat{a}) \cdot B^{2k+1}(\hat{b}) = \sum_{k=0}^{\infty} (-1)^k \frac{c^{2k+1}}{(2k+1)!} \hat{a}^{\otimes(2k+1)} \cdot \hat{b}^{\otimes(2k+1)}. \tag{3.19}$$

But $\hat{a}^{\otimes i} \cdot \hat{b}^{\otimes i} = \left(\hat{a} \cdot \hat{b}\right)^i$, since we can just evaluate the dot product separately on each tensor factor. Therefore

$$A(\hat{a}) \cdot B(\hat{b}) = \sum_{k=0}^{\infty} (-1)^k \frac{c^{2k+1}}{(2k+1)!} (\hat{a} \cdot \hat{b})^{2k+1} = \sin\left(c\hat{a} \cdot \hat{b}\right), \tag{3.20}$$

since the Taylor series for $\sin x = \sum_k (-1)^k x^{2k+1}/(2k+1)!$. It should now be apparent why the functions A and B were chosen as they were. Lemma 3.2.2 then implies that

$$\langle \alpha\beta \rangle = \frac{2}{\pi} \arcsin A(\hat{a}) \cdot B(\hat{b}) = \frac{2c}{\pi}\, \hat{a} \cdot \hat{b}. \tag{3.21}$$

It remains to check that $A(\hat{a})$ and $B(\hat{b})$ are unit vectors. We check this for $A(\hat{a})$:

$$A(\hat{a}) \cdot A(\hat{a}) = \sum_{k=0}^{\infty} \frac{c^{2k+1}}{(2k+1)!} \hat{a}^{\otimes(2k+1)} \cdot \hat{a}^{\otimes(2k+1)} = \sum_{k=0}^{\infty} \frac{c^{2k+1}}{(2k+1)!} = \sinh c = 1, \tag{3.22}$$

which explains why we chose c as we did. The calculation for $B(\hat{b})$ is almost identical. This verifies that we were correct in applying Lemma 3.2.2. ■

3.5 LHV model based on Krivine's upper bound on $K_G(3)$

In this section we present an LHV model for observables on the Werner state

$$\rho_p^W = p\,|\psi^-\rangle\langle\psi^-| + (1-p)\,\frac{\mathbb{1}}{4}, \tag{3.23}$$

valid for $p < 0.6595$. Although the LHV model in this section is just a special case of that which will be presented in Section 3.6, we present it separately for two reasons: (i) LHV models for qubit-qubit Werner states are of particular interest, and (ii) most physicists (ourselves included) are more familiar with spherical harmonics and Legendre polynomials than they are with the equivalent technical machinery in higher dimensions.

The protocol is similar to that developed in Section 3.4, in that it makes use of Protocol 3.2.1, but we use different functions A and B to obtain an LHV model that works for a larger range of parameters. Suppose Alice measures ρ_p^W along a direction $\hat{a} \in \mathbb{R}^3$, and Bob along $\hat{b} \in \mathbb{R}^3$.

In particular, let $A_3, B_3 : S_3 \to \bigoplus_{k=0}^{\infty} \bigoplus_{m=-(2k+1)}^{2k+1} \mathbb{R}^2$ and write

$$A_3(\hat{a}) = \bigoplus_{k=0}^{\infty} \bigoplus_{m=-(2k+1)}^{2k+1} A_3^{(2k+1,m)}(\hat{a}), \tag{3.24}$$

and similarly for B_3, which serves to define the "coordinates" of A_3 and B_3.

Let c_3 be the unique solution of

$$\frac{\sqrt{c_3}}{2}\int_0^{c_3} t^{-3/2}\sin t\,dt = 1 \tag{3.25}$$

in the interval $c_3 \in [0, \pi/2]$. Numerically we find that $c_3 \approx 1.0360$. We define

$$A_3^{(2k+1,m)}(\hat{a}) = (-1)^{k+1}\sqrt{\frac{4\pi^{3/2}J_{2k+3/2}(c_3)}{\sqrt{2c_3}}}\left(\operatorname{Re}Y_{2k+1}^m(\hat{a}), \operatorname{Im}Y_{2k+1}^m(\hat{a})\right), \tag{3.26}$$

$$B_3^{(2k+1,m)}(\hat{b}) = \sqrt{\frac{4\pi^{3/2}J_{2k+3/2}(c_3)}{\sqrt{2c_3}}}\left(\operatorname{Re}Y_{2k+1}^m(\hat{b}), \operatorname{Im}Y_{2k+1}^m(\hat{b})\right), \tag{3.27}$$

where Y_l^m are the spherical harmonics, and J_ν is the Bessel function of the first kind of order ν.

The protocol proceeds as follows:

Protocol 3.5.1 (LHV model for qubit-qubit Werner states)**. (Random Variables)** Alice and Bob share an infinite sequence $\lambda_1, \lambda_2, \ldots$ of real numbers, where each λ_i is drawn from a normal distribution with mean 0 and standard deviation 1. We write $\vec{\lambda} = (\lambda_1, \lambda_2, \ldots) \in l_\infty$.

(Alice) Alice outputs $\alpha = \operatorname{sgn}(A_3(\hat{a}) \cdot \vec{\lambda})$.

(Bob) Bob outputs $\beta = \operatorname{sgn}(B_3(\hat{b}) \cdot \vec{\lambda})$.

We claim that:

Theorem 3.5.2. *Protocol 3.5.1 results in correlations*

$$\langle\alpha\beta\rangle = -\frac{2c_3}{\pi}\hat{a}\cdot\hat{b}. \tag{3.28}$$

Before proving Theorem 3.5.2, we assemble a number of facts about special functions we shall need:

Lemma 3.5.3 (Addition Theorem for Spherical Harmonics)**.** *The spherical harmonics Y_n^m satisfy*

$$P_n(\hat{a}\cdot\hat{b}) = \frac{4\pi}{2n+1}\sum_{m=-n}^{n} Y_n^m(\hat{a})Y_n^{m*}(\hat{b}), \tag{3.29}$$

where P_n is the Legendre polynomial of order n, normalized such that $P_n(1) = 1$.

Lemma 3.5.4 (Orthogonality of Legendre polynomials)**.** *The Legendre polynomials satisfy*

$$\int_{-1}^{1} P_n(t)P_m(t)dt = \frac{2}{2n+1}\delta_{nm}. \tag{3.30}$$

Lemma 3.5.5 (Rodrigues' formula for P_n)**.** *The Legendre polynomial $P_n(x)$ can be written as*

$$P_n(x) = \frac{1}{2^n n!}\left(\frac{d}{dx}\right)^n (x^2-1)^n. \tag{3.31}$$

Lemma 3.5.6 (Definition of $J_\nu(x)$)**.** *The Bessel function of the first kind has integral representation*

$$J_\nu(x) = \frac{1}{\Gamma(\nu+1/2)\sqrt{\pi}} \left(\frac{x}{2}\right)^\nu \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \cos(x\sin t)\cos^{2\nu} t\, dt, \tag{3.32}$$

for $\nu > -1/2$ and $x \in \mathbb{R}$.

As before, to use our analysis of Protocol 3.2.1, we must first establish that the function A_3 maps unit vectors to unit vectors.

Lemma 3.5.7. *If $\hat{a}$ is a unit vector, then so is $A_3(\hat{a})$. Similarly, if $\hat{b}$ is a unit vector, then so is $B_3(\hat{b})$.*

Proof. We calculate

$$\begin{aligned}
A_3(\hat{a})\cdot A_3(\hat{a}) &= \sum_{k=0}^{\infty}\sum_{m=-(2k+1)}^{2k+1} A_3^{(2k+1,m)}(\hat{a})\cdot A_3^{(2k+1,m)}(\hat{a}) && (3.33)\\
&= \sum_{k=0}^{\infty} \frac{4\pi^{3/2}J_{2k+3/2}(c_3)}{4\sqrt{2c_3}}\,\mathrm{Re}\left[\sum_{m=-(2k+1)}^{2k+1} Y_{2k+1}^{m}(\hat{a})Y_{2k+1}^{m*}(\hat{a})\right] && (3.34)\\
&= \sum_{k=0}^{\infty} \frac{4\pi^{3/2}J_{2k+3/2}(c_3)}{\sqrt{2c_3}}\,\frac{4k+3}{4\pi}\,P_{2k+1}(1) && (3.35)\\
&= \sqrt{\frac{\pi}{2c_3}}\sum_{k=0}^{\infty}(4k+3)J_{2k+3/2}(c_3), && (3.36)
\end{aligned}$$

where we made use of Lemma 3.5.3.

It is a technical exercise to evaluate this sum (see Appendix A) and we find that

$$\sqrt{\frac{\pi}{2c_3}}\sum_{k=0}^{\infty}(4k+3)J_{2k+3/2}(c_3) = \frac{\sqrt{c_3}}{2}\int_0^{c_3} t^{-3/2}\sin t\, dt = 1, \tag{3.37}$$

by our choice of the constant c_3. This establishes that $A_3(\hat{a})$ is a unit vector. A similar analysis works for $B_3(\hat{b})$. ■

With these tools in hand, we are in a position to prove Theorem 3.5.2:

Proof of Theorem 3.5.2. We calculate

$$\begin{aligned}
A_3(\hat{a}) \cdot B_3(\hat{b}) &= \sum_{k=0}^{\infty} \sum_{m=-(2k+1)}^{2k+1} A_3^{(2k+1,m)}(\hat{a}) \cdot B_3^{(2k+1,m)}(\hat{b}) && (3.38)\\
&= \sum_{k=0}^{\infty} (-1)^{k+1} \frac{4\pi^{3/2} J_{2k+3/2}(c_3)}{4\sqrt{2c_3}} \mathrm{Re}\left[\sum_{m=-(2k+1)}^{2k+1} Y_{2k+1}^{m}(\hat{a}) Y_{2k+1}^{m*}(\hat{b}) \right] && (3.39)\\
&= \sum_{k=0}^{\infty} (-1)^{k+1} \frac{4\pi^{3/2} J_{2k+3/2}(c_3)}{\sqrt{2c_3}} \frac{4k+3}{4\pi} P_{2k+1}(\hat{a} \cdot \hat{b}) && (3.40)\\
&= \sqrt{\frac{\pi}{2c_3}} \sum_{k=0}^{\infty} (-1)^{k+1} (4k+3) J_{2k+3/2}(c_3) P_{2k+1}(\hat{a} \cdot \hat{b}). && (3.41)
\end{aligned}$$

We claim that this is the expansion of $-\sin(c_3 \hat{a} \cdot \hat{b})$ in Legendre polynomials. To check this, write

$$-\sin(c_3 t) = \sum_{k=0}^{\infty} a_{2k+1} P_{2k+1}(t), \tag{3.42}$$

where the expansion is guaranteed to exist by completeness of the Legendre polynomials. By the orthogonality of Legendre polynomials (Lemma 3.5.4), the coefficient a_{2k+1} is given by

$$a_{2k+1} = -\frac{4k+3}{2} \int_{-1}^{1} \sin(c_3 t) P_{2k+1}(t) dt. \tag{3.43}$$

We use Rodrigues' formula for P_{2k+1} (given in Lemma 3.5.5) and integrate by parts $2k+1$ times to obtain

$$\begin{aligned}
a_{2k+1} &= -\frac{4k+3}{2} \frac{1}{2^{2k+1}(2k+1)!} c_3^{2k+1} (-1)^{2k+1} (-1)^{k} \int_{-1}^{1} \cos(c_3 t)(t^2-1)^{2k+1} dt \\
&= \frac{(4k+3) c_3^{2k+1}}{2^{2k+2}(2k+1)!} (-1)^{k+1} \int_{-1}^{1} \cos(c_3 t)(1-t^2)^{2k+1} dt && (3.44)\\
&= \frac{(4k+3) c_3^{2k+1}}{2^{2k+2}(2k+1)!} (-1)^{k+1} \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \cos(c_3 \sin t') \cos^{2(2k+1)+1} t' dt' && (3.45)\\
&= \sqrt{\frac{\pi}{2c_3}} (-1)^{k+1} (4k+3) J_{2k+3/2}(c_3), && (3.46)
\end{aligned}$$

where we use the integral representation of $J_\nu(x)$ given in Lemma 3.5.6.

It follows that

$$A_3(\hat{a}) \cdot B_3(\hat{b}) = -\sin\left(c_3 \hat{a} \cdot \hat{b}\right), \tag{3.47}$$

and so, by Lemma 3.2.2, Protocol 3.5.1 gives correlations

$$\langle \alpha\beta \rangle = -\frac{2c_3}{\pi} \hat{a} \cdot \hat{b}. \tag{3.48}$$

This completes the proof. ■

3.6 LHV model based on Krivine's upper bound on $K_G(n)$

In this section, we generalize the LHV model of the previous section to higher dimensions. If we are interested in simulating the correlations of noisy states supported on Hilbert spaces of known dimension, then we can do slightly better than Protocol 3.4.1. In particular we saw in Section 3.7 that $K_G(2) = \sqrt{2}$, and in Section 3.4 that $\lim_{n\to\infty} K_G(n) = K_G \leq 1.78$. In this section we give better upper bounds (or better hidden variables models) for finite n.

We start with a number of definitions and facts. This section is rather technical: Ref. [50] provides an introduction to the required mathematics. Let H_k^n be the space of homogeneous polynomials in n variables of degree k, harmonic on $\mathbb{R}^n$. (A polynomial $h_k(\vec{v})$ is homogeneous of degree k if it satisfies $h_k(\lambda\vec{v}) = \lambda^k h_k(\vec{v})$ for $\lambda > 0$; it is harmonic if it satisfies Laplace's equation $\nabla^2 h_k(\vec{v}) = 0$.) The dimension of the space H_k^n is

$$N_k^n = (n+2k-2)(n+k-3)!/k!(n-2)!. \tag{3.49}$$

Let $\Delta_{k,j}^n$ be a real orthonormal basis for H_k^n for $j = 1, \ldots, N_k^n$. These functions are termed surface harmonics. They are normalized such that

$$\int_{S_k} \Delta_{k,j}^n \Delta_{k,j'}^n d\Omega = \delta_{jj'}. \tag{3.50}$$

Definition 3.6.1 (Gegenbauer polynomials)**.** The *Gegenbauer polynomial* of degree k and order ν is defined by the generating function

$$\sum_{k=0}^{\infty} C_k^\nu(x) r^k = \left(1 - 2rx + r^2\right)^{-\nu}. \tag{3.51}$$

Setting $\nu = 1/2$, we obtain the generating function for the Legendre polynomials, which implies that $C_k^{1/2}(x) = P_k(x)$. These polynomials are the generalization to higher dimensions of the Legendre polynomials. To make the comparison more explicit we have the following:

Definition 3.6.2 (Ultraspherical polynomial)**.** The ultraspherical polynomial of degree k and order n is

$$P_k^n(x) = \frac{C_k^{n/2-1}(x)}{C_k^{n/2-1}(1)} = \frac{k!(n-3)!}{(n+k-3)!}\, C_k^{n/2-1}(x). \tag{3.52}$$

Our terminology here is not standard: "Ultraspherical polynomial" is often synonymous with "Gegenbauer polynomial" as defined above; our ultraspherical polynomials are normalized, reparametrized

Gegenbauer polynomials.

There is a generalization of the addition theorem for spherical harmonics (Lemma 3.5.3) to higher dimensions:

Lemma 3.6.3 (Addition Theorem for Ultraspherical Polynomials (Theorem 11.4 of Ref. [50]))**.** *The surface harmonics $\Delta^n_{k,j}$ satisfy*

$$P^n_k(\hat{a}\cdot\hat{b}) = \frac{\omega_n}{N^n_k}\sum_{j=1}^{N^n_k}\Delta^n_{k,j}(\hat{a})\Delta^n_{k,j}(\hat{b}), \tag{3.53}$$

where $\omega_n = 2\pi^{n/2}/\Gamma(n/2)$ is the surface area of S_n.

Notice how Lemma 3.6.3 reduces to Lemma 3.5.3 when we set $n = 3$.

Next, we define a probability distribution on $[-1,+1]$ that mimics the distribution of $\hat{a}\cdot\hat{b}$ on $S_n \times S_n$:

Lemma 3.6.4. *Define*

$$\pi_n(dx) = \frac{\Gamma(n/2)}{\pi^{1/2}\Gamma((n-1)/2)}(1-x^2)^{(n-3)/2}dx. \tag{3.54}$$

Then for all continuous functions $F : [-1,+1] \to \mathbb{R}$,

$$\int_{-1}^{1} F(x)\pi_n(dx) = \int_{S_n}\int_{S_n} F(\hat{a}\cdot\hat{b})d\hat{a}d\hat{b}. \tag{3.55}$$

The ultraspherical polynomials are orthogonal with respect to this measure:

Lemma 3.6.5 (Orthogonality and completeness of ultraspherical polynomials (Eqs. (3.15.16) and (1.15.17) of Ref. [50]))**.** *The ultraspherical polynomials satisfy*

$$\int_{-1}^{1} P^n_k(x)P^n_l(x)\pi_n(dx) = \frac{1}{N^n_k}\delta_{kl}. \tag{3.56}$$

Furthermore, the ultraspherical polynomials form a complete system (with respect to π_n) for functions on $[-1,+1]$.

Just as in the three-dimensional case, we can write the function $\sin c_n x = \sum_{k=0}^{\infty} b^n_{2k+1}P^n_{2k+1}(x)$, where the coefficient $b^n_{2k+1} = N^n_{2k+1}\int_{-1}^{1}\sin c_n x P^n_{2k+1}(x)\pi_n(dx)$. Leaving aside the value of c_n for the moment, we can write down the LHV protocol. We define functions $A_n, B_n : S_n \to \bigoplus_{k=0}^{\infty}\bigoplus_{j=0}^{N^n_{2k+1}}\mathbb{R}$, with coordinates defined by

$$A_n(\hat{a}) = \bigoplus_{k=0}^{\infty}\bigoplus_{j=0}^{N^n_{2k+1}} A_n^{(2k+1,j)}(\hat{a}) \tag{3.57}$$

and

$$B_n(\hat{a}) = \bigoplus_{k=0}^{\infty} \bigoplus_{j=0}^{N_{2k+1}^n} B_n^{(2k+1,j)}(\hat{a}), \tag{3.58}$$

such that

$$\begin{aligned} A_n^{(2k+1,j)}(\hat{a}) &= (-1)^k \sqrt{\frac{b_{2k+1}^n N_{2k+1}^n}{\omega_n}}\, \Delta_{2k+1,j}^n(\hat{a}), & (3.59) \\ B_n^{(2k+1,j)}(\hat{b}) &= \sqrt{\frac{b_{2k+1}^n N_{2k+1}^n}{\omega_n}}\, \Delta_{2k+1,j}^n(\hat{b}). & (3.60) \end{aligned}$$

Protocol 3.6.6 (LHV model for correlations on arbitrary states)**.** **(Random Variables)** Alice and Bob share an infinite sequence $\lambda_1, \lambda_2, \ldots$ of real numbers, where each λ_i is drawn from a normal distribution with mean 0 and standard deviation 1. We write $\vec{\lambda} = (\lambda_1, \lambda_2, \ldots) \in l_\infty$.

(Alice) Alice outputs $\alpha = \mathrm{sgn}(A_n(\hat{a}) \cdot \vec{\lambda})$.

(Bob) Bob outputs $\beta = \mathrm{sgn}(B_n(\hat{b}) \cdot \vec{\lambda})$.

Theorem 3.6.7. *Protocol 3.6.6 results in correlations*

$$\langle \alpha\beta \rangle = \frac{2c_n}{\pi} \hat{a} \cdot \hat{b} \tag{3.61}$$

The proof that this protocol gives the correct correlations goes through in the same manner as the proof of Theorem 3.5.2 in the previous section. To determine the value of c_n, we require that $\hat{A}_n(\hat{a})$ be a unit vector, which is equivalent to the condition $\sum_k |b_{2k+1}^n| = 1$. We shall state Krivine's result without proof:

Theorem 3.6.8 (Krivine [33])**.** *The coefficient c_n is determined by*

$$\sum_k |b_{2k+1}^n| = \sum_k \frac{A_{2k+1}}{(2k+1)!} (c_n)^{2k+1} = 1, \tag{3.62}$$

where $A_1 = 1$ and

$$A_{2k+1} = \frac{(n-4)(n-8)\cdots(n-4k)}{(n+2)(n+6)\cdots(n+4k-2)}. \tag{3.63}$$

Using this expression, Krivine obtains $c_4 = 1$, and so on. The case $k = 8$ is of particular interest, in light of Corollary 2.3.5. We find $c_8 + c_8^3/15 = 1$, which gives $c_8 = 0.9439$. This implies $K_G(8) \leq 1.6641$, or that our LHV model for the joint correlation of projective measurements on $p\rho + (1-p)\mathbb{1}/4$ works for $p \leq 0.6009$.

3.7 LHV model based on Krivine's upper bound on $K_G(2)$

In this section, we deal with the scenario where Alice and Bob share the Werner state

$$\rho^W_{1/\sqrt{2}} = \frac{1}{\sqrt{2}} |\psi^-\rangle\langle\psi^-| + \left(1 - \frac{1}{\sqrt{2}}\right) \frac{\mathbb{1}}{4}, \tag{3.64}$$

and make measurements on this state along axes that lie in a fixed plane of the Poincaré sphere. In fact, it is sufficient if either Alice or Bob choses measurements from a plane, because then only the projection of the other's measurement axis onto that plane is relevant. We construct an LHV model for these measurements from Krivine's proof that $K_G(2) = \sqrt{2}$.

Fix a reference direction in the plane, and assume Alice measures her qubit at an angle x and Bob measures his at an angle y with $x, y \in [-\pi, \pi)$. Alice outputs a bit α and Bob a bit β depending on whether their qubit is aligned with, or opposite to, their measurement axis. We require that $\langle\alpha\beta\rangle = -\frac{1}{\sqrt{2}} \cos(x - y)$.

We shall make use of the following particularly nice fact about LHV models for simulating correlations of two-dimensional vectors:

Lemma 3.7.1. *Let $\lambda \in \mathbb{R}$. Suppose there is a protocol $\mathcal{P}$ that generates correlations $\langle\alpha\beta\rangle = F(x-y)$. Then there is a protocol $\mathcal{P}_\lambda$, which generates the correlations $\langle\alpha\beta\rangle = F(\lambda(x - y))$.*

Proof. Simply run the original protocol $\mathcal{P}$ on inputs λx and λy. ■

Note that we cannot obtain a similar result for vectors living in a space of dimension greater than two.

Before stating the protocol, we require a few definitions. Define a function $f : [-\pi, \pi) \to [-1, +1]$ such that $f(t) = f(-t)$, $f(t) = -f(\pi - t)$, and

$$f(t) = \begin{cases} 1 & \text{if } t \in [0, \pi/4], \\ 3\left(1 - \frac{2t}{\pi}\right) - 4\left(1 - \frac{2t}{\pi}\right)^3 & \text{if } t \in [\pi/4, \pi/2]. \end{cases} \tag{3.65}$$

(This is the mysterious part of the protocol. We explain the choice of f below.) We sketch the function f in Fig. 3.1. We can extend the function f to the real line by defining $f(t + 2\pi) = f(t)$. We also define a function $g : [-\pi, \pi) \to [-1, +1]$,

$$g(t) = \operatorname{sgn}(\cos t). \tag{3.66}$$

Let c_{2k+1} be the Fourier coefficients of f:

$$f(t) = \sum_{k=0}^{\infty} c_{2k+1} \cos(2k + 1)t, \tag{3.67}$$

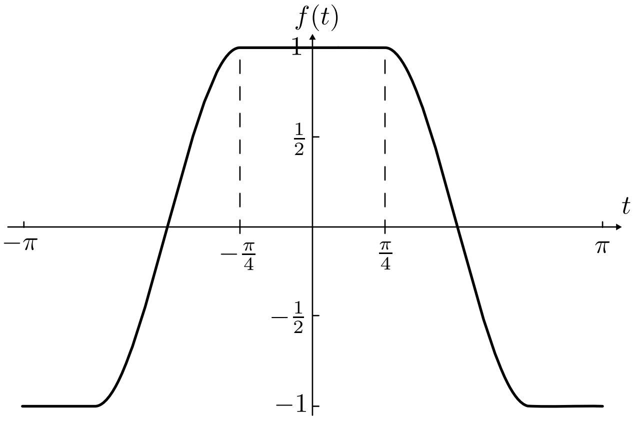


Figure 3.1: The function f defined by Eq. (3.65)

and d_{2k+1} be the Fourier coefficients of g:

$$g(t) = \sum_{k=0}^{\infty} d_{2k+1} \cos(2k+1)t. \tag{3.68}$$

Let

$$a_{2k+1} = \frac{c_{2k+1}d_{2k+1}}{2} = (-1)^k \frac{2}{\pi(2k+1)} c_{2k+1}, \tag{3.69}$$

where we have substituted for the coefficients d_{2k+1}. Finally, we define a sequence b_{2k+1}. Let $b_1 = 1/(\sqrt{2}\,a_1)$, and

$$b_{2k+1} = -\frac{1}{a_1} \sum_{d|(2k+1), d\neq 1} a_d b_{(2k+1)/d}, \tag{3.70}$$

for $k > 0$. Numerically, the first few terms are $b_1 = 0.9284$, $b_3 = -0.0597$, $b_5 = -0.0041$, and $b_7 = 0.0017$.

We require the following lemma to ensure that our protocol is well defined. We defer its proof until after presentation of the protocol.

Lemma 3.7.2. *The set* $\{|b_{2k+1}| : k = 0, 1, \dots\}$ *is a probability distribution, i.e.,* $\sum_k |b_{2k+1}| = 1$.

We are now in a position to define the protocol.

Protocol 3.7.3. (Random Variables) Alice and Bob share an integer $2n+1$ chosen randomly and distributed such that $\Pr(2n+1) = |b_{2n+1}|$. (In practice this can be accomplished by sharing a random number $r \in [0,1)$ and choosing n such that $\sum_{k=1}^{2n-1} |b_{2k+1}| \leq r < \sum_{k=1}^{2n+1} |b_{2k+1}|$.) They also share an angle $t \in [-\pi, \pi)$ drawn uniformly at random.

(Alice) Alice outputs

$$\alpha = -\operatorname{sgn}(b_{2n+1})\, g\left[(2n+1)(x-t)\right]. \tag{3.71}$$

(Bob) Bob outputs $\beta \in \{-1, +1\}$ such that

$$\langle \beta \rangle = f\left[(2n+1)(t-y)\right]. \tag{3.72}$$

In order to prove that Protocol 3.7.3 yields the correct correlations, we start by establishing some properties of the function $f(t)$. We shall need only these properties for the rest of the proof; the function $f(t)$, defined in Eq. (3.65), is chosen to establish the existence of a function that satisfies:

Lemma 3.7.4. *Let the function $f(t)$, defined by Eq. (3.65), have Fourier coefficients c_{2k+1}, as defined in Eq. (3.67). Let $\chi(2k+1) = \sqrt{2}\cos\left[(2k+1)\pi/4\right] (= \pm 1)$. Then*

(i) $f(t) = 1$ *for* $t \in [0, \pi/4]$;

(ii) $f(t) = f(-t)$ *and* $f(t) = -f(\pi - t)$;

(iii) $c_1 > 0$ *and* $\operatorname{sgn} c_{2k+1} = -(-1)^k \chi(2k+1)$ *for* $k > 0$;

(iv) $c_1 > \sum_{k=1}^{\infty} |c_{2k+1}|/(2k+1)$.

Proof. Properties (i) and (ii) are obvious. To establish (iii) and (iv) we calculate the Fourier coefficients of $f(t)$ directly. We find that

$$c_{2k+1} = \frac{768\cos\left[(2k+1)(\pi/4)\right]}{\pi^4(2k+1)^3}\left(\frac{1}{2k+1} - (-1)^k\frac{\pi}{4}\right). \tag{3.73}$$

Property (iii) follows immediately. To prove (iv), we either plug Eq. (3.73) into a computer algebra system, or calculate (following Krivine),

$$\begin{aligned}
\frac{\sqrt{2}\,\pi^4}{768}\sum_{k=1}^{\infty}\frac{|c_{2k+1}|}{(2k+1)} &= \sum_{k=1}^{\infty}\frac{1}{(2k+1)^4}\left(\frac{\pi}{4} - \frac{(-1)^k}{2k+1}\right) && (3.74)\\
&< \sum_{k=1}^{\infty}\frac{1}{(2k+1)^4}\left(\frac{\pi}{4} + \frac{1}{3}\right) && (3.75)\\
&= \left(\frac{\pi^4}{96} - 1\right)\left(\frac{\pi}{4} + \frac{1}{3}\right) && (3.76)\\
&< 1 - \pi/4 = \frac{\sqrt{2}\,\pi^4}{768}c_1, && (3.77)
\end{aligned}$$

using $\sum_{k=0}^{\infty} 1/(2k+1)^4 = \pi^4/96$. ■

Continuing, for the moment, to defer proof of Lemma 3.7.2 (which ensured our protocol was well defined), we prove

Theorem 3.7.5. *Protocol 3.7.3 results in correlations*

$$\langle \alpha\beta \rangle = -\frac{1}{\sqrt{2}}\cos(x-y). \tag{3.78}$$

Proof. Integrating over the random variables, we find that

$$\langle\alpha\beta\rangle = -\sum_n b_{2n+1} \int \frac{dt}{2\pi}\, g((2n+1)(x-t))f((2n+1)(t-y)). \tag{3.79}$$

Let $F = f \star g$ be the convolution of f and g:

$$F(x-y) = \int \frac{dt}{2\pi} f(x-t)g(t-y) = \sum_k a_{2k+1}\cos\left[(2k+1)(x-y)\right], \tag{3.80}$$

recalling the definition of a_{2k+1}. It follows that

$$\begin{aligned}
\langle\alpha\beta\rangle &= -\sum_n b_{2n+1}F\left[(2n+1)(x-y)\right] && (3.81)\\
&= -\sum_n\sum_m b_{2n+1}a_{2m+1}\cos\left[(2m+1)(2n+1)(x-y)\right] && (3.82)\\
&= -\sum_k \sum_{(2m+1)|(2k+1)} a_{2m+1}b_{(2k+1)/(2m+1)}\cos\left[(2k+1)(x-y)\right], && (3.83)
\end{aligned}$$

where the rearrangement of the summations is justified because the series $\sum_k a_{2k+1}$ and $\sum_k b_{2k+1}$ are both absolutely converging (the former by Eq. (3.69) and property (iv) of Lemma 3.7.4; the latter by Lemma 3.7.2).

The coefficients b_{2k+1} were chosen such that $a_1 b_1 = 1/\sqrt{2}$ and

$$\sum_{(2m+1)|(2k+1)} a_{2m+1}b_{(2k+1)/(2m+1)} = 0 \tag{3.84}$$

if $k > 0$ (this is just a rearrangement of Eq. (3.70)), so only the $k = 0$ term survives in Eq. (3.83) and we have $\langle\alpha\beta\rangle = -\frac{1}{\sqrt{2}}\cos(x-y)$. ■

Proof of Lemma 3.7.2. We first determine the signs of the b_k. The Fourier coefficients a_{2k+1} are given by Eq. (3.69) so that $a_1 > 0$ and $\operatorname{sgn} a_{2k+1} = -\chi(2k+1)$ for $k > 0$ by property (iii) of Lemma 3.7.4. We prove by induction that $\operatorname{sgn} b_{2k+1} = \chi(2k+1)$. We need the fact that $\chi(mn) = \chi(m)\chi(n)$ where m and n are odd integers, which is easily established (indeed it is sufficient to check for $m, n \in \{\pm1, \pm3\}$ since $\chi(m) = \chi(m+8)$). It is clear that $\operatorname{sgn} b_1 = \chi(1)$. The coefficient b_{2k+1} is defined inductively via

$$b_{2k+1} = -\frac{1}{a_1}\sum_{d|(2k+1), d\neq 1} a_d b_{(2k+1)/d}, \tag{3.85}$$

and $\operatorname{sgn}(a_d b_{(2k+1)/d}) = -\chi(d)\chi((2k+1)/d) = -\chi(2k+1)$ for all $1 < d < 2k+1$ by the inductive hypothesis. Hence all the terms in the sum have the same sign and we have $\operatorname{sgn} b_{2k+1} = \chi(2k+1)$.

So we must compute $\sum|b_{2k+1}| = \sum b_{2k+1}\chi(2k+1)$. To this end, define the Dirichlet series

$$D(s) = \sum_{k=0}^{\infty} \frac{a_{2k+1}\chi(2k+1)}{(2k+1)^s}. \tag{3.86}$$

From property (iv) of Lemma 3.7.4 it follows that $a_1 > |a_3| + |a_5| + \cdots$, so the series $D(s)$ is absolutely converging for $s \geq 0$ and has an inverse,

$$1/D(s) = \sqrt{2}\sum_{k=0}^{\infty} \frac{b_{2k+1}\chi(2k+1)}{(2k+1)^s}, \tag{3.87}$$

where the coefficients are given by Eq.(3.70), as can be verified by equating coefficients in the product term by term.

Substituting $s = 0$ gives

$$\begin{aligned}\sum b_{2k+1}\chi(2k+1) &= \left[\sqrt{2}\sum a_{2k+1}\chi(2k+1)\right]^{-1} && (3.88)\\ &= \left[2\sum a_{2k+1}\cos\left[(2k+1)\pi/4\right]\right]^{-1} && (3.89)\\ &= \frac{1}{2F(\pi/4)}. && (3.90)\end{aligned}$$

Now $F(\pi/4) = \int_{-\pi}^{\pi}\frac{dt}{2\pi}f(t)g(\frac{\pi}{4}-t)$. Noting that $g(\frac{\pi}{4}-t) = \operatorname{sgn}\cos(\frac{\pi}{4}-t) = 1$ if $t \in [-\pi/4, 3\pi/4]$, we find

$$F(\pi/4) = \int_{-\pi/4}^{3\pi/4}\frac{dt}{2\pi}f(t) - \int_{-5\pi/4}^{-\pi/4}\frac{dt}{2\pi}f(t) = \frac{2}{\pi}\int_0^{\pi/4}dt\,f(t), \tag{3.91}$$

using property (ii) of Lemma 3.7.4. Now, since $f(t) = 1$ if $t \in [0, \pi/4]$ by property (i) of the same lemma, we have $F(\pi/4) = 1/2$, which implies that $\sum_k|b_{2k+1}| = 1$. ■

It is instructive to look through the proof and note where we used each of the four properties of the function $f(t)$ enumerated in Lemma 3.7.4. Property (iv) ensures that the Dirichlet series $D(s)$, defined in Eq. (3.86), is absolutely converging for $s > 0$, which means we may write $1/D(s)$ as a Dirichlet series that itself is absolutely converging. Property (iii) ensures that we can compute the signs of the coefficients a_{2k+1} and b_{2k+1}. Finally, properties (i) and (ii) ensure that $F(\pi/4)$ is as large as possible, which makes $\sum_k|b_{2k+1}|$ as small as possible, which gives the best possible LHV model.

3.8 Discussion

In this chapter we have constructed explicit LHV models for the joint correlation resulting from making two-outcome projective measurements on quantum states. This work raises a number of questions:

1. It is rather unsatisfying to have results only for the joint correlation. Can we generalize these techniques to obtain LHV models for the full probability distribution?

2. Can we obtain better bounds on $K_G(n)$ and, in particular, on $K_G(3)$? I have attempted to improve the lower bound on $K_G(3)$ elsewhere, thus far without success [48]. I also suspect that Krivine's upper bound on $K_G(3)$ can be improved a little. This is because we know that the upper bound given by the same construction (i.e., that of Section 3.6) is not tight for $n = 2$. Perhaps a similar construction to that used in Section 3.7 is possible in the case $n = 3$, but the analysis will be messier due to the absence of Lemma 3.7.1 when $n > 2$.

3. Do all proofs of Grothendieck's inequality (and there are many) give rise to explicit LHV models? For example, I do not see how to extract an LHV model from the proof in Ref. [51] (see Theorem 6.27, p. 152).

Chapter 4

Bell inequalities with auxiliary communication

4.1 Introduction

We turn now to a different measure of nonlocality, the communication cost of classical simulation. We generalize Bell inequalities to the setting of local realistic theories augmented by a fixed amount of classical communication. Suppose two parties choose one of M two-outcome measurements and exchange one bit of information. We present the complete set of inequalities for M = 2, and the complete set of inequalities for the joint correlation observable for M = 3. We find that correlations produced by quantum theory satisfy both of these sets of inequalities. One bit of communication is therefore sufficient to simulate quantum correlations in both of these scenarios. This chapter is joint work with Dave Bacon.

Quantum correlations on spacelike separated systems cannot be reproduced classically. If, however, the systems are timelike separated, then classical simulation is possible, albeit at the expense of some communication, but *how much* is required? In particular, suppose a number of spatially separate parties share an entangled quantum state, and each makes a local measurement on their component. Then quantum correlations are manifest in the *joint probability distribution* of the parties' outcomes, dependent on each party's choice of measurement. If this probability distribution cannot be reproduced by a classical local realistic theory, then it violates some generalized Bell inequality [19]. This means some communication between the parties is required to reproduce the probability distribution, but Bell inequality violation does nothing to *quantify* how much. More generally, entanglement is a *resource* for performing information processing tasks, and an important goal of quantum information theory is to demarcate the difference between it and classical resources, such as shared randomness and classical communication channels. What classical resources are required to reproduce the joint probability distributions arising from local measurement on shared quantum states?

Within the setting of local realistic theories augmented by a fixed amount of *two-way* classical communication [52], we introduce the notion of *Bell inequalities with auxiliary communication.* These inequalities provide conditions on the joint probability distribution, which must be satisfied if such correlations can be simulated with shared randomness and a fixed amount of communication. Of particular significance are complete sets of such inequalities, which provide necessary and sufficient conditions. In the scenario where two parties choose one of M two-outcome measurements and exchange one bit of information, we present the complete set of inequalities for $M = 2$, and the complete set of inequalities for the joint correlation observable for $M = 3$. We find that quantum correlations satisfy all of these inequalities, irrespective of the particular quantum state or the specific measurements, and can therefore be explained in these settings by augmenting a local realistic theory with a single bit of communication. This is particularly remarkable for the $M = 3$ case, where one would naively expect a trit of auxiliary communication is required to simulate quantum correlations.

4.2 The model

We restrict attention to scenarios with two parties, A and B. In a *quantum measurement scenario* for this bipartite case, each party selects one of M different measurements and then—possibly after some delay, during which we might allow the parties to communicate—outputs one of K different outcomes. (Note that A and B may choose measurements from different M-element sets.) Such a measurement scenario results in a set of probabilities $0 \leq p(a, b|i, j) \leq 1$, where $p(a, b|i, j)$ is the probability that A outputs a and B outputs b, given that A selects measurement i and B selects measurement j. Discounting null outcomes (which can be incorporated as a separate outcome if desired), it follows that $\sum_{a=0}^{K-1} \sum_{b=0}^{K-1} p(a, b|i, j) = 1$, where $0 \leq i, j \leq M - 1$. A valid measurement scenario is any set of probabilities that satisfies these normalization constraints.

Given a particular measurement scenario, we investigate all *protocols* that the two parties might perform to produce the correct probabilities. A protocol consists of three stages: (i) *preparation* via the distribution of shared randomness, (ii) *communication* via the exchange of messages between the parties, and (iii) *output* of outcomes by each party as determined by information accessible to each party. A and B select their measurements after step (i) but before step (ii). If a protocol produces identical probabilities to the measurement scenario, then we say that the protocol has *simulated* the scenario.

Two informational resources are of interest: the quantity of shared randomness and the amount of communication between the parties. We focus on the amount of communication and define the *cost* of a protocol to be the maximum amount of communication required (as opposed to the average amount of communication, see [53, 54, 55, 56, 57]). In the preparation phase of a protocol, we allow A and B to share an infinite amount of classical information and, in particular, continuous variables.

In the parlance of foundational studies of quantum theory, these are known as local hidden variables (LHVs) [58]. A protocol with no communication (step (ii) missing) is usually called an *LHV theory.* In such a theory, each party's output depends on the shared randomness and on which measurement the particular party has locally selected, but not on the measurement choice of the other party.

The protocols we investigate are therefore an extension of LHV theories, where we allow the parties to communicate after selecting measurements [52]. This allows some "which-measurement" information to propagate between the parties. We emphasize that a protocol of this form simulates the joint probability distribution resulting from a set of quantum measurements, *but not* the quantum measurements themselves: It is not possible to replace local measurements made by two spacelike separated parties on an entangled quantum state by classical communication. Even in this case, however, the amount of two-way communication required to reproduce the joint probability distribution provides a measure of the *nonlocality* of the correlations. From an information processing perspective, this model provides a fair setting for the comparison of quantum correlations and classical resources required to reproduce them.

Of particular significance in this respect is the result of Brassard, Cleve, and Tapp [52], who demonstrated that the correlations produced by two-outcome projective measurements on an EPR pair can be simulated by a local realistic theory augmented by eight bits of communication. Surprisingly, a single bit of communication is sufficient, as we shall see in the next chapter.

Little, however, is known for more general states and more general measurements. The goal of this paper is to illuminate how such bounds can be achieved by generalizing Bell inequalities to what we term *Bell inequalities with auxiliary communication.*

4.3 Bell polytopes

We described how to construct Bell inequalities without auxiliary communication in Section 1.2.3. For completeness, we reproduce the argument here. Bell inequalities [7, 25] describe necessary conditions on the probabilities $p(a, b|i, j)$, which must be satisfied if these probabilities are to be produced by a local realistic theory. When a set of these conditions is also sufficient, we say that we have a complete set of Bell inequalities. The construction of complete sets of Bell inequalities is an exercise in convex geometry.

Consider a deterministic protocol, i.e., one in which no randomness, shared or otherwise, is used. (This corresponds to a protocol consisting only of step (iii) above, with the additional requirement that this step is deterministic.) Each party's output can only depend on their local which-measurement information, so that all such protocols can be completely characterized by two functions $\alpha, \beta : \mathbb{Z}_M \to \mathbb{Z}_K$, which describe the outcomes of the two parties' measurements: If A selects measurement i, she outputs $\alpha(i)$ and if B selects measurement j, he outputs $\beta(j)$. The probabilities for the scenario are then $p(a,b|i,j) = \delta^a_{\alpha(i)}\delta^b_{\beta(j)}$.

By listing the components, we may view the probabilities $p(a,b|i,j)$ as vectors $\vec{p}$ in $\mathbb{R}^D$ with $D = M^2(K^2-1)$ (recall the constraint $\sum_{a,b} p(a,b|i,j) = 1$). To each pair of functions $\{\alpha,\beta\}$, there corresponds a deterministic protocol, so the set of all deterministic protocols is a finite collection of such vectors $\{\vec{d}_\zeta|\zeta = 1,...,K^{2M}\}$.

Now consider the effect of allowing randomness. Any unshared randomness can always be replaced by shared randomness on which the other party does not act [59], so we may continue to assume step (iii) is deterministic. But then every set of random variables in step (i) corresponds to a particular deterministic protocol. Therefore the set of *all possible protocols that use randomness and no communication* is described by a convex sum of the deterministic protocols without communication

$$\vec{p} = \sum_\zeta \lambda_\zeta \vec{d}_\zeta, \quad \sum_\zeta \lambda_\zeta = 1, \quad \lambda_\zeta \geq 0. \tag{4.1}$$

The set of all protocols therefore corresponds to a region Ω_{MK} in $\mathbb{R}^D$, which is a polytope because there is a finite number of extreme vectors $\vec{d}_\zeta$ [20]. This permits an alternative description: Instead of describing the polytope Ω_{MK} as the convex combination of a finite set of extreme points, we can describe it by specifying a complete (finite) set of facet inequalities. A facet inequality is a pair $\{\vec{f}, c\}$ that defines a half-space of $\mathbb{R}^D$ via the inequality $\vec{f}\cdot\vec{p} \leq c$. Complete sets of facet inequalities $\vec{f}_\eta, c_\eta$ are satisfied if and only if $\vec{p}$ is in Ω_{MK}:

$$\vec{p} \in \Omega_{MK} \text{ iff } \vec{f}_\eta \cdot \vec{p} \leq c_\eta, \ \forall\eta. \tag{4.2}$$

Each facet is therefore a Bell inequality and complete sets of facet inequalities are complete sets of Bell inequalities. Complete sets are known in the two-party case when $M = 2, K = 2$ [21], $M = 3, K = 2$ [22], and also when extra symmetry constraints are imposed [24].

4.4 Bell inequalities with auxiliary communication

We now turn to the main focus of this chapter: extending the formalism of Bell inequalities to protocols that permit communication after the parties have chosen their measurements. Again consider a deterministic protocol, but now allow for the communication (possibly two way) of at most r bits of information between the parties after selection of measurements. Such a protocol is completely characterized by two functions $\alpha, \beta : \mathbb{Z}_M \otimes \mathbb{Z}_M \to \mathbb{Z}_K$, which describe the outcomes of the two parties' measurements, but now each party's output can also depend on which measurement the other party selects: If A selects measurement i and B measurement j, A outputs $\alpha(i,j)$ and B outputs $\beta(i,j)$. The probabilities for such a deterministic protocol are then $p(a,b|i,j) = \delta^a_{\alpha(i,j)}\delta^b_{\beta(i,j)}$. While α and β can now depend on which measurement the other party selects, not all functions

$\alpha(i,j)$, $\beta(i,j)$ are necessarily accessible, if the parties exchange at most r bits of communication. The set of possible functions $\alpha(i,j)$, $\beta(i,j)$ for protocols that use at most r bits of communication is the subject of the field of communication complexity [60, 59]. For example, with a single bit of communication from A to B, $\alpha(i,j)$ is independent of B's measurement j, and $\beta(i,j)$ can depend only on a partition of the set of possible i's into two sets. Despite this complication, deterministic protocols still correspond to a finite set of vectors of probabilities $\vec{d}_\zeta^{(r)}$ in $\mathbb{R}^D$.

If we now allow randomness, the set of accessible probabilities $\Omega_{MK}^{(r)}$ is given by the convex combination of the deterministic probabilities, $\vec{p} = \sum_\zeta \lambda_\zeta \vec{d}_\zeta^{(r)}$, $\sum_\zeta \lambda_\zeta = 1$, $\lambda_\zeta \geq 0$. Again, $\Omega_{MK}^{(r)}$ is a convex combination of a finite number of extreme points—a polytope—and can be described by a finite set of facet inequalities: $\vec{p} \in \Omega_{MK}^{(r)}$ iff $\vec{f}_\eta^{(r)} \cdot \vec{p} \leq c_\eta, \forall \eta$. The complete set of facet inequalities for $\Omega_{MK}^{(r)}$ is a complete set of Bell inequalities with r bits of communication. An important limit arises when $r \geq 2\log_2 M$ because then each party can tell the other exactly which measurement they have selected. In this setting, all deterministic protocols can be executed by the two parties: The probability distribution $p(a,b|i,j)$ is unrestricted. This implies that Bell inequalities with auxiliary communication are trivial when $M = 1$.

Additionally, for $r \geq \log_2 M$, Bell inequalities with auxiliary communication, although not necessarily trivial, are never violated by probability distributions arising from local measurements on a shared quantum state. In fact this is true for any probability distribution satisfying the *no-one-way-signaling conditions*: $p_{a|i,j} \equiv \sum_{b=0}^{K-1} p(a,b|i,j) = p_{a|i}$ is independent of j for all a and i. In other words, A's marginal probability distribution is independent of B's choice of measurement. In such cases, it is sufficient that only A communicate her measurement choice. The simulation procedure is as follows: For each of A's measurements i, the parties share a random variable $\tilde{a}_i$ drawn from the probability distribution $\{a, p_{a|i}\}$ (i.e. $\tilde{a}_i = a$ with probability $p_{a|i}$). Suppose A chooses measurement i and B chooses measurement j. A outputs $\tilde{a}_i$ and sends her measurement choice i to B. B then outputs $\tilde{b}_{\tilde{a}_i,i,j}$, where $\tilde{b}_{a,i,j}$ is drawn from the probability distribution $\{b, p_{\tilde{a}_i,b|i,j}\}$. (The roles of A and B in the no-one-way-signaling conditions and protocol may be reversed.)

4.4.1 Complete set of Bell inequalities with auxiliary communication

Consider the simplest case $M = K = 2$ and $r = 1$ bit. The polytope $\Omega_{2,2}^{(1)}$ is 12 dimensional and has 112 extreme vectors. Using both the primal-dual algorithm and the double description method [61, 62] for facet enumeration, we find that this polytope has 48 facets. Sixteen facets describe trivial inequalities, $p(a,b|i,j) \geq 0$ $(0 \leq i,j,a,b \leq 1)$. Another 16 facets are of the form

$$p(a_1,b_1|0,0) + p(a_2,b_2|0,1) + p(a_3,b_3|1,0) + p(a_4,b_4|1,1) \leq 2, \tag{4.3}$$

with $(a_1a_2a_3a_4) \in \{(0101),(1010),(0110),(1001)\}$ and $(b_1b_2b_3b_4) \in \{(0011),(1100),(0110),(1001)\}$. The remaining 16 facets are given by

$$p(a,0|i,j) + p(a,1|i,j) + p(0,b|\bar{i},\bar{j}) + p(1,b|\bar{i},\bar{j}) - p(a,b|i,\bar{j}) \geq 0 \tag{4.4}$$

$(0 \leq i,j,a,b \leq 1)$, where $\bar{0} = 1$ and $\bar{1} = 0$. The above inequalities completely describe the region of probability distributions that can be created with one bit of communication. There are probability distributions that violate these inequalities: For example, if $p(a,b|i,j) = \delta_j^a \delta_i^b$, Eq. (4.3) with $(a_1a_2a_3a_4) = (0101)$ and $(b_1b_2b_3b_4) = (0011)$ is maximally violated; substitution gives $4 \not\leq 2$.

It is straightforward to check that any probability distribution satisfying the no-signaling conditions satisfies inequalities Eq. (4.3) and Eq. (4.4). Finally, consider the probability distribution $p(a,b|i,j) = \frac{1}{2}\left(\delta_0^a \delta_i^b + \delta_j^a \delta_0^b\right)$. This probability distribution violates the no-signaling conditions (in both directions), but satisfies Eq. (4.3) and Eq. (4.4), thus indicating that these inequalities are strictly stronger than no-signaling.

4.4.2 Complete set of Bell inequalities for the joint observable

The above complete Bell inequalities with auxiliary communication were used to bound the allowed probabilities $p(a,b|i,j)$ for protocols using a specified amount of communication. In quantum theory we are often not interested in all of the probabilities for a measurement scenario, but only on the value of a particular joint observable. This simplifies our computational task, because we may project the polytope $\Omega_{MK}^{(r)}$ onto a lower dimensional subspace and only enumerate the facets of the projected polytope, as we shall explain in the following. We term a complete set of facet inequalities for this convex set a complete set of Bell joint observable inequalities with auxiliary communication. These generalize the CHSH inequality [25] to protocols with communication.

Consider a measurement scenario with probabilities $p(a,b|i,j)$ and identify measurement outcomes with values of local observables. The joint observable for the ith and jth measurements of A and B is then defined as

$$c_{i,j} = \sum_{a=0}^{K-1} \sum_{b=0}^{K-1} A_a B_b p(a,b|i,j), \tag{4.5}$$

where A_a and B_b are the values of the local observable corresponding to measurement outcomes a and b, respectively. As for the full measurement scenario, we may list the components of the joint observable to form a vector $\vec{c}$ in $\mathbb{R}^D$ with $D = M^2$ (compare $D = M^2(K^2-1)$ for the full probability distribution). For deterministic protocols with at most r bits of communication, the allowed functions α and β are the same as in the previous section, but now correspond to vectors with components $c_{i,j} = A_{\alpha(i,j)} B_{\beta(i,j)}$. Since the map given by Eq. (4.5) is linear, the vectors corresponding to joint observables accessible using randomness remain convex combinations of the vectors accessible via deterministic protocols.

We now specialize to the scenario where each party has local ±1-valued observables ($K = 2$) and they exchange $r = 1$ bit of communication. The joint correlation observable then has components $c_{i,j} = p(0,0|i,j) + p(1,1|i,j) - p(0,1|i,j) - p(1,0|i,j)$. If $M = 2$, we obtain only trivial inequalities $-1 \leq c_{i,j} \leq 1$. That the inequalities are trivial also follows from Eqs. (4.3) and (4.4), for in this scenario all possible joint observables can be obtained from probability distributions that satisfy the no-signaling conditions [17, 18].

If $M = 3$, the polytope has 320 extreme vectors. Using again both the primal-dual algorithm and double description method for facet enumeration we find that this polytope has 498 facets. Eighteen of these describe the trivial inequalities $-1 \leq c_{i,j} \leq 1$. The remaining 480 facets can be described by the inequalities

$$\sum_{i,j=0}^{2} M_{i,j} c_{i,j} \leq 1, \tag{4.6}$$

where $M_{i,j}$ is either

$$M_1 = \frac{1}{6}\begin{pmatrix} 0 & -1 & 1 \\ -1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix},\ M_2 = \frac{1}{11}\begin{pmatrix} 1 & 2 & -2 \\ 2 & 1 & 2 \\ -2 & 2 & 1 \end{pmatrix}, \tag{4.7}$$

or any matrix obtained from these two matrices by (i) permuting the rows and/or columns of the matrix and/or (ii) multiplying any subset of the rows and columns of the matrix by -1. The full set of inequalities is a complete set of Bell joint observable inequalities for $M = 3$.

Let us show that quantum theory satisfies all of the above Bell joint observable inequalities with auxiliary communication. We do this for a single one of the inequalities and the other inequalities all follow by a similar argument. For ±1-valued observables $\mathbf{A}_i$ and $\mathbf{B}_j$ and the joint quantum state ρ, a particular inequality looks like $\mathrm{Tr}\,[\rho \mathbf{T}_i] \leq 1$, where $\mathbf{T}_i$ is the operator corresponding to matrix M_i, e.g., $\mathbf{T}_1 = [\mathbf{A}_1(-\mathbf{B}_2 + \mathbf{B}_3) + \mathbf{A}_2(-\mathbf{B}_1 + \mathbf{B}_2 + \mathbf{B}_3) + \mathbf{A}_3(\mathbf{B}_1 + \mathbf{B}_2 + \mathbf{B}_3)]/6$. $\mathrm{Tr}\,[\rho\mathbf{T}]$ is bounded by the sup norm of $\mathbf{T}$, $|\mathbf{T}| = \sup_{|\psi\rangle} \|\mathbf{T}|\psi\rangle\| / \||\psi\rangle\|$ and further $\mathrm{Tr}\,[\rho\mathbf{T}] \leq |\mathbf{T}^k|^{1/k}$. Calculation of $\mathbf{T}^k$ yields a polynomial in $\mathbf{A}_i$, $\mathbf{B}_i$, and $\mathbf{I}$. Since $|\mathbf{X} + \mathbf{Y}| \leq |\mathbf{X}| + |\mathbf{Y}|$ and $|\mathbf{P}| \leq 1$ for any product $\mathbf{P}$ of $\mathbf{A}_i$, $\mathbf{B}_i$, and $\mathbf{I}$, it follows that $|\mathbf{T}^k|$ is less than or equal to the sum of the absolute value of the coefficients in the polynomial expansion of $\mathbf{T}^k$. By computer calculation we find that the sum of the absolute value of the coefficients of $\mathbf{T}_1^4$ is $\frac{155}{162}$ so that $|\mathbf{T}_1| \leq \sqrt[4]{\frac{155}{162}}$. Thus this Bell inequality with auxiliary communication is satisfied. Similar arguments using $\mathbf{T}_1^4$ or $\mathbf{T}_2^5$ show that all of the inequalities Eq. (4.6) are satisfied. Therefore in the scenario where each party chooses one of three two-outcome measurements, a single bit of communication is sufficient to simulate the joint correlation observable in quantum theory for all quantum states and all quantum observables.

4.5 Conclusion

Bell inequalities with auxiliary communication are a powerful new tool for understanding the *cost* of producing quantum correlations. In all the cases we considered, it was sufficient to augment local realistic theories with a single bit of communication to simulate the quantum correlations.

Chapter 5

Communication cost of simulating Bell correlations

5.1 Introduction

In this chapter, we continue our study of the communication cost of simulating quantum correlations. For the simplest and most important case of local projective measurements on an entangled Bell pair state, we show that exact simulation is possible using local hidden variables augmented by just one bit of classical communication. Certain quantum teleportation experiments, which teleport a single qubit, therefore admit a local hidden variables model. This chapter is joint work with Dave Bacon.

Bell pairs are the maximally entangled states of two quantum bits (qubits) and are the basic resource currency of bipartite quantum information theory. Various equivalences are known: One shared Bell pair plus two bits of classical communication can be used to teleport one qubit [63] and, conversely, one shared Bell pair plus a single qubit of communication can be used to send two bits of classical communication via superdense coding [64].

Consider the gedanken experiment of Einstein, Podolsky, and Rosen [6] (EPR), as reformulated by Bohm [65]. Two spatially separate parties, Alice and Bob, each have a spin-$\frac{1}{2}$ particle, or qubit. The global spin wave function is the entangled Bell singlet state (also known as an EPR pair) $|\psi\rangle = \frac{1}{\sqrt{2}}(|\uparrow\rangle_A|\downarrow\rangle_B - |\downarrow\rangle_A|\uparrow_B\rangle)$. The spin states $|\uparrow\rangle$, $|\downarrow\rangle$ are defined with respect to a local set of coordinate axes: $|\uparrow\rangle$ corresponds to spin along the local $+\hat{z}$ direction, while $|\downarrow\rangle$ corresponds to spin along the local $-\hat{z}$ direction. Alice and Bob each measure their particle's spin along a direction parameterized by a three-dimensional unit vector: Alice measures along $\hat{a}$, Bob along $\hat{b}$. Alice and Bob obtain results, $\alpha \in \{+1, -1\}$ and $\beta \in \{+1, -1\}$, respectively, which indicate whether the spin was pointing along $(+1)$ or opposite (-1) the direction each party chose to measure. Locally, Alice and Bob's outcomes appear random, with expectation values $\langle\alpha\rangle = \langle\beta\rangle = 0$, but their joint probabilities are correlated such that $\langle\alpha\beta\rangle = -\hat{a}\cdot\hat{b}$. We refer to these correlations as *Bell correlations*.

It is not possible to reproduce these correlations using a protocol that draws on random variables shared between Alice and Bob, but does not allow communication after they have selected measurements [7]. So how much communication *is* required to exactly simulate them [53, 52, 55, 66, 2, 54, 56]? Naively, Alice can just tell Bob the direction of her measurement $\hat{a}$ (or vice versa), but this requires an infinite amount of communication. The question of whether a simulation can be done with a finite amount of communication was raised independently by Maudlin [53], Brassard, Cleve, Tapp [52], and Steiner [55]. Their approaches differ in how the communication cost of the simulation is defined: Brassard *et al.* take the cost to be the number of bits sent in the worst case, Steiner, the average. (Steiner's model is weaker because the amount of communication in the worst case can be unbounded, although such cases occur with probability zero.) Brassard *et al.* present a protocol that simulates Bell correlations using exactly eight bits of communication (since improved to six bits [66]). In light of the previous chapter, where we could prove no nontrivial lower bounds, the only lower bound for the amount of communication is given by Bell's theorem: At least some communication is needed. Here we present a simple protocol that uses just one bit of communication.

5.2 Simulation protocol

We first note three simple properties of Bell correlations: (i) if $\hat{a} = \hat{b}$, then we must have $\alpha = -\beta$: Alice and Bob must output perfectly anticorrelated bits; (ii) either party can reverse their measurement axis and flip their output bit; and (iii) the joint probability is only dependent on $\hat{a}$ and $\hat{b}$ via the combination $\hat{a} \cdot \hat{b}$. In his original paper, Bell gave a local hidden variables model that reproduces these three properties for all possible axes, but his model fails to reproduce Bell correlations because the statistical correlations when $\hat{a} \neq \hat{b}$ are not as strong as those of quantum mechanics [7]. The protocol we describe below is inspired by Bell's original protocol. Property (iii) implies that we may restrict attention to rotationally invariant protocols, for which all probabilities depend only on $\hat{a} \cdot \hat{b}$ and not $\hat{a}$ and $\hat{b}$ separately, by *randomizing over all inputs with the same dot product.* More precisely, suppose P is any protocol that simulates the correlations. Then define a new protocol P' whose hidden variables consist of (i) those required by P, and (ii) a random rotation $R \in \mathrm{SO}(3)$. Protocol P' then consists of running protocol P on $R\hat{a}$ and $R\hat{b}$ in place of $\hat{a}$ and $\hat{b}$.

We now describe our protocol. Alice and Bob share two random variables $\hat{\lambda}_1$ and $\hat{\lambda}_2$, which are real three-dimensional unit vectors. They are chosen independently and distributed uniformly over the unit sphere.

The protocol proceeds as follows:

1. Alice outputs $\alpha = -\operatorname{sgn}(\hat{a} \cdot \hat{\lambda}_1)$.

2. Alice sends a single bit $c \in \{-1, +1\}$ to Bob where $c = \operatorname{sgn}(\hat{a} \cdot \hat{\lambda}_1) \operatorname{sgn}(\hat{a} \cdot \hat{\lambda}_2)$.

3. Bob outputs $\beta = \text{sgn}\left[\hat{b} \cdot (\hat{\lambda}_1 + c\hat{\lambda}_2)\right]$,

where we have used the sgn function defined by $\text{sgn}(x) = +1$ if $x \geq 0$ and $\text{sgn}(x) = -1$ if $x < 0$. A geometric description of our protocol is given in Fig. 5.1. We note immediately that Bob obtains *no information* about Alice's output from the communication.

We now prove that the protocol reproduces the correct expectation values. Each party's output changes sign under the symmetry $\hat{\lambda}_1 \leftrightarrow -\hat{\lambda}_1$, $\hat{\lambda}_2 \leftrightarrow -\hat{\lambda}_2$, so $\langle\alpha\rangle = \langle\beta\rangle = 0$ because $\hat{\lambda}_1$ and $\hat{\lambda}_2$ are uniformly distributed. The joint expectation value $\langle\alpha\beta\rangle$ can be calculated using

$$\langle\alpha\beta\rangle = E\left\{ -\text{sgn}(\hat{a} \cdot \hat{\lambda}_1) \sum_{d=\pm 1} \frac{(1+cd)}{2} \;\text{sgn}\left[\hat{b} \cdot (\hat{\lambda}_1 + d\hat{\lambda}_2)\right] \right\}, \tag{5.1}$$

where $E\{x\} = \frac{1}{(4\pi)^2} \int d\hat{\lambda}_1 \int d\hat{\lambda}_2 \, x$, $c = \text{sgn}(\hat{a} \cdot \hat{\lambda}_1)\,\text{sgn}(\hat{a} \cdot \hat{\lambda}_2)$ and we have used the trick that $(1+cd)/2 = 1$ if $c = d$ and 0 if $c \neq d$. After substituting for c and expanding Eq. (1), we obtain the sum of four terms (because each term inside the summation sign is itself the sum of two terms) and, using $\text{sgn}(\hat{a} \cdot \hat{\lambda}_1)\, c = \text{sgn}(\hat{a} \cdot \hat{\lambda}_2)$, we note that the four terms are related by the symmetries $\hat{\lambda}_1 \leftrightarrow \hat{\lambda}_2$ or $\hat{\lambda}_2 \leftrightarrow -\hat{\lambda}_2$, so each has the same expectation value. Hence

$$\langle\alpha\beta\rangle = E\left\{2 \;\text{sgn}(\hat{a} \cdot \hat{\lambda}_1) \;\text{sgn}\left[\hat{b} \cdot \left(\hat{\lambda}_2 - \hat{\lambda}_1\right)\right]\right\}. \tag{5.2}$$

This integral may be evaluated with the help of the two diagrams shown in Fig. 5.2, with the result that $\langle\alpha\beta\rangle = -\hat{a} \cdot \hat{b}$, as required.

Our protocol exactly simulates the quantum mechanical probability distribution for projective measurements on the singlet Bell pair state. If a large number of simulations is performed in parallel, the communication may be compressed. To see this, assume Alice's measurement vector $\hat{a}$ is uniformly distributed (if not, we randomize, as outlined above). Then, if $\hat{\lambda}_1 \cdot \hat{\lambda}_2 = \cos\eta$, Alice sends -1 with probability η/π and 1 with probability $1 - \eta/\pi$, so that the communication can be compressed to $\int_0^{\pi/2} \sin\eta \, d\eta H(\eta/\pi) \approx 0.85$ bits, where $H(\eta/\pi)$ is the Shannon entropy. This encoding depends on the shared unit vectors $\hat{\lambda}_1$ and $\hat{\lambda}_2$: A third party without access to the hidden variables will observe Alice sending uniformly distributed bits to Bob.

Our protocol is easily modified to simulate joint measurements on any maximally entangled state of two qubits, because every such state is related to the singlet by a local change of basis and thus may be simulated by rotating and/or reflecting the input vectors $\hat{a}$ and $\hat{b}$, before running our protocol.

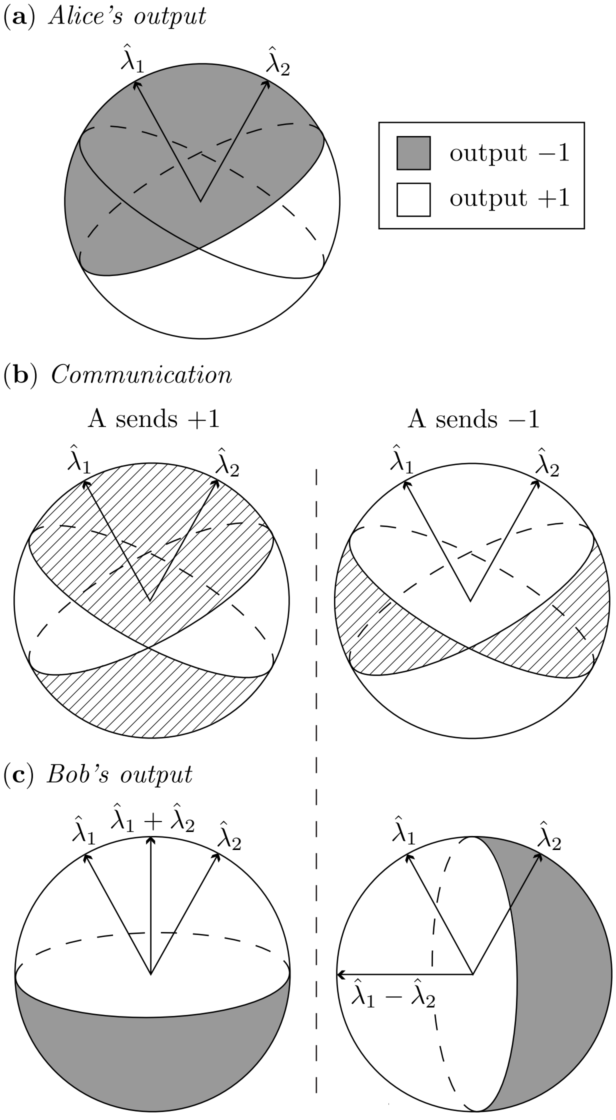
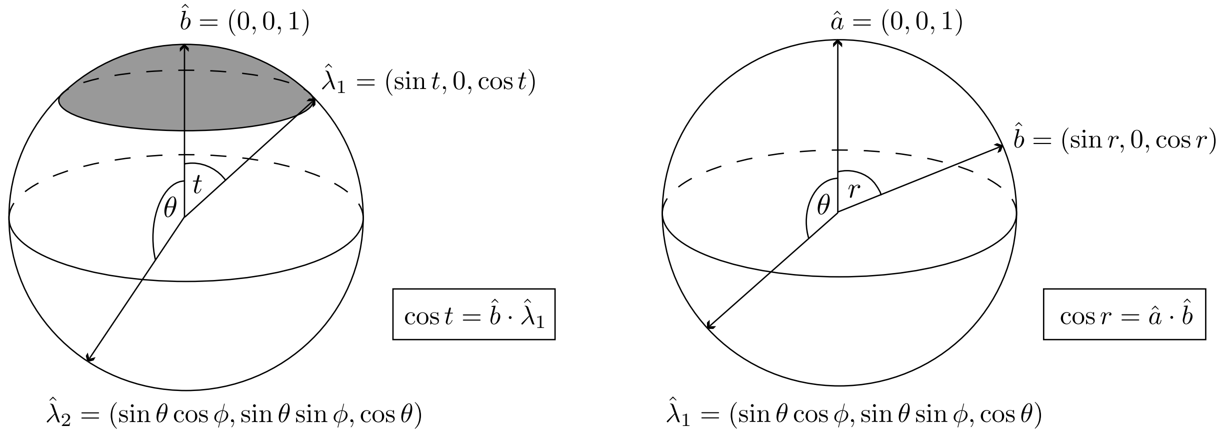


Figure 5.1: *The protocol*: The shared unit vectors $\hat{\lambda}_1$ and $\hat{\lambda}_2$ described in the text divide the Bloch sphere into four quadrants, as shown. Alice and Bob's actions depend on which quadrant their respective measurement axes lie in, and in Bob's case, the bit he receives from Alice. (**a**) *Alice's output*: if $\hat{a}$ lies in the shaded region, Alice outputs -1; in the unshaded region, she outputs $+1$. (**b**) *The communication*: Alice sends $c = +1$ if her measurement axis lies in the N or S quadrants, and -1 otherwise. (**c**) *Bob's output*: this depends on the bit received from Alice. The shading is as for (**a**).


$\hat{b} = (0,0,1)$
$\hat{\lambda}_1 = (\sin t, 0, \cos t)$
t
θ
$\cos t = \hat{b} \cdot \hat{\lambda}_1$
$\hat{\lambda}_2 = (\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta)$
$\hat{a} = (0,0,1)$
$\hat{b} = (\sin r, 0, \cos r)$
r
θ
$\cos r = \hat{a} \cdot \hat{b}$
$\hat{\lambda}_1 = (\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta)$

(**a**) $\dfrac{1}{4\pi}\displaystyle\int d\hat{\lambda}_2 \,\mathrm{sgn}\left[\hat{b}\cdot(\hat{\lambda}_2 - \hat{\lambda}_1)\right] = -\hat{b}\cdot\hat{\lambda}_1$

(**b**) $\dfrac{1}{2\pi}\displaystyle\int d\hat{\lambda}_1 \,\mathrm{sgn}(\hat{a}\cdot\hat{\lambda}_1)\hat{b}\cdot\hat{\lambda}_1 = \hat{a}\cdot\hat{b}$

Figure 5.2: *Construction used to evaluate Eq. (5.2)*: (**a**) We first integrate over $\hat{\lambda}_2$, taking $\hat{b}$ to point along the positive z-axis [67]. Observe that $\mathrm{sgn}\left[\hat{b}\cdot(\hat{\lambda}_2 - \hat{\lambda}_1)\right]$ is positive in the top spherical cap (shaded) and negative otherwise. The area of the top spherical cap is $A_+ = 2\pi\int_0^t \sin\theta d\theta = 2\pi(1 - \cos t)$ where $\cos t = \hat{b}\cdot\hat{\lambda}_1$, hence $\int d\hat{\lambda}_2 \ \mathrm{sgn}\left[\hat{b}\cdot(\hat{\lambda}_2 - \hat{\lambda}_1)\right] = A_+ - (4\pi - A_+) = -4\pi\cos t = -4\pi\hat{b}\cdot\hat{\lambda}_1$. (**b**) We now take $\hat{a}$ to point along the positive z-axis [68], set $\hat{b} = (\sin r, 0, \cos r)$, and integrate over $\hat{\lambda}_1$, obtaining $\int d\hat{\lambda}_1 \ \mathrm{sgn}(\hat{a}\cdot\hat{\lambda}_1)\hat{b}\cdot\hat{\lambda}_1 = \int_0^\pi \sin\theta d\theta \int_0^{2\pi} d\phi \ \ \mathrm{sgn}(\cos\theta)\left(\cos\theta\cos r + \sin\theta\cos\phi\sin r\right) = 2\pi\cos r = 2\pi\hat{a}\cdot\hat{b}$.

5.3 Application to teleportation experiments

Now consider the following experiment: Alice prepares a qubit in a state unknown to Bob. She then teleports the qubit to Bob, who performs a projective measurement on it, along a direction unknown to Alice. We shall show that this experiment admits a local hidden variables description. We first note that quantum teleportation experiments do not purport to test whether quantum mechanics allows a local hidden variables model; rather they aim to distinguish quantum teleportation from other protocols Alice and Bob might carry out using classical communication, but no entanglement [69]. From this point of view, teleportation experiments represent "investigations *within* quantum mechanics" [70], rather than comparisons of quantum mechanics with classical local hidden variables models [71]. With this distinction in mind, it is still interesting to ask whether teleportation experiments *can* be explained by a local hidden variables model.

If one allows an infinite amount of classical communication from Alice to Bob, then there is a trivial local hidden variables model, for Alice can just send a classical description of the state to Bob, who then simulates his measurement. We now give a local hidden variables model that requires only two bits of communication, which is the *same* amount as the quantum teleportation protocol. The construction is based on Ref. [54], where the procedure is termed "classical teleportation." It is sufficient to consider the case where Alice prepares the qubit in a pure state, which we suppose has spin aligned along the axis $\hat{a}$. We suppose Bob's measurement is aligned along the axis $\hat{b}$. Alice and Bob share uniformly distributed random three-dimensional unit vectors $\hat{\lambda}_1$ and $\hat{\lambda}_2$ (which can be thought of as hidden variables carried by the Bell pair used for teleportation). The protocol is as follows:

1. Alice sends $c_1 = \text{sgn}(\hat{a} \cdot \hat{\lambda}_1)$ and $c_2 = \text{sgn}(\hat{a} \cdot \hat{\lambda}_2)$ to Bob.

2. Bob outputs $\beta = \text{sgn}\left[\hat{b} \cdot \left(c_1\hat{\lambda}_1 + c_2\hat{\lambda}_2\right)\right]$.

It is easy to verify that $\langle\beta\rangle = \hat{a}\cdot\hat{b}$, as required. We also note that the two bits sent appear completely random to a party without access to the hidden variables.

It is usual in teleportation experiments to have (i) a third party Victor supply Alice with a quantum state unknown to her, and (ii) Bob hand off the teleported state to Victor (or another party) to measure, rather than measuring it himself. Such a distinction is not important for the question we address, because the qubit transmitted from Victor to Alice, for example, can carry hidden variables describing its state. The point is that local hidden variables are *hidden*: Although it is convenient to describe a local hidden variables model as if Alice and Bob had access to the hidden variables, the model still exists even if the hidden variables are inaccessible to them. There is no way for the experimenters to tell whether their experiment is described by quantum theory or by "gremlins" within their apparatus, executing the local hidden variables protocol described above.

Are there quantum teleportation experiments that do not have such a local hidden variables description? One obvious possibility is an experiment that teleports entanglement itself. But there is a more subtle possibility. If we allow Bob to measure the qubit using elements of a positive operator-valued measure, then there may not be a local hidden variables description that respects the two-bit classical communication bound. More generally, if Alice teleports n qubits (which requires $2n$ bits of communication) and Bob makes a joint measurement on them, then it is known that any exact local hidden variables theory requires that Alice send at least a constant times 2^n bits of communication in the worst case [52]. Whether this holds for protocols with bounded error is an important open question.

Finally, using the classical teleportation protocol, we obtain a (not necessarily optimal) protocol to simulate joint projective measurements on partially entangled states of two qubits, which uses two bits of communication: Alice first simulates her measurement and determines the post-measurement state of Bob's qubit; Alice and Bob then execute the classical teleportation protocol.

5.4 Conclusion

The results presented here offer an intriguing glimpse into the nature of correlations produced in quantum theory. If we interpret Bell inequality violation to mean that some communication is necessary to simulate Bell correlations, then our results prove that the minimal amount, one bit, is all that is necessary for projective measurements on Bell pairs. Is our straightforward protocol an indication of a deep structure in quantum correlations? We hope that our protocol and the development of a general theory of the communication cost of simulating quantum correlations will help shed light on this fundamental question.

Chapter 6

Communication cost of simulating the quantum joint correlation

6.1 Introduction

In this chapter, we show how to simulate exactly the quantum joint correlation of two-outcome measurements using five bits of communication. This combines ideas from Chapters 3 and 5. This chapter is joint work with Oded Regev.

Suppose that Alice and Bob share a state ρ on $\mathbb{C}^d \otimes \mathbb{C}^d$, on which they perform local two-outcome measurements. Label Alice's observable by $\mathbf{A}$, Bob's by $\mathbf{B}$, Alice's outcome by α, and Bob's by β. Recall that the joint correlation is

$$\langle \alpha\beta \rangle = \operatorname{tr}\left(\mathbf{A} \otimes \mathbf{B}\rho\right). \tag{6.1}$$

Let $n = 2d^2$. Let $\hat{a} \in \mathbb{R}^n$ be the vector associated with Alice's measurement in accordance with Lemma 2.3.1, and let $\hat{b} \in \mathbb{R}^n$ be the vector associated with Bob's measurement.

6.2 A communication primitive

We first present an LHV model augmented by five bits of communication that does not reproduce the correct correlations, but which we shall use as a primitive in a protocol that does.

Protocol 6.2.1. (Random Variables) Alice and Bob share unit vectors $\hat{\lambda}_i \in \mathbb{R}^n$ for $i = 1, 2, \ldots, 5$, chosen independently and uniformly at random from the unit sphere, together with a bit c chosen uniformly at random from $\{-1, +1\}$.

(Alice) Let $\alpha_i = \operatorname{sgn}(\hat{a} \cdot \hat{\lambda}_i)$ for $i = 1, 2, \ldots, 5$. Alice outputs c and sends to Bob the five bits α_i ($i = 1, 2, \ldots, 5$).

(Bob) Let $\beta_i = \operatorname{sgn}(\hat{b} \cdot \hat{\lambda}_i)$ for $i = 1, 2, \ldots, 5$. Let $\gamma = \mathrm{MAJ}(\alpha_1\beta_1, \alpha_2\beta_2, \ldots, \alpha_5\beta_5)$ where MAJ is the majority function. Bob outputs $\beta = \gamma c$.

It is useful to define some functions. Let $f : [-1, 1] \to [0, 1]$ be defined by

$$f(x) = \frac{1}{\pi} \cos^{-1} x. \tag{6.2}$$

Let $g : [0, 1] \to [-1, +1]$ be defined by

$$g(y) = 1 - 20y^3 + 30y^4 - 12y^5. \tag{6.3}$$

Let $h(x) = g \circ f(x)$.

Lemma 6.2.2. *Protocol 6.2.1 gives correlations*

$$\langle \alpha\beta \rangle = h\left(\hat{a} \cdot \hat{b}\right) \tag{6.4}$$

Proof. Because the unit vectors $\hat{\lambda}_i$ are chosen independently, the events $\alpha_i = \beta_i$ are independent for different i. Let $p = \Pr(\alpha_i \neq \beta_i) = (1 - \langle \alpha_i \beta_i \rangle)/2$, which is independent of i. By Lemma 3.2.2, $p = f(\hat{a} \cdot \hat{b})$. Thus

$$\Pr(\gamma = -1) = p^5 + 5p^4(1-p) + 10p^3(1-p)^2. \tag{6.5}$$

Noting that $\langle \alpha\beta \rangle = 1 - 2\Pr(\gamma = -1)$ completes the proof. ■

6.3 Simulation of the joint correlation using five bits of communication

Just as in Chapter 3, we need to do some preliminary work before we can state the protocol.

Lemma 6.3.1. *Let c_{2k+1} be the coefficients in a series expansion of $h(x)$ about $x = 0$, i.e.,*

$$h(x) = \sum_{k=0}^{\infty} c_{2k+1} x^{2k+1}. \tag{6.6}$$

Then $c_1 > 0$ and $c_{2k+1} < 0$ for all $k > 0$.

We prove this lemma in Appendix B.

Let d_{2k+1} be the coefficients in a series expansion of $h^{-1}(x)$ about $x = 0$, i.e.,

$$h^{-1}(x) = \sum_{k=0}^{\infty} d_{2k+1} x^{2k+1}. \tag{6.7}$$

We note that $\sum_k d_{2k+1} = h^{-1}(1) = 1$. In fact, all the d_{2k+1} are positive. This follows from

Lemma 6.3.2. *Let $f : [-1, +1] \to [-1, +1]$ be a function and let f^{-1} be its inverse. Suppose that $f(x) = \sum_{k=1}^{\infty} c_k x^k$ and $f^{-1}(y) = \sum_{k=1}^{\infty} d_k y^k$, with $c_1 > 0$ and $c_k \leq 0$ for all $k > 1$. Then $d_k > 0$ for all k.*

Proof. There is an explicit formula for d_k (see, e.g., Ref. [72]):

$$d_k = \frac{1}{nc_1^n} \sum_{s,t,u,\ldots} (-1)^{s+t+u+\cdots} \frac{n(n+1)\cdots(n-1+s+t+u+\cdots)}{s!t!u!\cdots} \left(\frac{c_2}{c_1}\right)^s \left(\frac{c_3}{c_1}\right)^t \cdots, \tag{6.8}$$

where $s + 2t + 3u + \cdots = n - 1$. Then every term in the sum is nonnegative, from which it follows that $d_k \geq 0$ for all k. ■

Our simulation procedure will be as follows: We map $\hat{a}$ and $\hat{b}$ to new vectors, $C(\hat{a})$ and $C(\hat{b})$, respectively, which live in a much larger space. We then run Protocol 6.2.1 on $C(\hat{a})$ and $C(\hat{b})$. The trick is in choosing an appropriate function C.

To this end, let $C : \mathbb{R}^n \to \bigoplus_{k=0}^{\infty} (\mathbb{R}^n)^{\otimes(2k+1)}$. The range of C a direct sum of tensor products of $\mathbb{R}^n$. We write $C(\vec{v}) = \bigoplus_{k=0}^{\infty} C^{2k+1}(\vec{v})$, and term the functions $C^{2k+1}(\vec{v})$ "coordinates" of $C(\vec{v})$.

Define

$$C^{2k+1}(\vec{v}) = \sqrt{d_{2k+1}}\, \vec{v}^{\,\otimes(2k+1)}, \tag{6.9}$$

where $\vec{v}^{\,\otimes(2k+1)}$ denotes the vector $\vec{v} \otimes \vec{v} \otimes \cdots \otimes \vec{v}$ with $2k+1$ tensor factors. Note that this is well defined, since $d_{2k+1} \geq 0$ for all k.

The LHV model is as follows:

Protocol 6.3.3. (Random Variables) Alice and Bob share an infinite sequence $\lambda_1, \lambda_2, \ldots$ of real numbers, where each λ_i is drawn from a normal distribution with mean 0 and standard deviation 1. They also share a bit c chosen uniformly at random from $\{-1, +1\}$. For $i = 1, 2, \ldots, 5$, we write $\vec{\lambda}_i = (\lambda_i, \lambda_{i+5}, \lambda_{i+10} \ldots) \in l_\infty$.

(Alice) Let $\alpha_i = \mathrm{sgn}(C(\hat{a}) \cdot \hat{\lambda}_i)$ for $i = 1, 2, \ldots, 5$. Alice outputs c and sends to Bob the five bits α_i $(i = 1, 2, \ldots, 5)$.

(Bob) Let $\beta_i = \mathrm{sgn}(C(\hat{b}) \cdot \hat{\lambda}_i)$ for $i = 1, 2, \ldots, 5$. Let $\gamma = \mathrm{MAJ}(\alpha_1\beta_1, \alpha_2\beta_2, \ldots, \alpha_5\beta_5)$ where MAJ is the majority function. Bob outputs $\beta = \gamma c$.

Theorem 6.3.4. *Protocol 6.3.3 results in correlations*

$$\langle \alpha\beta \rangle = \hat{a} \cdot \hat{b}. \tag{6.10}$$

Proof. In order to apply Lemma 6.2.2, we have to check that $C(\hat{v})$ is a unit vector whenever $\hat{a}$ is.

This is straightforward:

$$C(\hat{v}) \cdot C(\hat{v}) = \sum_{k=0}^{\infty} d_{2k+1} \hat{v}^{\otimes(2k+1)} \cdot \hat{v}^{\otimes(2k+1)} = \sum_{k=0}^{\infty} d_{2k+1} = h^{-1}(1) = 1. \tag{6.11}$$

Lemma 6.2.2 now implies that Protocol 6.3.3 results in correlations

$$\langle \alpha\beta \rangle = h\left(C(\hat{a}) \cdot C(\hat{b}) \right). \tag{6.12}$$

But

$$C(\hat{a}) \cdot C(\hat{b}) = \sum_{k=0}^{\infty} d_{2k+1} \hat{a}^{\otimes(2k+1)} \cdot \hat{b}^{\otimes(2k+1)} = \sum_{k=0}^{\infty} d_{2k+1} \left(\hat{a} \cdot \hat{b} \right)^{2k+1} = h^{-1} \left(\hat{a} \cdot \hat{b} \right). \tag{6.13}$$

It follows that Protocol 3.4.1 results in correlations

$$\langle \alpha\beta \rangle = h\left(C(\hat{a}) \cdot C(\hat{b}) \right) = h \circ h^{-1} \left(\hat{a} \cdot \hat{b} \right) = \hat{a} \cdot \hat{b}, \tag{6.14}$$

which concludes the proof. ■

6.4 Discussion

The protocol presented above made use of 5 bits of communication. We have another protocol that only requires two bits of communication, which we shall present at a later date. In the previous chapter, we saw that one bit of communication sufficed in the case $n = 3$. We suspect that one bit is insufficient in the limit $n \to \infty$.

We have yet to explore other applications of this technique. For example, Alon and Naor have presented an approximation algorithm for the cut-norm of a matrix that is based on Krivine's proof of Grothendieck's inequality [41]. Perhaps this extension of Krivine's technique has a similar application.

Part II

Monogamy of Nonlocal Correlations

Chapter 7

Monogamy of nonlocal no-signaling correlations

7.1 Introduction

One of the remarkable properties of quantum entanglement is that it is monogamous: If Alice (A), Bob (B), and Charlie (C) each have a qubit, and A and B are maximally entangled, then C's qubit is completely uncorrelated with either A's or B's. This property is inherently nonclassical: If A, B, and C have bits instead of qubits, and A's bit is always the same as B's bit, then there is no restriction on how A's bit is correlated with C's bit. In this work, we consider the correlations that result from making local measurements on a multipartite quantum system. Some such quantum correlations violate Bell inequalities [7]. We show how these correlations, termed *nonlocal*, can also be monogamous.

Consider, for example, the well-known Clauser-Horne-Shimony-Holt (CHSH) inequality [25]. Two parties, A and B, share a quantum state ρ, and each chooses one of two observables to measure on their component of the state. Define the *CHSH operator*

$$\mathcal{B}_{\mathrm{CHSH}} = \mathbf{A}_1 \otimes (\mathbf{B}_1 + \mathbf{B}_2) + \mathbf{A}_2 \otimes (\mathbf{B}_1 - \mathbf{B}_2), \tag{7.1}$$

where $\mathbf{A}_1$ and $\mathbf{A}_2$ ($\mathbf{B}_1$ and $\mathbf{B}_2$) are A's (B's) observables and are Hermitian operators with spectrum in $[-1,+1]$. Then the CHSH inequality states that $|\langle\mathcal{B}_{\mathrm{CHSH}}\rangle_{\mathrm{LHV}}| \leq 2$, for all local hidden variable (LHV) models, but there are observables on the singlet state of two qubits $|\psi^-\rangle = (|01\rangle - |10\rangle)/\sqrt{2}$, such that $\langle\mathcal{B}_{\mathrm{CHSH}}\rangle_{\mathrm{QM}} = \mathrm{tr}\,(\mathcal{B}_{\mathrm{CHSH}}|\psi^-\rangle\langle\psi^-|) = 2\sqrt{2}$. Thus these correlations cannot be described by an LHV model. In fact, it is true that $|\mathrm{tr}\,(\mathcal{B}_{\mathrm{CHSH}}\rho)| \leq 2\sqrt{2}$ for all observables $\mathbf{A}_1$, $\mathbf{A}_2$, $\mathbf{B}_1$, $\mathbf{B}_2$, and all states ρ. Such a bound on the maximum quantum value of a Bell inequality is termed a Tsirelson bound [28]. Although we do not know how to calculate the best such bound for an arbitrary Bell inequality, a number of ad hoc techniques have been developed [28, 29, 31].

In this chapter, we introduce a new technique for obtaining Tsirelson bounds. Since local measurements on spatially separated components of a multipartite quantum system can be carried out simultaneously, such measurements cannot be used to send a signal from one party to another. The outcomes of local measurements on an entangled quantum state are therefore described by a *no-signaling* probability distribution. Maximization over no-signaling probability distributions can be cast as a linear program, and so we obtain an upper bound by solving this linear program. (For the CHSH inequality there is a no-signaling probability distribution such that $\langle \mathcal{B}_{\mathrm{CHSH}} \rangle_{\mathrm{NS}} = 4$, so this technique is not useful [17]).

Suppose three parties, A, B, and C, share an entangled quantum state of arbitrary dimension. We use this technique to (i) bound the trade-off between (A and B) and (A and C)'s violation of the CHSH inequality, and (ii) demonstrate that forcing B and C to be classically correlated prevents A and B from violating the odd cycle Bell inequality of Ref. [29].

7.2 Framework

We cast our results in the language of *cooperative games of incomplete information*, also called nonlocal games [29]. Let $V : \mathbb{Z}_2^m \times \mathbb{Z}_n^m \to [0, 1]$ be a function and let π be a probability distribution on $\mathbb{Z}_n^m$. (A game defined in this way can be converted to a "standard" game (where V is a predicate) by adding questions.) Then V and π define a m-player nonlocal game $G(V, \pi)$ as follows: A referee chooses a set of questions $(q_1, q_2, \ldots, q_m) \in \mathbb{Z}_n^m$ randomly, according to π, and sends question q_i to player i. Each player must answer with a bit a_i. The players are not permitted to communicate after receiving the questions, but they may agree on a strategy before receiving them. They win with probability $V(a_1, a_2, \ldots, a_m | q_1, q_2, \ldots, q_m)$ (where the $|$ in $V(\cdot|\cdot)$ separates answers from questions). The *classical value* of a game $G(V, \pi)$, denoted $\omega_c(G)$, is the maximum probability with which the players can win, assuming they use purely classical strategies. The *quantum value*, denoted $\omega_q(G)$, is the maximum winning probability, assuming they are allowed to share entanglement. The *no-signaling value*, denoted $\omega_{ns}(G)$, is the maximum winning probability, assuming the players are allowed (black box) access to any no-signaling probability distribution. It is clear that $\omega_c(G) \leq \omega_q(G) \leq \omega_{ns}(G)$.

7.2.1 The CHSH game

We describe how to interpret the CHSH inequality within this framework. The CHSH game G_{CHSH} is defined by setting $n = 2$, letting π be the uniform distribution on $\mathbb{Z}_2 \times \mathbb{Z}_2$ and letting $V(a_1, a_2 | q_1, q_2) = [a_1 \oplus a_2 = q_1 \wedge q_2]$, where $a_1 \oplus a_2$ is the "exclusive-or" of bits a_1 and a_2, $q_1 \wedge q_2$ is the "and" of bits q_1 and q_2, and $[\phi]$ is 1 if ϕ is true and 0 otherwise. Then the winning probability of a particular strategy is $1/2 + \langle \mathcal{B}_{\mathrm{CHSH}} \rangle / 8$, where $\mathcal{B}_{\mathrm{CHSH}}$ is the CHSH operator of Eq. (7.1) and $\langle \cdot \rangle$ is the appropriate expectation value for the strategy. It follows that $\omega_c(G_{\text{CHSH}}) = 3/4$, $\omega_q(G_{\text{CHSH}}) = 1/2 + 1/(2\sqrt{2}) \approx 0.85$, and $\omega_{ns}(G_{\text{CHSH}}) = 1$.

7.3 Main technique

An m-party no-signaling probability distribution is a set of probabilities $p(\{a_i\}_{i=1}^m | \{q_i\}_{i=1}^m)$, subject to

1. *Positivity*: $p(\{a_i\}|\{q_i\}) \geq 0$;
2. *Normalization*: For all $\{q_i\}$, $\sum_{\{a_i\}} p(\{a_i\}|\{q_i\}) = 1$;
3. *No-signaling*: For each subset $S \subset \mathbb{Z}_m$ of the m parties, the marginal probability distribution on $\mathbb{Z}_m - S$ must be independent of the inputs of the parties in S. In particular, $\sum_{\{a_i : i \in S\}} p(\{a_i\}|\{q_i\})$ must be independent of $\{q_i : i \in S\}$ for all $\{a_i : i \notin S\}$ and for all $\{q_i : i \notin S\}$.

The no-signaling value of G is given by

$$\omega_{ns}(G) = \max_p \sum_{\{a_i\},\{q_i\}} \pi(\{q_i\}) V(\{a_i\}|\{q_i\}) p(\{a_i\}|\{q_i\}),$$

subject to the three sets of linear constraints enumerated above. We observe that $\omega_{ns}(G)$ is the solution to a linear program in variables $p(\{a_i\}|\{q_i\})$. Solving this program for $\omega_{ns}(G)$ gives an upper bound on $\omega_q(G)$. Moreover, even if we cannot solve the linear program, we can obtain an upper bound on $\omega_{ns}(G)$ by constructing a solution to the dual program (see Ref. [73] for an introduction to convex optimization).

7.4 Applications

7.4.1 An analogue of the CKW theorem for nonlocal quantum correlations

Suppose three parties, A, B, and C, each have a qubit. There is a well-known theorem of Coffman, Kundu, and Wootters (CKW) that describes the trade-off between how entangled A is with B, and how entangled A is with C. It states that $\mathcal{C}_{AB}^2 + \mathcal{C}_{AC}^2 \leq 4 \det \rho_A$, where $\mathcal{C}_{AB}$ is the concurrence between A and B, $\mathcal{C}_{AC}$ is the concurrence between A and C, and ρ_A is the reduced density matrix of A [74].

To derive a similar expression for correlations, we consider a generalization of the CHSH game to three players, suggested by Michael Nielsen (see also [75]). In the game G'_{CHSH}, the referee sends bits chosen uniformly at random to each of the three players, and with probability 1/2 checks if $a_1 \oplus a_2 = q_1 \wedge q_2$ and with probability 1/2 checks if $a_1 \oplus a_3 = q_1 \wedge q_3$. Formally, π is uniform on $\mathbb{Z}_2^3$ and $V(a_1, a_2, a_3|q_1, q_2, q_3) = [a_1 \oplus a_2 = q_1 \wedge q_2]/2 + [a_1 \oplus a_3 = q_1 \wedge q_3]/2$. Then the winning probability of a particular strategy is $1/2 + \langle \mathcal{B}_{\text{CHSH}}^{\text{AB}} \rangle /16 + \langle \mathcal{B}_{\text{CHSH}}^{\text{AC}} \rangle /16$, where the superscripts denote on which parties the CHSH operator acts. It is easy to see that $\omega_c(G'_{\text{CHSH}}) = 3/4$ (a strategy where everyone always answers 0 achieves this, and this strategy is the best possible, by the CHSH inequality applied to AB and BC separately). It turns out that $\omega_{ns}(G'_{\text{CHSH}}) = 3/4$ too, as is easily verified using linear programming software, which implies the following:

Theorem 7.4.1. *Suppose three parties, A, B, and C, share no-signaling correlations, where each chooses to measure one of two observables. Then*

$$\left|\langle \mathcal{B}_{CHSH}^{AB} \rangle\right| + \left|\langle \mathcal{B}_{CHSH}^{AC} \rangle\right| \leq 4. \tag{7.2}$$

Theorem 7.4.1 establishes a trade-off between AB's and AC's violation of the CHSH inequality. In particular, CHSH correlations are monogamous: If AB violate the CHSH inequality, then AC cannot, as has been shown independently for no-signaling correlations by Masanes, Acín, and Gisin [76]. (In fact, Theorem 7.4.1 may be obtained easily from Result 3 of Ref. [76] by symmetrization of B and C.) Note that if AB and AC each share an EPR pair, there are measurements such that either $\operatorname{tr}\left(\mathcal{B}_{\text{CHSH}}^{\text{AB}} \rho\right)$ or $\operatorname{tr}\left(\mathcal{B}_{\text{CHSH}}^{\text{AC}} \rho\right)$ is $2\sqrt{2}$, which at first appears to contradict Theorem 1. It does not: In Theorem 1 we insist that A's observables are the *same* in $\operatorname{tr}\left(\mathcal{B}_{\text{CHSH}}^{\text{AB}} \rho\right)$ and $\operatorname{tr}\left(\mathcal{B}_{\text{CHSH}}^{\text{AC}} \rho\right)$. This is in the same spirit as the requirement of CKW that B and C are entangled with the same qubit of A. One can generalize this result:

Corollary 7.4.2. *Suppose $N+2$ parties A, B_0, $B_1, \ldots$, B_N share a quantum state and each chooses to measure one of two observables. Then A violates the CHSH inequality with at most one of the B_i.*

Proof. Suppose A violates the CHSH inequality with both B_j and B_k, $j \neq k$. Trace out the rest of the B_i's. We obtain a contradiction with Theorem 7.4.1. ■

For no-signaling probability distributions, we also have a converse of Theorem 7.4.1: For any pair $\left(\langle \mathcal{B}_{\text{CHSH}}^{\text{AB}} \rangle, \langle \mathcal{B}_{\text{CHSH}}^{\text{AC}} \rangle\right)$ consistent with Ineq. (7.2), there is a no-signaling probability distribution with these expectation values. This is because we can write $\left(\langle \mathcal{B}_{\text{CHSH}}^{\text{AB}} \rangle, \langle \mathcal{B}_{\text{CHSH}}^{\text{AC}} \rangle\right)$ as a convex combination of $(4, 0)$, $(0, 4)$ and $(0, 0)$, each of which is achieved by a no-signaling probability distribution. Thus Ineq. (7.2) establishes precisely which $\left(\langle \mathcal{B}_{\text{CHSH}}^{\text{AB}} \rangle, \langle \mathcal{B}_{\text{CHSH}}^{\text{AC}} \rangle\right)$ are allowed. For quantum theory, it turns out that the allowed region is described by $\langle \mathcal{B}_{\text{CHSH}}^{\text{AB}} \rangle^2 + \langle \mathcal{B}_{\text{CHSH}}^{\text{AC}} \rangle^2 \leq 8$, as we shall see in the next chapter.

7.4.2 Classical correlation restricts Bell inequality violation

Consider a two-player game $G(V,\pi)$ being played by A and B$_0$. We show that forcing B$_0$ to be *classically* correlated with additional players B$_1$, B$_2$, $\dots$, and B$_N$ restricts the advantage that A and B_0 can gain by sharing entanglement. For $N \geq 1$, define the Nth *extension* of a two-player game $G(V,\pi)$ to be the $N+2$ player game $G_N(V_N, \pi_N)$, with π_N defined by choosing (q_1, q_2) according to π and setting $q_2 = q_3 = q_4 = \cdots = q_{N+2}$; and $V_N(\{a_i\}|\{q_i\}) = V(a_1, a_2|q_1, q_2) \times [a_2 = a_3 = a_4 = \cdots = a_{N+2}]$. We also set $G_0 = G$. The idea is that we send B$_0$'s question to the other B$_i$'s and the players win if (i) the answers A and B$_0$ give satisfy the winning condition of $G(V,\pi)$ and (ii) all the B$_i$'s agree.

Theorem 7.4.3. *Let $G(V,\pi)$ be a two-player game. Then the values of its extensions satisfy (i) $\omega_c(G_N) = \omega_c(G)$ for all N, and (ii) $\omega_{ns}(G_N)$ is a nonincreasing sequence in N, with $\omega_{ns}(G_{n-1}) = \omega_c(G)$, where n is the number of questions for B_0 in G.*

Proof. To prove (i), we observe that we can assume an optimal classical strategy for G is deterministic, which immediately gives a strategy for G_N with the same value. Conversely, we may convert a strategy for G_N to one for G by ignoring B$_1$ through B$_N$. Hence $\omega_c(G_N) = \omega_c(G)$. To prove (ii), it is clear that $\omega_{ns}(G_N)$ is a nonincreasing sequence in N, because any strategy for G_N gives a strategy for G_M, $M < N$, by ignoring B$_{M+1}$ through B$_N$. Finally, consider a no-signaling strategy for G_{n-1}, with probabilities $p\left(\{a_i\}|\{q_i\}\right)$. We define a classical strategy for G as follows: A and B$_0$ share bits c_0, c_2, $\dots$, c_{n-1}, drawn according to the distribution $\sum_{a_1} p(a_1, a_2 = c_0, a_3 = c_1, \dots, a_{n+1} = c_{n-1}|q_1, q_2 = 0, q_3 = 1, \dots, q_{n+1} = n-1)$, which is guaranteed to be independent of q_1 because p is no-signaling. Then if A and B$_0$ are asked questions q_1 and q_2, B$_0$ answers c_{q_2} and A answers with a bit a_1 drawn from the distribution

$$p_A(a_1) = \frac{p(a_1, c_0, c_1, \dots, c_{n-1}|q_1, 0, 1, \dots, n-1)}{\sum_{a_1'} p(a_1', c_0, c_1, \dots, c_{n-1}|q_1, 0, 1, \dots, n-1)}.$$

One can check that this classical strategy for G wins at least as often as that of the no-signaling strategy for G_{n-1}. Hence $\omega_{ns}(G_{n-1}) = \omega_c(G)$. ■

A similar result was already known for $\omega_q(G_N)$, and the proof is analogous [77]. Theorem 7.4.3 places a limit on how polygamous the no-signaling violation of a Bell inequality can be.

7.4.3 The odd cycle game

Our final example is taken from Ref. [29], and illustrates how violation of a Bell inequality can preclude classical correlation with another party. We start with a two-player game based on an interactive proof for graph colorability. Imagine that the two players, A and B, are trying to convince the referee that an odd cycle of length n is two-colorable (which it is not, as n is odd). The referee sends them each the name of a vertex, such that the two vertices are either the same or adjacent. A and B each send one of two colors back to the referee. The referee requires that, when the vertices are the same, the two colors should agree and, when the vertices are adjacent, the colors should differ. Formally, we define a two-player game G_{OC} as follows: Let $n \geq 3$ be an odd integer, let π be uniform over the set $\{(q_1, q_2) \in \mathbb{Z}_n \times \mathbb{Z}_n \,:\, q_1 = q_2 \text{ or } q_1 + 1 \equiv q_2 \pmod{n}\}$ and let V be defined by $V(a_1, a_2|q, q) = [a_1 = a_2]$, $V(a_1, a_2|q, q+1) = [a_1 \neq a_2]$. It is established in Refs. [78, 29] that $\omega_c(G_{\mathrm{OC}}) = 1 - 1/2n$ and $\omega_q(G_{\mathrm{OC}}) = \cos^2(\pi/4n)$.

Now consider the first extension of this game, which we denote G'_{OC}. Formally, G'_{OC} is defined by using the same distribution as G_{OC} on (q_1, q_2), and setting $q_2 = q_3$. The function V is defined by $V(a_1, a_2, a_3|q, q, q) = [a_1 = a_2 = a_3]$, $V(a_1, a_2, a_3|q, q+1, q+1) = [a_1 \neq a_2 = a_2]$. Theorem 7.4.3 implies that $\omega_c(G'_{\mathrm{OC}}) = \omega_c(G_{\mathrm{OC}}) = 1 - 1/2n$. We shall show that $\omega_{ns}(G'_{\mathrm{OC}}) = 1 - 1/2n$ also. This result is remarkable because it establishes that adding just one additional player is sufficient to prevent A and B from gaining advantage by sharing entanglement, rather than the $n-1$ additional players required in Theorem 7.4.3.

There are a number of symmetries we can use to simplify the problem. Without changing the probability of winning: (i) all parties can flip their outputs, or (ii) all parties can add $(\bmod\, n)$ an integer m to their inputs, or (iii) B and C can exchange roles. For a given no-signaling strategy, let $p(a, b, c|i, j, k)$ be the probability that (A,B,C) answer (a, b, c) when asked (i, j, k). Then we can take $p(a, b, c|i, j, k)$ to be symmetric under these three symmetries. In particular, symmetry (i) implies we can restrict attention to $a = 0$, symmetry (ii) to $i = 0$. Therefore, let $r(b, c|j, k) = p(0, b, c|0, j, k)$. We shall use symmetry (iii) to give extra constraints, rather than to reduce the number of parameters.

We rewrite the primary linear program in these variables, labeling the constraints. Our goal is to maximize

$$\omega_{ns}(G'_{\mathrm{OC}}) = \frac{1}{2} \max_r \left[r(0, 0|0, 0) + r(1, 1|1, 1)\right], \tag{7.3}$$

subject to:

- (Normalization) $n(j, k)$: $\sum_{b,c} r(b, c|j, k) = 1$, for $0 \leq j, k < n$.

- (Symmetry) $s(b, c|j, k)$: $r(b, c|j, k) = r(c, b|k, j)$, for $b, c \in \{0, 1\}$, $0 \leq j, k < n$. Note that when $b = c$ and $j = k$ this constraint is trivial.

- (No-signaling conditions, A to BC) $y(d|j, k)$: $p(0, d|j, j + k) + p(1, \bar{d}|j, j + k) = p(0, d|0, k) + p(1, \bar{d}|0, k)$, for $d \in \{0, 1\}$, $1 \leq j < n$, $0 \leq k < n$, where the sum $j + k$ is taken mod n.

- (No-signaling conditions, B to AC) $z(d|j, k)$: $p(0, d|j, k) + p(1, d|j, k) = p(0, d|0, k) + p(1, d|0, k)$, for $d \in \{0, 1\}$, $1 \leq j < n$, $0 \leq k < n$, where the sum $j + k$ is taken mod n.

We omit the no-signaling conditions in the other directions (BC to A and AC to B), which do not further constrain the solution.

Each constraint in the primary linear program corresponds to a variable in the dual, as labeled above. The objective of the dual program is to minimize

$$\frac{1}{2n}\sum_{j,k} n(j,k), \tag{7.4}$$

subject to the constraints $\mu(0,0|0,0), \mu(1,1|1,1) \geq n$, $\mu(b,c|j,k) \geq 0$, for all $b,c \in \{0,1\}$, $0 \leq j,k < n$, where

$$\begin{aligned} \mu(b,c|j,k) &= n(j,k) + s(b,c|j,k) - s(c,b|k,j) \\ &\quad + [j=0] \sum_{j'=1}^{n-1} \left(y\left(\frac{1-bc}{2}\middle| j',k\right) + z(c|j',k) \right) \\ &\quad - [j \neq 0] \left(y\left(\frac{1-bc}{2}\middle| j,k-j\right) + z(c|j,k) \right). \end{aligned} \tag{7.5}$$

We now give an explicit solution to the dual. The nonzero variables are:

$$\begin{aligned} n(0,0) &= 2n-1; \\ s(0,1|0,0) &= 3n/2, \\ s(0,1|1,0) &= -n+1, \\ s(0,0|0,1) &= -n+1, \\ s(0,1|1,1) &= -n/2, \\ s(0,0|j,j+1) &= (-1)^j \text{ for } j = 1,2,\ldots,n-1; \\ s(0,1|j,j+1) &= -(-1)^j \text{ for } j = 1,2,\ldots,n-1; \end{aligned}$$

$$\begin{aligned} y(0|1,0) &= -2n+3, \\ y(0|1,k) &= -n+k+5/2+(-1)^k/2 \\ &\quad \text{for } k = 1,2,\ldots,n-1; \\ y(1|1,0) &= 3-3n/2, \\ y(1|1,1) &= -n+4, \\ y(1|j,1) &= -(-1)^j \text{ for } j = 2,3,\ldots,n-1, \\ y(1|1,k) &= -n+k+5/2+(-1)^k/2 \\ &\quad \text{for } k = 2,3,\ldots,n-2, \end{aligned}$$

$$\begin{aligned} y(1|1, n-1) &= -n+3, \\ y(1|j, n-1) &= 1-(-1)^j \text{ for } j = 2, 3, \ldots, n-1; \end{aligned}$$

$$\begin{aligned} z(0|1,0) &= n-3, \\ z(0|1,1) &= 2n-3, \\ z(0|1,2) &= n-4, \\ z(0|j,j-1) &= -1 \text{ for } j = 2, 3, \ldots, n-1, \\ z(0|j,j+1) &= (-1)^j \text{ for } j = 2, 3, \ldots, n-1, \\ z(0|1,k) &= n-k-3/2+(-1)^k/2 \\ & \quad \text{for } k = 3, 4, \ldots, n-1; \\ z(1|j,j-1) &= -1+(-1)^j \text{ for } j = 1, 2, \ldots, n-1, \\ z(1|1,k) &= n-k-3/2+(-1)^k/2 \\ & \quad \text{for } k = 1, 2, \ldots, n-1. \end{aligned}$$

All other variables are zero.

For this solution, it's tedious but straightforward to establish that $\mu(0,0|0,0) = \mu(1,1|1,1) = n$, $\mu(0,1|0,0) = 2n$, $\mu(1,0|0,1) = 2n-2$, $\mu(0,0|j,j-1) = 1+(-1)^j$ for $j = 1, 2, \ldots, n-1$, $\mu(1,1|k+1,k) = 1+(-1)^k$ for $k = 1, 2, \ldots, n-1$, and $\mu(b,c|j,k) = 0$, otherwise. Thus our solution satisfies the constraints. This solution was constructed by solving the linear program for small n, finding a consistent set of equality conditions by generalizing from the small n case, and inverting these constraints to yield the solution. Substituting into Eq. (7.4), we find that $\omega_{ns}(G'_{\mathrm{OC}}) \leq 1-1/2n$, which proves the following theorem:

Theorem 7.4.4. *For the first extension of the odd cycle game,* $\omega_c(G'_{OC}) = \omega_q(G'_{OC}) = \omega_{ns}(G'_{OC}) = 1-1/2n$.

Thus sharing entanglement (or indeed no-signaling correlations) gives no advantage for G'_{OC}. In the context of interactive proof systems, we can interpret the fact that $\omega_q(G_{\mathrm{OC}}) > \omega_c(G_{\mathrm{OC}})$ in the two-player game as saying that sharing entanglement allows the provers to cheat, because it increases the probability with which they are able to convince the referee that the odd cycle is two-colorable. Theorem 7.4.4 shows that we can counter this by adding an extra prover, and forcing B to be classically correlated with her. This placed no extra burden on classical provers, because an optimal classical strategy is deterministic, but it prevents quantum provers from gaining any advantage by sharing entanglement. We hope that a similar approach will help in determining the power of the complexity classes QMIP and MIP* [29].

This result also has applications to cryptography [79, 80]. Suppose that A and B are trying to share a secret key, and that C is eavesdropping on them. If A and B observe correlations that would cause them to win the two-player odd cycle game with probability greater than $1 - 1/2n$, then this limits how correlated C can be with B. Indeed, Barrett, Hardy, and Kent have presented a key distribution protocol along these lines [79], which is provably secure against no-signaling eavesdroppers.

Chapter 8

Monogamy of nonlocal quantum correlations

8.1 Introduction

Suppose three parties, A, B, and C, share any quantum state ρ (of arbitrary dimension) and each chooses to measure one of two observables. In the previous chapter, we proved Theorem 7.4.1, which established a trade-off between AB and AC's violation of the CHSH inequality, viz.:

$$\left|\left\langle \mathcal{B}^{\mathrm{AB}}_{\mathrm{CHSH}} \right\rangle\right| + \left|\left\langle \mathcal{B}^{\mathrm{AC}}_{\mathrm{CHSH}} \right\rangle\right| \leq 4. \tag{8.1}$$

The proof of Theorem 7.4.1 only required that the correlations be no-signaling. In this chapter, we prove a stronger monogamy trade-off relationship for correlations that are realizable with quantum resources. This chapter is joint work with Frank Verstraete.

We shall only be concerned with the CHSH inequality in this chapter, so we drop the subscript indicating the Bell inequality: $\langle \mathcal{B}_{\mathrm{AB}} \rangle \equiv \left\langle \mathcal{B}^{\mathrm{AB}}_{\mathrm{CHSH}} \right\rangle$. We establish the following monogamy trade-off relation:

Theorem 8.1.1. *Suppose that three parties, A, B, and C, share a quantum state (of arbitrary dimension) and each chooses to measure one of two observables. Then*

$$\langle \mathcal{B}_{AB} \rangle^2 + \langle \mathcal{B}_{AC} \rangle^2 \leq 8. \tag{8.2}$$

The important point is that we obtain Tsirelson's bound, $\langle \mathcal{B}_{\mathrm{AB}} \rangle^2 \leq 8$, as a simple corollary. Eq. (8.2) is the best possible bound: There are states and measurements achieving any values of $\langle \mathcal{B}_{\mathrm{AB}} \rangle$ and $\langle \mathcal{B}_{\mathrm{AC}} \rangle$ that satisfy it. We illustrate the monogamy trade-offs for various theories in Fig. 8.1.

We prove Theorem 8.1.1 in two parts. We first show that is sufficient to restrict to states with support on a qubit at each site. We can then relax the requirement that A's measurements be the same in $\langle \mathcal{B}_{\mathrm{AB}} \rangle$ and $\langle \mathcal{B}_{\mathrm{AC}} \rangle$, maximizing over the measurements in $\langle \mathcal{B}_{\mathrm{AB}} \rangle$ and $\langle \mathcal{B}_{\mathrm{AC}} \rangle$ separately, but keeping the state fixed.

8.2 Dimensional reduction

We start by establishing a bound on the dimension of the quantum state required to maximally violate certain Bell inequalities. This result was originally proved by Masanes [81]. The main ingredient—a canonical decomposition for a pair of subspaces of $\mathbb{C}^n$—is described in more detail in, e.g., Ref. [82].

Lemma 8.2.1. *Consider any Bell inequality in the setting where m parties each choose from two two-outcome measurements. Then the maximum quantum value of the Bell inequality is achieved by a state that has support on a qubit at each site. Furthermore, we can assume this state has real coefficients and that the observables are real and traceless.*

Proof. For $i \in \{1, 2\}$, assume party k has observables $M_{k,i}$, acting on a Hilbert space $\mathcal{H}_k$. By extending the local Hilbert spaces $\mathcal{H}_k$, we can assume for all k and for all $i = 1, 2$ that (i) $\mathcal{H}_k = \mathbb{C}^{2d}$ for some fixed d, (ii) $M_{k,i}$ has eigenvalues ± 1, and (iii) $\operatorname{tr} M_{k,i} = 0$. The first condition states that all local spaces have the same dimension $2d$, the latter two that each observable corresponds to a projective measurement onto a d-dimensional subspace and its complement. We also define $M_{k,0} = \mathbb{1}_{2d}$, the identity operator on $\mathcal{H}_k$. We can write a generic Bell operator in the setting stated in the lemma as

$$\mathcal{B} = \sum_{i_1=0}^{2} \sum_{i_2=0}^{2} \cdots \sum_{i_m=0}^{2} c_{i_1 i_2 \cdots i_m} \bigotimes_{k=1}^{m} M_{k,i_k}, \tag{8.3}$$

where the coefficients $c_{i_1 i_2 \cdots i_m}$ are arbitrary real numbers. Our goal is find the quantum value of this Bell operator, which is maximum of $B \equiv \langle \psi | \mathcal{B} | \psi \rangle$ over states $|\psi\rangle$ and measurements $M_{k,i}$.

We now choose a local basis for each $\mathcal{H}_k$ such that party k's observables have a simple form. We start by taking $M_{k,1} = \left[\begin{array}{c:c} \mathbb{1}_d & 0 \\ \hdashline 0 & -\mathbb{1}_d \end{array}\right]$. This leaves us the freedom to specify the basis within the two $d \times d$ blocks on which $M_{k,1}$ is constant. Let $M_{k,2} = 2PP^\dagger - \mathbb{1}_{2d}$ (we suppress the dependence on k), where P is a $2d \times d$ matrix with orthonormal columns, which span the $+1$–eigenspace of $M_{k,2}$. Write $P = \left[\begin{array}{c} P_1 \\ \hdashline P_2 \end{array}\right]$, where P_1 and P_2 are $d \times d$ matrices. The rows of P are orthonormal, which implies $P^\dagger P = P_1^\dagger P_1 + P_2^\dagger P_2 = \mathbb{1}_d$, so $P_1^\dagger P_1$ and $P_2^\dagger P_2$ are simultaneously diagonalizable. This means there is a singular value decomposition of the form

$$P_1 = U_1^\dagger D_1 V, \tag{8.4}$$

$$P_2 = U_2^\dagger D_2 V, \tag{8.5}$$

where U_1, U_2 and V are $d \times d$ unitary matrices and D_1 and $D_2 = \sqrt{\mathbb{1}_d - D_1^2}$ are nonnegative (real) diagonal matrices. Changing basis according to the unitary $U_1 \oplus U_2$, which leaves $M_{k,1}$ invariant, it follows that $M_{k,2} = \left[\begin{array}{c:c} 2D_1^2 - \mathbb{1}_d & 2D_1D_2 \\ \hdashline 2D_1D_2 & 2D_2^2 - \mathbb{1}_d \end{array}\right]$, where each of the $d \times d$ blocks is diagonal. We relabel our basis vectors so that $M_{k,1} = \bigoplus_{j=1}^d Z$, $M_{k,2} = \bigoplus_{j=1}^d (\cos\theta_j Z + \sin\theta_j X)$, where $2D_1^2 - \mathbb{1}_d = \operatorname{diag}(\cos\theta_1, \cos\theta_2, \ldots, \cos\theta_d)$ and X and Z are the usual Pauli operators. Hence our operators are real and preserve a $\oplus_{j=1}^d \mathbb{C}^2$ subspace of $\mathcal{H}_k$. They are traceless on each $\mathbb{C}^2$ space.

We wish to maximize $B = \langle\psi|\mathcal{B}|\psi\rangle$ over the state $|\psi\rangle$ and the measurements $M_{k,i}$. Fix k, and let $\rho_{k,j}$ be the reduced density matrix obtained by projecting $|\psi\rangle$ onto the j'th $\mathbb{C}^2$ factor of the $\oplus_{j=1}^d \mathbb{C}^2$ subspace induced by $M_{k,1}$ and $M_{k,2}$ at site k. Then $B = \sum_{j=1}^d \operatorname{tr} \mathcal{B}\rho_{k,j}$ is a convex sum over the $\mathbb{C}^2$ factors, whereupon it follows that the maximum is achieved by a state with support on a qubit at site k. Since this argument works for all k, the maximum of B is achieved by a state that has support on a qubit on each site.

Finally, write $|\psi\rangle = |\psi_1\rangle + i|\psi_2\rangle$, where $|\psi_1\rangle$ and $|\psi_2\rangle$ are real. Then $\langle\psi|\mathcal{B}|\psi\rangle = \langle\psi_1|\mathcal{B}|\psi_1\rangle + \langle\psi_2|\mathcal{B}|\psi_2\rangle$ since $\mathcal{B}$ is real, which is the same expression we would obtain if the state were a real mixture of $|\psi_1\rangle$ and $|\psi_2\rangle$. Hence the maximum of B is achieved by a state with real coefficients. ■

8.3 Monogamy trade-off relation

The region $\mathcal{R}$ of allowed values of $(\langle\mathcal{B}_{\mathrm{AB}}\rangle, \langle\mathcal{B}_{\mathrm{AC}}\rangle)$ is convex and can therefore be described by an (infinite) family of half-space inequalities,

$$c_{\mathrm{AB}}\langle\mathcal{B}_{\mathrm{AB}}\rangle + c_{\mathrm{AC}}\langle\mathcal{B}_{\mathrm{AC}}\rangle \leq d, \tag{8.6}$$

with $c_{\mathrm{AB}}, c_{\mathrm{AC}}, d \in \mathbb{R}$. The left-hand side of Eq. (8.6) is a Bell operator, as defined in Eq. (8.3), which means we can apply Lemma 8.2.1 to conclude that extreme points of $\mathcal{R}$ are achieved by real states on three qubits, with measurements of the form $M = \cos\theta Z + \sin\theta X$. Theorem 8.1.1 will emerge as a corollary of:

Lemma 8.3.1. *Let $|\psi\rangle$ be a pure state in $\mathbb{C}^2 \otimes \mathbb{C}^2 \otimes \mathbb{C}^2$ with real coefficients. Then the maximum of $\langle\mathcal{B}_{\mathrm{AB}}\rangle = \langle\psi|\mathcal{B}_{\mathrm{AB}}|\psi\rangle$ over real traceless observables A_1, A_2, B_1, B_2 is*

$$\max_{A_i, B_j} \langle\mathcal{B}_{\mathrm{AB}}\rangle = 2\sqrt{1 + \langle Y_{\mathrm{A}}Y_{\mathrm{B}}\rangle^2 - \langle Y_{\mathrm{A}}Y_{\mathrm{C}}\rangle^2 - \langle Y_{\mathrm{B}}Y_{\mathrm{C}}\rangle^2}, \tag{8.7}$$

where Y is the usual Pauli operator, $\langle Y_{\mathrm{A}}Y_{\mathrm{B}}\rangle = \operatorname{tr}(Y_A \otimes Y_B \otimes \mathbb{1}\,\rho)$, and so on. Cyclic permutations of Eq. (8.7) hold for $\langle\mathcal{B}_{\mathrm{AC}}\rangle$ and $\langle\mathcal{B}_{\mathrm{BC}}\rangle$.

Proof. We consider $\rho_{AB} = \operatorname{tr}_C |\psi\rangle\langle\psi|$, which is a real state on $\mathbb{C}^2 \otimes \mathbb{C}^2$. Horodecki and family have calculated the maximum quantum value of the CHSH operator for a state on $\mathbb{C}^2 \otimes \mathbb{C}^2$ [37]. Their analysis simplifies in our case because the state and measurements are real. Define

$$T_{\text{AB}} = \begin{bmatrix} \langle X_A X_B \rangle & \langle X_A Z_B \rangle \\ \langle Z_A X_B \rangle & \langle Z_A Z_B \rangle \end{bmatrix}. \tag{8.8}$$

For $i = 1, 2$, write $A_i = \hat{a}_i \cdot \vec{\sigma}_r$, $B_i = \hat{b}_i \cdot \vec{\sigma}_r$, where $\hat{a}_i$ and $\hat{b}_i$ are two-dimensional unit vectors and $\vec{\sigma}_r = (X, Z)$. Define

$$\hat{b}_1 + \hat{b}_2 = 2\cos\theta \hat{d}_1, \quad \hat{b}_1 - \hat{b}_2 = 2\sin\theta \hat{d}_2, \tag{8.9}$$

where $\theta \in [0, \pi/2]$ and $\hat{d}_1$ and $\hat{d}_1$ are orthogonal unit vectors. Then

$$\begin{aligned} \frac{1}{2} \max_{A_i, B_j} \langle \mathcal{B}_{\text{AB}} \rangle &= \max_{\hat{d}_i, \theta, \hat{a}_i} \cos\theta \hat{a}_1^t T_{\text{AB}} \hat{d}_1 + \sin\theta \hat{a}_2^t T_{\text{AB}} \hat{d}_2 && (8.10) \\ &= \max_{\hat{d}_i, \theta} \cos\theta \left\| T_{\text{AB}} \hat{d}_1 \right\| + \sin\theta \left\| T_{\text{AB}} \hat{d}_2 \right\| && (8.11) \\ &= \max_{\hat{d}_i} \sqrt{\left\| T_{\text{AB}} \hat{d}_1 \right\|^2 + \left\| T_{\text{AB}} \hat{d}_2 \right\|^2} && (8.12) \\ &= \sqrt{\operatorname{tr}\left(T_{\text{AB}} T_{\text{AB}}{}^t \right)}. && (8.13) \end{aligned}$$

This is just the Frobenius norm of T_{AB} and it is straightforward to check that, for pure states on $\mathbb{C}^2 \otimes \mathbb{C}^2 \otimes \mathbb{C}^2$ with real coefficients, it is equal to half the right-hand side of Eq. (8.7). ■

Lemma 8.3.2. *For a pure state $|\psi\rangle$ with real coefficients in $\mathbb{C}^2 \otimes \mathbb{C}^2 \otimes \mathbb{C}^2$,*

$$\max_{A_i, B_j, C_k} \langle \mathcal{B}_{\text{AB}} \rangle^2 + \langle \mathcal{B}_{\text{AC}} \rangle^2 = 8\left(1 - \langle Y_{\text{B}} Y_{\text{C}} \rangle^2\right). \tag{8.14}$$

Proof. Lemma 8.3.1, applied to $\langle \mathcal{B}_{\text{AB}} \rangle$ and $\langle \mathcal{B}_{\text{AC}} \rangle$ separately, immediately implies:

$$\begin{aligned} \max_{A_i, B_j, C_k} \langle \mathcal{B}_{\text{AB}} \rangle^2 + \langle \mathcal{B}_{\text{AC}} \rangle^2 &\leq \max_{A_i, B_j} \langle \mathcal{B}_{\text{AB}} \rangle^2 + \max_{A_i, C_k} \langle \mathcal{B}_{\text{AC}} \rangle^2 \\ &= 8\left(1 - \langle Y_{\text{B}} Y_{\text{C}} \rangle^2\right). && (8.15) \end{aligned}$$

The reason we do not have equality is that the measurements A_i achieving the maximum in $\langle \mathcal{B}_{\text{AB}} \rangle$ and $\langle \mathcal{B}_{\text{AC}} \rangle$ may be different. We have to show they can be chosen to be the same. Define T_{AC} in analogy with Eq. (8.8) and write the vectors corresponding to C's measurements as

$$\hat{c}_1 + \hat{c}_2 = 2\cos\theta \hat{e}_1, \quad \hat{c}_1 - \hat{c}_2 = 2\sin\theta \hat{e}_2, \tag{8.16}$$

in analogy with Eq. (8.9) for B's observables. One can check that $[T_{\text{AB}} T_{\text{AB}}^t, T_{\text{AC}} T_{\text{AC}}^t] = 0$ for all pure states $|\psi\rangle$ with real coefficients. Hence there are orthonormal vectors a'_1 and a'_2 that are simultaneous eigenvectors of $T_{\mathrm{AB}}T^t_{\mathrm{AB}}$ and $T_{\mathrm{AC}}T^t_{\mathrm{AC}}$. Next, note that the term being maximized in Eq. (8.12), $\|T_{\mathrm{AB}}\hat{d}_1\|^2 + \|T_{\mathrm{AB}}\hat{d}_2\|^2$, is actually independent of the $\hat{d}_i$ (recall that $\hat{d}_1 \cdot \hat{d}_2 = 0$), so we are free to choose the $\hat{d}_i$ as we please. Take $\hat{d}_i = T^t_{\mathrm{AB}}\hat{a}'_i$ for $i = 1, 2$ and, similarly, take $\hat{e}_i = T^t_{\mathrm{AC}}\hat{a}'_i$. Alice's measurement vector $\hat{a}_i$ in the AB maximization of the previous lemma was taken to be the unit vector along $T_{\mathrm{AB}}\hat{d}_i$, but this is $T_{\mathrm{AB}}T^t_{\mathrm{AB}}\hat{a}'_i \propto \hat{a}'_i$ so $\hat{a}_i = \hat{a}'_i$. The same will hold in the AC maximization. Hence we can choose A's measurement vectors to be the same in both cases, and we have equality in Eq. (8.14). ■

Combining Lemmas 8.2.1 and 8.3.2, we obtain Theorem 8.1.1. Lemma 8.3.2 also implies that any $\langle\mathcal{B}_{\mathrm{AB}}\rangle$ and $\langle\mathcal{B}_{\mathrm{AC}}\rangle$ compatible with Eq. (8.14) are achievable. In particular, the state

$$|\psi\rangle = c_-\left(|010\rangle + |011\rangle\right) + c_+\left(|100\rangle + |101\rangle\right), \tag{8.17}$$

where

$$c_\pm = \frac{1}{2}\sqrt{1 \pm \sqrt{2}\sin t}, \tag{8.18}$$

and $0 \leq t \leq \pi/4$, gives $\langle\mathcal{B}_{\mathrm{AB}}\rangle = 2\sqrt{2}\,\cos t$, $\langle\mathcal{B}_{\mathrm{AC}}\rangle = 2\sqrt{2}\,\sin t$.

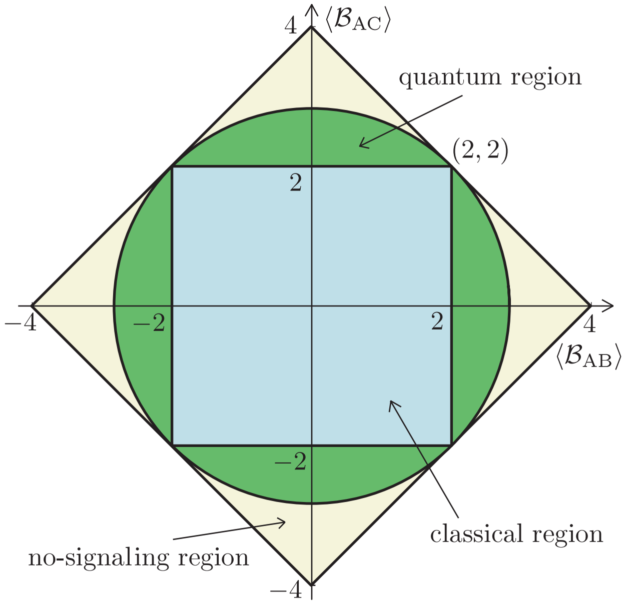

4
$\langle \mathcal{B}_{\mathrm{AC}} \rangle$
quantum region
(2, 2)
2
−4
−2
2
4
$\langle \mathcal{B}_{\mathrm{AB}} \rangle$
−2
classical region
no-signaling region
−4

Figure 8.1: Accessible values of $\langle \mathcal{B}_{\mathrm{AB}} \rangle$ and $\langle \mathcal{B}_{\mathrm{AC}} \rangle$ for classical theories (interior of square), quantum theory (interior of circle), and no-signaling theories (interior of diamond).

Appendix A

Calculation of the constant c_3

In Section 3.5, we left the identity in Eq. (3.37),

$$I = \sqrt{\frac{\pi}{2x}} \sum_{k=0}^{\infty} (4k+3) J_{2k+3/2}(x) = \frac{\sqrt{x}}{2} \int_0^x t^{-3/2} \sin t \, dt, \tag{A.1}$$

unproven. We now prove it. Although this identity would follow from Krivine's Theorem 3.6.8, we choose to prove the identity directly, so that our analysis of the three-dimensional case is self-contained.

Our starting point is the series representation of $J_\nu(x)$:

$$J_\nu(x) = \sum_{n=0}^{\infty} \frac{(-1)^n}{n!\,\Gamma(n+\nu+1)} \left(\frac{x}{2}\right)^{2n+\nu}. \tag{A.2}$$

Substituting into Eq. (A.1),

$$\begin{aligned} I &= \frac{\sqrt{\pi}}{2} \sum_{k=0}^{\infty} \sum_{n=0}^{\infty} (4k+3) \frac{(-1)^n}{n!\,\Gamma(n+2k+5/2)} \left(\frac{x}{2}\right)^{2n+2k+1} & \text{(A.3)} \\ &= \sum_{l=0}^{\infty} \frac{(-1)^l \sqrt{\pi}}{2^{2(l+1)}} \left(\sum_{k=0}^{l} \frac{(4k+3)(-1)^k}{(l-k)!\Gamma(l+k+5/2)} \right) x^{2l+1}, & \text{(A.4)} \end{aligned}$$

where we have reordered the sums and substituted $l = n + k$.

To calculate the coefficient of x^{2l+1} we note that

$${}_2F_1(1, -l; 5/2+l; t) = l!\,\Gamma(l+5/2) \sum_{k=0}^{l} \frac{(-1)^k}{(l-k)!\,\Gamma(k+l+5/2)} t^k, \tag{A.5}$$

where ${}_2F_1(a, b; c; t)$ is the hypergeometric function. Differentiating Eq. (A.5) w.r.t. t, we obtain

$$\begin{aligned} l!\,\Gamma(l+5/2) \sum_{k=0}^{l} \frac{(4k+3)(-1)^k}{(l-k)!\Gamma(l+k+5/2)}\, x^k & \qquad \text{(A.6)} \\ = 4t\,{}_2F_1'(1, -l; 5/2+l; t) + 3\,{}_2F_1(1, -l; 5/2+l; t) & \qquad \text{(A.7)} \\ = 4t\frac{-l}{5/2+l}\,{}_2F_1(2, -l+1; 7/2+l; t) + 3\,{}_2F_1(1, -l; 5/2+l; t). & \qquad \text{(A.8)} \end{aligned}$$

Substituting $t = 1$ and using Gauss's hypergeometric theorem

$${}_2F_1(a, b; c; 1) = \Gamma(c)\Gamma(c-a-b)/\Gamma(c-a)\Gamma(c-b) \qquad \text{(A.9)}$$

yields

$$I = \sum_{l=0}^{\infty} \frac{(-1)^l}{(2l+1)!(4l+1)}\, x^{2l+1}. \qquad \text{(A.10)}$$

To convert this into the integral representation given in Eq. (A.1), we start with the power series expression for $\sin t$, multiply by $t^{-3/2}$, then integrate term by term.

Appendix B

Signs of the derivatives of $h(x)$

Here we prove Lemma 6.3.1:

Lemma B.0.3 (6.3.1). *Let c_{2k+1} be the coefficients in a series expansion of*

$$h(x) = 1 - \frac{2\left(\cos^{-1} x\right)^3}{\pi^5}\left(10\pi^2 - 15\pi\cos^{-1} x + 6\left(\cos^{-1} x\right)^2\right). \tag{B.1}$$

about $x = 0$, i.e.,

$$h(x) = \sum_{k=0}^{\infty} c_{2k+1} x^{2k+1}. \tag{B.2}$$

Then $c_1 > 0$ and $c_{2k+1} < 0$ for all $k > 0$.

Proof. We calculate $c_1 = h'(0) = 15/4\pi > 0$. For the other coefficients, consider the second derivative

$$h''(x) = -\frac{60}{\pi^5} h_1(x) h_2(x), \tag{B.3}$$

where

$$h_1(x) = \frac{\frac{\pi^2}{4} - \left(\sin^{-1} x\right)^2}{\sqrt{1-x^2}}, \tag{B.4}$$

$$h_2(x) = 4\frac{\sin^{-1} x}{\sqrt{1-x^2}} - \frac{x}{\sqrt{1-x^2}} h_1(x). \tag{B.5}$$

In Lemma B.0.5 we show that $h_1(x) = \sum_k a_{2k} x^{2k}$ with $a_{2k} > 0$ for all k, and in Lemma B.0.6 that $h_2(x) = \sum_k b_{2k+1} x^{2k+1}$ with $b_{2k+1} > 0$ for all k. This implies that all the coefficients in the series expansion of $h''(x)$ are negative, i.e., $c_{2k+1} < 0$ for all $k > 0$. ■

Lemma B.0.4 (Some series expansions)**.** *The following are valid for* $|x| < 1$*:*

$$\begin{aligned}
\frac{1}{1-x^2} &= \sum_{k=0}^{\infty} x^{2k}, && \text{(B.6)}\\
\frac{1}{\sqrt{1-x^2}} &= \sum_{k=0}^{\infty} d_{2k} x^{2k} = \sum_{k=0}^{\infty} \frac{(2k-1)!!}{(2k)!!} x^{2k}, && \text{(B.7)}\\
\left(\sin^{-1} x\right)^2 &= \sum_{k=0}^{\infty} f_{2k+2} x^{2k+2} = \sum_{k=0}^{\infty} \frac{(2k)!!}{(2k+1)!!(k+1)} x^{2k+2}, && \text{(B.8)}\\
\left(\sin^{-1} x\right)^3 &= \sum_{k=0}^{\infty} g_{2k+3} x^{2k+3} = 6\sum_{k=0}^{\infty} \frac{(2k+1)!!^2}{(2k+3)!} \left(\sum_{j=0}^{k} \frac{1}{(2j+1)^2}\right) x^{2k+3}. && \text{(B.9)}
\end{aligned}$$

The third and fourth series expansions are Eqs. (1.645.2) and (1.645.3) of Ref. [83].

Lemma B.0.5. *Let* $h_1(x) = \sum_k a_{2k} x^{2k}$*. Then* $a_{2k} > 0$ *for all* k*.*

Proof. We write

$$\begin{aligned}
h_1(x) &= \frac{\pi^2}{4}\frac{1}{\sqrt{1-x^2}} - \frac{1}{3}\frac{d}{dx}\left(\sin^{-1} x\right)^3 && \text{(B.10)}\\
&= \frac{\pi^2}{4} d_0 + \sum_{k=1}^{\infty} \left(\frac{\pi^2}{4} d_{2k} - \frac{2k+1}{3}\, g_{2k+1}\right) x^{2k}, && \text{(B.11)}
\end{aligned}$$

which immediately gives $a_0 = \pi^2 d_0/4 > 0$. For $k > 0$,

$$\begin{aligned}
a_{2k} &= \frac{\pi^2}{4} d_{2k} - \frac{2k+1}{3}\, g_{2k+1} && \text{(B.12)}\\
&= \frac{2(2k-1)!!}{(2k)!!}\left(\frac{\pi^2}{8} - \sum_{j=0}^{k-1} \frac{1}{(2j+1)^2}\right) && \text{(B.13)}\\
&> \frac{2(2k-1)!!}{(2k)!!}\left(\frac{\pi^2}{8} - \sum_{j=0}^{\infty} \frac{1}{(2j+1)^2}\right) && \text{(B.14)}\\
&= 0. && \text{(B.15)}
\end{aligned}$$

Thus all the coefficients in the series expansion of $h_1(x)$ are positive. ■

Lemma B.0.6. *Let* $h_2(x) = \sum_k b_{2k+1} x^{2k+1}$*. Then* $b_{2k+1} > 0$ *for all* k*.*

Proof. We have

$$\begin{aligned} h_2(x) &= -\frac{\pi^2}{4}\frac{x}{1-x^2} + \frac{x\left(\sin^{-1}x\right)^2}{1-x^2} + 2\frac{d}{dx}\left(\sin^{-1}x\right)^2 && \text{(B.16)} \\ &= -\frac{\pi^2}{4}\sum_{k=0}^{\infty} x^{2k+1} + \sum_{k=0}^{\infty}\sum_{j=0}^{\infty} f_{2k+2}x^{2(k+j)+3} + 2\sum_{k=0}^{\infty}(2k+2)f_{2k+2}x^{2k+1} && \text{(B.17)} \\ &= \sum_{k=0}^{\infty}\left[4(k+1)f_{2k+2} - \frac{\pi^2}{4} + \sum_{j=0}^{k-1} f_{2j+2}\right]x^{2k+1}, && \text{(B.18)} \end{aligned}$$

so that for all $k \geq 0$,

$$\begin{aligned} b_{2k+1} &= 4(k+1)f_{2k+2} - \frac{\pi^2}{4} + \sum_{j=0}^{k-1} f_{2j+2} && \text{(B.19)} \\ &= 4(k+1)f_{2k+2} - \sum_{j=k}^{\infty} f_{2j+2} \qquad \left(\text{since } \frac{\pi^2}{4} = \left(\sin^{-1}1\right)^2\right) && \text{(B.20)} \\ &= \frac{4(2k)!!}{(2k+1)!!} - \sum_{j=k}^{\infty}\frac{(2j)!!}{(2j+1)!!(j+1)}. && \text{(B.21)} \end{aligned}$$

Let

$$S_k = \sum_{j=k}^{\infty} s_{k,j} = \frac{(2k+1)!!}{(2k)!!}\sum_{j=k}^{\infty}\frac{(2j)!!}{(2j+1)!!(j+1)}. \tag{B.22}$$

Then $S_0 = \pi^2/4 < 4$ and $s_{k,j} = s_{k-1,j}(1+2k)/(2k)$. Summing this over $j = k, k+1, \ldots$, we obtain the recurrence

$$S_k = \frac{1+2k}{2k}\left(S_{k-1} - \frac{1}{k}\right). \tag{B.23}$$

We need to show that $S_k < 4$ for all $k > 0$. To this end, suppose $S_{k-1} > 3 + 1/k$ for some k. Then

$$S_k - S_{k-1} = \frac{S_{k-1}}{2k} - \frac{1+2k}{2k^2} > \frac{1}{2k}, \tag{B.24}$$

which implies that S_k is unbounded as $k \to \infty$. We obtain a contradiction by calculating the limit of S_k as $k \to \infty$. This is straightforward: For large j,

$$\frac{(2j)!!}{(2j+1)!!} \sim \frac{1}{2}\sqrt{\frac{\pi}{j}}, \tag{B.25}$$

so that

$$\lim_{k\to\infty} S_k = \sqrt{k}\int_{x=k}^{\infty}\frac{1}{x^{3/2}}\,dx = 2. \tag{B.26}$$

Hence $S_k < 4$ for all k, which implies $b_{2k+1} > 0$ for all k. ■

Bibliography

[1] A. Acin, N. Gisin, and B. Toner, "Grothendieck's constant and local models for noisy entangled quantum states," *Phys. Rev. A*, vol. 73, no. 6, p. 062105, 2006.

[2] D. Bacon and B. F. Toner, "Bell inequalities with auxiliary communication," *Phys. Rev. Lett.*, vol. 90, p. 157904, 2003.

[3] B. F. Toner and D. Bacon, "Communication cost of simulating Bell correlations," *Phys. Rev. Lett.*, vol. 91, p. 187904, 2003.

[4] J. A. Reeds, "A new lower bound on the real Grothendieck constant," 1991, unpublished note, available at http://www.dtc.umn.edu/~reedsj/bound2.dvi.

[5] A. M. Davie, "Lower bound for K_G," 1984, unpublished note.

[6] A. Einstein, P. Podolsky, and N. Rosen, "Can quantum-mechanical description of physical reality be considered complete?" *Phys. Rev.*, vol. 47, pp. 777–780, 1935.

[7] J. S. Bell, "On the Einstein-Podolsky-Rosen paradox," *Physics*, vol. 1, pp. 195–200, 1964.

[8] A. Aspect, P. Grangier, and G. Roger, "Experimental realization of Einstein-Podolsky-Rosen-Bohm gedankenexperiment: A new violation of Bell's inequalities," *Phys. Rev. Lett.*, vol. 49, no. 2, pp. 91–94, 1982.

[9] A. Aspect, J. Dalibard, and G. Roger, "Experimental test of Bell's inequalities using time-varying analyzers," *Phys. Rev. Lett.*, vol. 49, no. 25, pp. 1804–1807, 1982.

[10] W. Tittel, J. Brendel, H. Zbinden, and N. Gisin, "Violation of Bell inequalities by photons more than 10 km apart," *Phys. Rev. Lett.*, vol. 81, no. 17, pp. 3563–3566, 1998.

[11] G. Weihs, T. Jennewein, C. Simon, H. Weinfurter, and A. Zeilinger, "Violation of Bell's inequality under strict Einstein locality conditions," *Phys. Rev. Lett.*, vol. 81, no. 23, pp. 5039–5043, 1998.

[12] M. A. Rowe, D. Kielpinski, V. Meyer, C. A. Sackett, W. M. Itano, C. Monroe, and D. J. Wineland, "Experimental violation of a Bell's inequality with efficient detection," *Nature*, vol. 409, pp. 791–4, 2001.

[13] P. W. Shor, "Algorithms for quantum computation: discrete logarithms and factoring," in *Proceedings of the 35th Annual Symposium on the Foundations of Computer Science.* Los Alamitos, CA: IEEE, 1994, pp. 124–134.

[14] R. Raz, "Exponential separation of quantum and classical communication complexity," in *Proceedings of the 31st ACM Symposium on Theory of Computing.* New York: ACM Press, 1999, pp. 358–367.

[15] C. H. Bennett and G. Brassard, "Quantum cryptography: Public key distribution and coin tossing," in *Proceedings of IEEE International Conference on Computers, Systems, and Signal Processing.* IEEE, 1984, pp. 175–179.

[16] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information.* New York: Cambridge University Press, 2000.

[17] S. Popescu and D. Rohrlich, "Nonlocality as an axiom," *Found. Phys.*, vol. 24, p. 379, 1994.

[18] L. A. Khalfin and B. S. Tsirelson, "Quantum and quasi-classical analogs of Bell inequalities," in *Symposium on the Foundations of Modern Physics*, P. Lahti and P. Mittelstaedt, Eds. Singapore: World Scientific, 1985, pp. 441–460.

[19] R. F. Werner and M. M. Wolf, "Bell inequalities and entanglement," *Quantum Information and Computation*, vol. 1, no. 3, pp. 1–25, 2001.

[20] R. T. Rockafellar, *Convex Analysis.* Princeton, New Jersey: Princeton University Press, 1970.

[21] A. Fine, "Hidden variables, joint probability, and the Bell inequalities," *Phys. Rev. Lett.*, vol. 48, no. 5, pp. 291–295, 1982.

[22] I. Pitowsky and K. Svozil, "Optimal tests of quantum nonlocality," *Phys. Rev. A*, vol. 64, no. 1, p. 014102, 2001.

[23] D. Collins and N. Gisin, "A relevant two qubit Bell inequality inequivalent to the CHSH inequality," *J. Phys. A: Math. Gen.*, vol. 37, pp. 1775–1787, 2004.

[24] A. Garg and N. D. Mermin, "Farkas's lemma and the nature of reality: statistical implications for quantum correlations," *Found. Phys.*, vol. 14, pp. 1–39, 1984.

[25] J. F. Clauser, M. A. Horne, A. Shimony, and R. A. Holt, "Proposed experiment to test local hidden-variable theories," *Phys. Rev. Lett.*, vol. 23, pp. 880–884, 1969.

[26] W. van Dam, "Implausible consequences of superstrong nonlocality," 2005, quant-ph/0501159.

[27] G. Brassard, H. Buhrman, N. Linden, A. A. Methot, A. Tapp, and F. Unger, “Limit on nonlocality in any world in which communication complexity is not trivial,” *Phys. Rev. Lett.*, vol. 96, no. 25, p. 250401, 2006, quant-ph/0508042.

[28] B. S. Cirel’son, “Quantum generalizations of Bell’s inequality,” *Lett. Math. Phys.*, vol. 4, pp. 93–100, 1980.

[29] R. Cleve, P. Høyer, B. Toner, and J. Watrous, “Consequences and limits of nonlocal strategies,” in *Proceedings of the 19th IEEE Conference on Computational Complexity (CCC 2004)*, 2004, pp. 236–249, quant-ph/0404076.

[30] W. van Dam, R. D. Gill, and P. D. Grünwald, “The statistical strength of nonlocality proofs,” *IEEE Trans. Inf. Th.*, vol. 51, no. 8, pp. 2812–2835, 2005.

[31] B. S. Tsirelson, “Quantum analogues of the Bell inequalities. The case of two spatially separated domains,” *J. Soviet Math.*, vol. 36, pp. 557–570, 1987.

[32] B. Tsirelson, “Some results and problems on quantum Bell-type inequalities,” *Hadronic J. Suppl.*, vol. 8, no. 4, pp. 329–345, 1993.

[33] J. L. Krivine, “Constantes de Grothendieck et fonctions de type positif sur les sphères,” *Adv. Math.*, vol. 31, pp. 16–30, 1979.

[34] R. F. Werner, “Quantum states with Einstein-Podolsky-Rosen correlations admitting a hidden-variable model,” *Phys. Rev. A*, vol. 40, pp. 4277–4281, 1989.

[35] A. C. Doherty, P. A. Parrilo, and F. M. Spedalieri, “Distinguishing separable and entangled states,” *Phys. Rev. Lett.*, vol. 88, no. 187904, 2002.

[36] A. Peres, “Separability criterion for density matrices,” *Phys. Rev. Lett.*, vol. 77, pp. 1413–1415, 1996.

[37] M. Horodecki, P. Horodecki, and R. Horodecki, “Separability of mixed states: necessary and sufficient conditions,” *Phys. Lett. A*, vol. 223, no. 1, pp. 1–8, 1996.

[38] A. Acín, N. Gisin, L. Masanes, and V. Scarani, “Bell’s inequalities detect efficient entanglement,” *Int. J. Quant. Inf.*, vol. 2, p. 23, 2004, quant-ph/0310166.

[39] N. Gisin, “Bell’s inequality holds for all non-product states,” *Phys. Lett. A*, vol. 154, no. 5, pp. 201–202, 1991.

[40] J. Barrett, “Nonsequential positive-operator-valued measurements on entangled mixed states do not always violate a Bell inequality,” *Phys. Rev. A*, vol. 65, p. 042302, 2002.

[41] N. Alon and A. Naor, "Approximating the cut-norm via Grothendieck's inequality," in *Proc. of the 36th ACM STOC*. New York: ACM Press, 2004, pp. 72–80.

[42] B. M. Terhal, A. C. Doherty, and D. Schwab, "Symmetric extensions of quantum states and local hidden variable theories," *Phys. Rev. Lett.*, vol. 90, p. 157903, 2003.

[43] S. R. Finch, *Mathematical Constants*, ser. Encyclopedia of Mathematics and its Applications. Cambridge, UK: Cambridge University Press, 2003.

[44] J. Larsson, "Modeling the singlet state with local variables," *Phys. Lett. A*, vol. 256, no. 4, pp. 245–252, 1999.

[45] A. Grothendieck, "Résumé de la théorie métrique des produits tensoriels topologiques," *Boletim Da Sociedade de Matemática de São Paulo*, vol. 8, p. 1, 1953.

[46] P. C. Fishburn and J. A. Reeds, "Bell inequalities, Grothendieck's constant, and root two," *SIAM J. Discrete Math.*, vol. 7, no. 1, pp. 48–56, 1994.

[47] L. Gurvits and H. Barnum, "Largest separable balls around the maximally mixed bipartite quantum state," *Phys. Rev. A*, vol. 66, no. 6, p. 062311, 2002.

[48] B. F. Toner, in preparation.

[49] A. Garg, "Detector error and Einstein-Podolsky-Rosen correlations," *Phys. Rev. D*, vol. 28, no. 4, pp. 785–790, 1983.

[50] H. Bateman, *Higher Transcendental Functions*, A. Erdélyi, Ed. New York: McGraw-Hill, 1953.

[51] R. A. Ryan, *Introduction to Tensor Products of Banach Spaces*, ser. Springer Monographs in Mathematics. London: Springer, 2002.

[52] G. Brassard, R. Cleve, and A. Tapp, "Cost of exactly simulating quantum entanglement with classical communication," *Phys. Rev. Lett.*, vol. 83, pp. 1874–1877, 1999.

[53] T. Maudlin, "Bell's inequality, information transmission, and prism models," in *PSA 1992, Volume 1*, D. Hull, M. Forbes, and K. Okruhlik, Eds. East Lansing: Philosophy of Science Association, 1992, pp. 404–417.

[54] N. Cerf, N. Gisin, and S. Massar, "Classical teleportation of a quantum bit," *Phys. Rev. Lett.*, vol. 84, pp. 2521–2525, 2000.

[55] M. Steiner, "Towards quantifying non-local information transfer: finite-bit non-locality," *Phys. Lett. A*, vol. 270, pp. 239–244, 2000.

[56] S. Massar, D. Bacon, N. Cerf, and R. Cleve, "Classical simulation of quantum entanglement without local hidden variables," *Phys. Rev. A*, vol. 63, p. 052305, 2001.

[57] A. Coates, "A quantum measurement scenario which requires exponential classical communication for simulations," 2002, quant-ph/0203112.

[58] J. S. Bell, *Speakable and unspeakable in quantum mechanics.* New York: Cambridge University Press, 1993.

[59] E. Kushilevitz and N. Nisan, *Communication Complexity.* New York: Cambridge University Press, 1997.

[60] A. C. Yao, "Some complexity questions related to distributed computing," in *Proceedings of the 11th ACM Symposium on Theory of Computing.* ACM Press, 1979, pp. 209–213.

[61] D. Bremner, K. Fukuda, and A. Marzetta, "Primal-dual methods for vertex and facet enumeration," *Disc. and Comp. Geom.*, vol. 20, pp. 333–357, 1998.

[62] K. Fukuda and A. Prodon, "Double description method revisited," in *Combinatorics and Computer Science*, 1995, pp. 91–111, available at http://citeseer.ist.psu.edu/fukuda96double.html.

[63] C. H. Bennett, G. Brassard, C. Crépeau, R. Jozsa, A. Peres, and W. K. Wootters, "Teleporting an unknown quantum state via dual classical and Einstein-Podolsky-Rosen channels," *Phys. Rev. Lett.*, vol. 70, pp. 1895–1899, 1993.

[64] C. H. Bennett and S. J. Wiesner, "Communication via one- and two-particle operators on Einstein-Podolsky-Rosen states," *Phys. Rev. Lett.*, vol. 69, pp. 2881–2884, 1992.

[65] D. Bohm, *Quantum Theory.* New York: Prentice-Hall, 1951.

[66] J. A. Csirik, "The cost of exactly simulating a Bell pair using classical communication," *Phys. Rev. A*, vol. 66, p. 014302, 2002.

[67] K. H. Schatten, "Hidden-variable model for the Bohm Einstein-Podolsky-Rosen experiment," *Phys. Rev. A*, vol. 18, pp. 103–104, 1993.

[68] N. Gisin and B. Gisin, "A local hidden variable model of quantum correlations exploiting the detection loophole," *Phys. Lett. A*, vol. 260, pp. 323–327, 1999.

[69] D. Boschi, S. Branca, F. De Martini, L. Hardy, and S. Popescu, "Experimental realization of teleporting an unknown pure quantum state via dual classical and Einstein-Podolsky-Rosen channels," *Phys. Rev. Lett.*, vol. 80, no. 6, pp. 1121–1125, 1998.

[70] S. L. Braunstein, C. A. Fuchs, H. J. Kimble, and P. van Loock, "Quantum versus classical domains for teleportation with continuous variables," *Phys. Rev. A*, vol. 64, no. 2, p. 022321, 2001.

[71] L. Hardy, "Disentangling nonlocality and teleportation," 1999, quant-ph/9906123.

[72] P. M. Morse and H. Feshbach, *Methods of Theoretical Physics, Part I.* New York: McGraw-Hill, 1953.

[73] S. Boyd and L. Vandenberghe, *Convex Optimization.* Cambridge: Cambridge University Press, 2004, available online at http://www.stanford.edu/˜ boyd/cvxbook/.

[74] V. Coffman, J. Kundu, and W. K. Wootters, "Distributed entanglement," *Phys. Rev. A*, vol. 61, p. 052306, 2000.

[75] V. Scarani and N. Gisin, "Quantum communication between n partners and Bell's inequalities," *Phys. Rev. Lett.*, vol. 87, p. 117901, 2001.

[76] L. Masanes, A. Acín, and N. Gisin, "General properties of nonsignaling theories," *Phys. Rev. A*, vol. 73, no. 1, p. 012112, 2006.

[77] R. F. Werner, "An application of Bell's inequalities to a quantum state extension problem," *Lett. Math. Phys.*, vol. 17, pp. 359–363, 1989.

[78] S. Braunstein and C. Caves, "Wringing out better Bell inequalities," *Ann. Phys. (NY)*, vol. 202, no. 1, pp. 22–56, 1990.

[79] J. Barrett, L. Hardy, and A. Kent, "No signaling and quantum key distribution," *Phys. Rev. Lett.*, vol. 95, p. 010503, 2005.

[80] A. Acín, N. Gisin, and Lluis Masanes, "From Bell's theorem to secure quantum key distribution," *Phys. Rev. Lett.*, vol. 97, no. 12, p. 120405, 2006, quant-ph/0510094.

[81] L. Masanes, "Extremal quantum correlations for N parties with two dichotomic observables per site," 2005, quant-ph/0512100.

[82] R. Bhatia, *Matrix Analysis*, ser. Graduate texts in mathematics. New York: Springer-Verlag, 1996, vol. 169, section VII.1.

[83] I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series, and Products*, 5th ed., A. Jeffrey, Ed. San Diego: Academic Press, 1994.