On the Hardness of the Quantum Separability Problem and the Global Power of Locally Invariant Unitary Operations

by

Sevag Gharibian

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Master of Mathematics
in
Computer Science

Waterloo, Ontario, Canada, 2008

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

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

Abstract

Given a bipartite density matrix ρ of a quantum state, the Quantum Separability problem (QUSEP) asks — is ρ entangled, or separable? In this thesis, we first strengthen Gurvits' 2003 NP-hardness result for QUSEP by showing that the Weak Membership problem over the set of separable bipartite quantum states is *strongly* NP-hard, meaning it is NP-hard even when the error margin is as large as inverse polynomial in the dimension, i.e. is "moderately large". Previously, this NP-hardness was known only to hold in the case of inverse exponential error. We observe the immediate implication of NP-hardness of the Weak Membership problem over the set of entanglement-breaking maps, as well as lower bounds on the maximum (Euclidean) distance possible between a bound entangled state and the separable set of quantum states (assuming $P \neq NP$).

We next investigate the entanglement-detecting capabilities of locally invariant unitary operations, as proposed by Fu in 2006. Denoting the subsystems of ρ as A and B, such that $\rho_B = \mathrm{Tr}_A(\rho)$, a locally invariant unitary operation U^B is one with the property $U^B \rho_B {U^B}^\dagger = \rho_B$. We investigate the maximum shift (in Euclidean distance) inducible in ρ by applying $I \otimes U^B$, over all locally invariant choices of U^B. We derive closed formulae for this quantity for three cases of interest: (pseudo)pure quantum states of arbitrary dimension, Werner states of arbitrary dimension, and two-qubit states. Surprisingly, similar to recent anomalies detected for non-locality measures, the first of these formulae demonstrates the existence of non-maximally entangled states attaining shifts as large as maximally entangled ones. Using the latter of these formulae, we demonstrate for certain classes of two-qubit states an equivalence between the Fu criterion and the CHSH inequality. Among other results, we investigate the ability of locally invariant unitary operations to detect bound entanglement.

Acknowledgements

I thank Richard Cleve for helpful discussions, guidance, and playing devil's advocate for the material in this thesis.

With regard to the work on hardness of the separability problem, I thank Ben Toner for ideas which spurred the line of thought followed here, Stephen Vavasis for helpful discussions regarding convex optimization, and John Watrous for helpful discussions regarding quantum channels and the Jamiołkowski isomorphism.

For the work on locally invariant unitary operations, I thank my co-conspirators Dagmar Bruß and Hermann Kampermann, who were gracious enough to allow me to base part of my thesis on our joint work. I am also indebted to Marco Piani for suggesting the approach which led to the connection between the Fu criterion and the CHSH inequality, and André Méthot for intriguing discussions regarding non-locality.

I would like to thank those who happily agreed to be readers for my thesis, namely Richard Cleve, John Watrous, and Ashwin Nayak.

On a personal note, I would like to thank my family for always being supportive in every possible scenario, and my colleagues and friends for the many entertaining and intellectually stimulating discussions that make life memorable.

Last but not least, I would like to thank my über-patient and supportive Mareike Müller, for nurturing with loving care my sometimes workaholic nature.

Dedication

For the love of science and the unravelling beauty of the quantum world around us.

Contents

List of Figures

Chapter 1

Introduction

Once upon a time, there lived many a prominent physicist who found it difficult to accept a world in which entanglement existed. In 1935, Einstein, Podolsky and Rosen stated this mindset in their (in)famous “EPR” paper [22], in which they rejected quantum mechanics as a complete physical theory, due to what Einstein liked to call “spukhafte Fernwirkung” (“spooky action at a distance”). Over 70 years and many Bell inequalities [5] later, however, it is hard to imagine a quantum world *without* entanglement. With continuing applications for entanglement emerging in the fields of quantum computation and information, researchers are scrambling just to detect and quantify what is now generally believed to be a valuable resource. Yet, it turns out that the phenomenon once dubbed “the characteristic trait of quantum mechanics” by Schrödinger himself [57] is in fact unlikely to be detectable efficiently by *any* algorithm (in the general case), unless $\mathbf{P} = \mathbf{NP}$ [31].

Indeed, since the early 1990’s, it has been observed by the quantum information and computing community that entanglement lends itself to a number of interesting applications, such as quantum teleportation [7], superdense coding [8], quantum parallelism [58], quantum communication complexity [17, 30], and quantum cryptography [23]. In response, there has been an endless slew of proposed entanglement detection criteria and measures, such as the Positive Partial Transpose (PPT) criterion [55], Von Neumann entropy of reduced states [20], entanglement of formation [10], majorization criterion [54], relative entropy of entanglement [60], squashed entanglement [15], entanglement cost [10], distillable entanglement [10], negativity [61], and the first need for positive but not completely positive maps in physics [35], to name but a few. Yet, the problem of determining in general whether a given quantum state is entangled or separable, informally dubbed the Quantum Separability problem (QUSEP), was proven NP-hard by Gurvits in 2003 [31]. This was accomplished by phrasing it as an instance of the Weak Membership problem, where the Weak Membership problem over arbitrary convex set K is informally defined as the problem of deciding whether a given point p is in K, given a certain allowed margin of error. The hardness result was then shown via a reduction from the NP-complete problem KNAPSACK to Weak Membership over the (convex) set of separable quantum states.

In this thesis, we hence address the following. We first discuss the hardness result of Gurvits [31], specifically in light of an observation made by Aaronson [43] that Gurvits' proof has a shortcoming — it demonstrates NP-hardness of Weak Membership over the set of separable quantum states only if the error margin allowed is at most exponentially small in the dimension of the corresponding Hilbert space [43], i.e. is "very small". We thus first strengthen Gurvits' result by demonstrating a simple reduction which combines the previous work of Gurvits [31], Ioannou [43], and Liu [48], and show that the Weak Membership problem over the set of separable quantum states remains NP-hard even if the error allowed is as large as inverse polynomial in the dimension, i.e. is "moderately large". This translates to a property known as *strong* NP-hardness. We remark that this implies immediate lower bounds on the maximum distance possible between a bound entangled state and the separable set, as well as strong NP-hardness of the Weak Membership problem over the set of entanglement-breaking maps.

In defense of the many proposed attempts at detecting entanglement, on the other hand, we reference a spoken quote the author of this thesis finds amusing. As stated by Christos Papadimitriou in a visit to the University of Waterloo on April 25, 2008: "...proving NP-hardness is only the opening move in an interesting game between two players". In this spirit, the second half of this thesis investigates the abilities of an entanglement detection criterion proposed by Fu in 2006 [25], based on the concept of locally invariant unitary operations. What makes this criterion interesting, as we demonstrate, is that it exhibits strange anomalies similar to those recently observed for non-locality measures [21, 1, 12, 50]. Specifically, we find that under this criterion, certain non-maximally entangled quantum states are able to "perform" equally well as maximally entangled ones. Yet, this criterion is provably *not* a non-locality measure [25]. Among other results, we also reveal connections between this criterion and the CHSH inequality [16] for the case of two-qubit quantum states.

This thesis is organized as follows. In Chapter 2, we show strong NP-hardness of the Weak Membership problem over the set of separable quantum states. Chapter 3 investigates Fu's entanglement detection criterion based on locally invariant unitary operations, which we remark is based on joint work with Dagmar Bruß and Hermann Kampermann of the Heinrich-Heine-Universität Düsseldorf. We close in Chapter 4 with concluding comments and open problems. We have attempted to make each chapter as self-contained as possible.

For an introduction to quantum computing and information, we recommend [53]. For detailed background on the quantum separability problem, we refer the reader to [43]. Good surveys on entanglement and entanglement measures and detection criteria are [42, 13, 6]. Non-locality measures and the respective anomalies mentioned above are discussed in [50].

Chapter 2

Strong NP-Hardness of Weak Membership Over the Set of Separable Quantum States

2.1 A Brief History

Given a bipartite quantum state ρ in Hilbert space $\mathcal{H}^M \otimes \mathcal{H}^N$, specified in terms of its density matrix with respect to some known orthonormal basis, it has been a longstanding open problem to crack the Quantum Separability problem (QUSEP) — is ρ separable, or equivalently, unentangled? A natural way to think about this problem is in terms of membership in a convex set, since the set of separable quantum states is convex. Let us denote the set of separable bipartite quantum states in $\mathcal{H}^M \otimes \mathcal{H}^N$ as $\mathcal{S}_{M,N}$. Then, the resulting membership problem is referred to as Strong Membership over $\mathcal{S}_{M,N}$, denoted SMEM($\mathcal{S}_{M,N}$), where the title *strong* emphasizes that no margin of error is allowed in determining membership in $\mathcal{S}_{M,N}$.

In order to understand QUSEP, one can equivalently study SMEM($\mathcal{S}_{M,N}$). Is this, however, a computationally sound formulation to consider? Consider as input a quantum state $\rho \in \mathcal{H}^M \otimes \mathcal{H}^N$, such that ρ sits directly on the border of $\mathcal{S}_{M,N}$ (therefore $\rho \in \mathcal{S}_{M,N}$). In specifying the entries of the density matrix of ρ, we must use some finite precision, meaning that certain entries, such as irrational ones, may have to be approximated. In doing so, it is entirely possible that ρ is accidentally shifted outside $\mathcal{S}_{M,N}$, thus altering the answer to QUSEP, and making the problem undecidable for such a case. In this vein, it seems computationally natural to allow some margin of error in deciding membership in $\mathcal{S}_{M,N}$. We hence consider the problem of *Weak* Membership over $\mathcal{S}_{M,N}$ (WMEM$_\beta$($\mathcal{S}_{M,N}$)), where the name *weak* now derives from the fact that we introduce a parameter β quantifying an allowed margin of error in determining membership in $\mathcal{S}_{M,N}$.

In 2003, it was shown by Gurvits [31] that WMEM$_\beta$($\mathcal{S}_{M,N}$) is NP-hard, via a reduction from the NP-complete problem KNAPSACK, of which the NP-complete problem PARTITION can easily be seen to be a special case. It was soon pointed out by Scott Aaronson [43], however, that PARTITION is only *weakly* NP-hard, since it can be solved efficiently if the values of its numerical parameters are polynomially bounded in the dimension (in this case, via a dynamic programming approach [27]). A similar dynamic programming approach also exists for KNAPSACK [19], making KNAPSACK weakly NP-hard as well. It therefore follows that hardness for $\mathrm{WMEM}_\beta(\mathcal{S}_{M,N})$ following the reduction of [31] holds only when $\beta \leq 1/\exp(M,N)$ [43], leaving open the possibility of an approximation algorithm when weaker precision is demanded.

A second attempt to strengthen this result was then made by Gurvits [43], via the following reduction from the NP-complete problem CLIQUE [27], where RSDF and $\mathrm{WVAL}_\alpha(\mathcal{S}_{M,N})$ are the problems Robust Semidefinite Feasibility and Weak Validity over $\mathcal{S}_{M,N}$, respectively (all necessary formal definitions are given in Section 2.2, and by $\leq_K$ and $\leq_C$, we indicate a Karp and Cook reduction[1], respectively):

$$\mathrm{CLIQUE} \leq_K \mathrm{RSDF} \leq_K \mathrm{WVAL}_\alpha(\mathcal{S}_{M,N}) \leq_C \mathrm{WMEM}_\beta(\mathcal{S}_{M,N}). \tag{2.1}$$

This approach fares better, proving *strong* NP-hardness of $\mathrm{WVAL}_\alpha(\mathcal{S}_{M,N})$, where we define a problem as strongly NP-hard if it is NP-hard even if the values of its numerical parameters are polynomially bounded in the length of its input [27]. In this case, we have $\mathrm{WVAL}_\alpha(\mathcal{S}_{M,N})$ is NP-hard for $\alpha \leq 1/\operatorname{poly}(M,N)$, for α its error parameter [43]. The only known deterministic polynomial time Turing reduction from $\mathrm{WVAL}_\alpha(K)$ to $\mathrm{WMEM}_\beta(K)$ for arbitrary choice of $\alpha > 0$ (for $K \subseteq \mathbb{R}^n$ any convex set circumscribed in an origin-centered ball of radius R, and containing an inscribed ball centered at a known point $p \in K$ of radius r), however, is via the Yudin-Nemirovski theorem [67], which is based on the shallow-cut ellipsoid method (there is, however, a *randomized* reduction following from the random walk algorithm for convex optimization of Bertsimas and Vempala [11]). It turns out that following this reduction also results in exponential scaling for β [43], leaving us with the same problem as before — $\mathrm{WMEM}_\beta(\mathcal{S}_{M,N})$ is known to be NP-hard only for error parameter $\beta \leq 1/\exp(M,N)$.

The main result we show in this chapter is hence as follows.

Theorem 1. $\mathrm{WMEM}_\beta(\mathcal{S}_{M,N})$ *is strongly NP-hard, or equivalently, is NP-hard for* $\beta \leq 1/\operatorname{poly}(M,N)$.

In order to show this result, we observe that there exists a recent non-ellipsoidal Turing reduction by Liu [48] from the problem Weak Optimization[2] ($\mathrm{WOPT}_\epsilon(K)$)

[1] A Cook reduction is a polynomial time Turing reduction from problem A to B, where, intuitively, a Turing reduction demonstrates how to solve A using an algorithm for B possibly multiple times. A Karp reduction, on the other hand, invokes the algorithm for B once, the output of which it returns as the solution for A.

[2] What we call Weak Optimization can actually be shown to be equivalent to the problem Weak Validity in [29]. We follow the naming convention in [48], however, in order to avoid confusion when applying the reduction therein.

to $\mathrm{WMEM}_\beta(K)$, which runs in polynomial time if we demand at most "moderate" precision in solving Weak Optimization, i.e. $\epsilon \geq 1/\operatorname{poly}(M,N)$ for ϵ the error parameter for Weak Optimization. For $\epsilon = 1/\exp(M,N)$, we remark that the reduction is no longer poly-time, a point we discuss in further depth later. We thus begin by following the reduction CLIQUE $\leq_K$ RSDF from Equation 2.1. We then show a Karp reduction from RSDF to $\mathrm{WOPT}_\epsilon(\mathcal{S}_{M,N})$. We finish with the Turing reduction of Liu from $\mathrm{WOPT}_\epsilon(K)$ to $\mathrm{WMEM}_\beta(K)$. The modified reduction chain (in simplified form, for now) is then (where $\leq_T$ denotes a Turing reduction):

$$\text{CLIQUE} \leq_K \text{RSDF} \leq_K \mathrm{WOPT}_\epsilon(\mathcal{S}_{M,N}) \leq_T \mathrm{WMEM}_\beta(\mathcal{S}_{M,N}). \tag{2.2}$$

Since our goal is to prove NP-hardness of $\mathrm{WMEM}_\beta(\mathcal{S}_{M,N})$ for $\beta = 1/\operatorname{poly}(M,N)$, note that the precision restriction on the last link (following the Liu reduction) that $\epsilon \geq 1/\operatorname{poly}(M,N)$ will not be an issue.

We also discuss two possible applications of Theorem 1. First, the hardness results shown here imply immediate lower bounds on the maximum (Euclidean) distance possible between a bound entangled [36] state and the separable set, $\mathcal{S}_{M,N}$. Second, we observe that NP-hardness of Weak Membership over the set of separable quantum states immediately implies NP-hardness of determining whether a completely positive linear map is an entanglement-breaking map [37].

This chapter is organized as follows. In Section 2.2, we introduce and discuss necessary definitions for the problems in Equation 2.2. In Section 2.3, we show the desired reduction from CLIQUE to $\mathrm{WMEM}_\beta(\mathcal{S}_{M,N})$. Section 2.4 discusses applications of Theorem 1 to bound entanglement and entanglement-breaking maps. We close in Section 2.5 with some brief comments and open problems.

2.2 Pertinent Definitions and Discussion

There are six main problems we need to define in order to show our desired reduction chain: CLIQUE, Robust Semidefinite Feasibility (RSDF), Weak Optimization, Weak Separation, Weak Membership, and Quantum Separability (QUSEP). We begin by presenting their definitions here, along with pertinent discussion. All norms are taken as the Euclidean norm (indicated $\|\ \|_2$ or $\|\ \|_{\mathrm{F}}$, the latter denoting the Frobenius norm for matrices). We denote a (column) vector v by $\mathbf{v}$, its conjugate transpose as $\mathbf{v}^\dagger$, and its ith entry as v_i. We use the notation $\langle \alpha \rangle$ to signify the number of bits necessary to encode an entity α. Specifically, if $\alpha = a/b$ is rational, we define $\langle \alpha \rangle = \langle a \rangle + \langle b \rangle$. For a matrix A, we let $\langle A \rangle = \sum_{i,j} \langle A_{ij} \rangle$ (similarly for vectors).

First, the NP-complete problem CLIQUE is stated as follows.

Definition 1 (CLIQUE)**.** Given a simple graph G on n vertices, and $c \leq n$, for $n, c \in \mathbb{Z}^+$, decide, with respect to the complexity measure $\langle G \rangle + \langle c \rangle$:

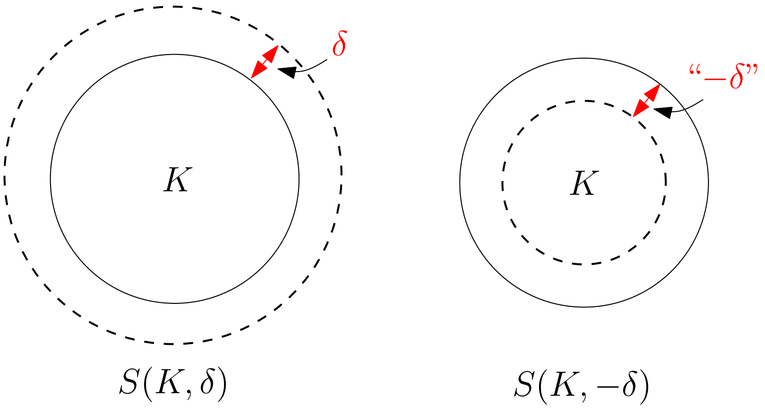

Figure 2.1: An intuitive picture of the sets $S(K, \delta)$ and $S(K, -\delta)$, respectively. We use the term "$-\delta$" in the latter illustration to stress the fact that K is contracted.

> If the number of vertices in the largest set of pairwise adjacent vertices in G is at least c, output "YES".
> Otherwise, output "NO".

Here, we take $\langle G\rangle = \langle A_G\rangle$, where A_G is the $n \times n$ adjacency matrix for G, such that $A_G[i, j] = 1$ if vertices i and j are connected by an edge, and $A_G[i, j] = 0$ otherwise. Next, let us define the problem Robust Semidefinite Feasibility (RSDF).

Definition 2 (RSDF)**.** Given k rational, symmetric $l \times l$ matrices $B_1, \ldots, B_k$, and $\zeta, \eta \in \mathbb{Q}$, with $\zeta, \eta \geq 0$, define $g(B_1, \ldots, B_k) = \max_{\mathbf{x}\in\mathbb{R}^l, \|\mathbf{x}\|_2=1} \sum_{i=1}^{k} (\mathbf{x}^{\mathrm{T}} B_i \mathbf{x})^2$. Then, decide, with respect to the complexity measure $lk + \sum_{1\leq i\leq k} \langle B_i\rangle + \langle\zeta\rangle + \langle\eta\rangle$:

> If $g(B_1, \ldots, B_k) \geq \zeta + \eta$, output "YES".
> If $g(B_1, \ldots, B_k) \leq \zeta - \eta$, output "NO".

Note that we have assumed $\zeta \geq 0$ without loss of generality above, since $g(B_1, \ldots, B_k) \geq 0$. This will be necessary later in Lemma 5, when we need to take $\sqrt{g(B_1, \ldots, B_k)}$. Note also that we have defined RSDF as a *promise* problem, meaning we are promised the input falls into one of two disjoint cases which may separated by a non-zero gap, and we are asked to distinguish between the two cases. One could equivalently lift the promise and allow input falling in the "gap" or "error" region (such that $\zeta - \eta < g(B_1, \ldots, B_k) < \zeta + \eta$) — in this case, for any such input, we would consider any output to be correct (i.e. "YES" or "NO"), a point we leverage later.

Moving on, for Weak Optimization, Weak Separation, and Weak Membership, we first define a few notions related to convex optimization. For any Euclidean space $\mathcal{E}$, where we call a space Euclidean if it is a finite dimensional vector space over $\mathbb{C}$ or $\mathbb{R}$ with a fixed inner product, a set $\mathcal{C} \subseteq \mathcal{E}$ is called *convex* if for all $x, y \in \mathcal{C}$ and $p \in [0, 1]$, we have $px + (1-p)y \in \mathcal{C}$. We define an *open ball* of radius r about a vector $\mathbf{v} \in \mathcal{E}$ as $B_r(\mathbf{v}) := \{\mathbf{u} \in \mathcal{E} \mid \|\mathbf{v} - \mathbf{u}\|_2 < r\}$. The set $\mathcal{C}$ is then called *open* if for all $\mathbf{v} \in \mathcal{E}$, there exists some $r > 0$ such that $\mathcal{B}_r(\mathbf{v}) \subseteq \mathcal{C}$. The complement of such a set is called *closed*. Further, the set $\mathcal{C}$ is *bounded* if $\mathcal{C} \subseteq \mathcal{B}_R(\mathbf{0})$, for $\mathbf{0}$ the origin and some real $R > 0$. By the Heine-Borel Theorem [64], $\mathcal{C}$ is called *compact* if and only if it is closed and bounded. Finally, following the terminology of [29], we call $\mathcal{C}$ *well-bounded* if it is bounded and *centered*, the latter meaning $\mathcal{C}$ contains a ball of some radius $r > 0$, the center of which need not be known. If the center of this contained ball is some *known* $\mathbf{p} \in \mathcal{C}$, however, then $\mathcal{C}$ is instead called p-centered.

Let us now put these definitions to use. Let $K \subseteq \mathbb{R}^m$ be a convex and compact set, and define the following related sets:

$$\begin{aligned} S(K,\delta) &:= \{\mathbf{x} \in \mathbb{R}^m \mid \exists\, \mathbf{y} \in K \text{ s.t. } \|\mathbf{x}-\mathbf{y}\|_2 \leq \delta\} && (2.3)\\ S(K,-\delta) &:= \{\mathbf{x} \in K \mid S(\mathbf{x},\delta) \subseteq K\}. && (2.4)\end{aligned}$$

Roughly, $S(K,\delta)$ can be thought of as extending the border of K by δ (in Euclidean distance), and $S(K,-\delta)$ can be thought of as taking the core of K, which is δ away from the border of K. We pictorially demonstrate this intuition in Figure 2.1. For the problems we define here, we require that K be well-bounded and p-centered, which ensures that K is bounded and full-dimensional. We will see shortly that $\mathcal{S}_{M,N}$ is in fact such a set. Finally, we set the encoding size of K as $\langle K\rangle = m + \langle r\rangle + \langle R\rangle + \langle \mathbf{p}\rangle$. Weak Optimization, Weak Separation, and Weak Membership over the set K are then defined as follows.

Definition 3 (Weak Optimization (WOPT$_\epsilon(K)$)[48])**.** Given $\mathbf{c} \in \mathbb{Q}^m$, such that $\|\mathbf{c}\|_2 = 1$, and $\gamma, \epsilon \in \mathbb{Q}$, such that error parameter $\epsilon > 0$, decide, with respect to the complexity measure $\langle K\rangle + \langle \mathbf{c}\rangle + \langle \gamma\rangle + \langle \epsilon\rangle$:

If there exists $\mathbf{y} \in S(K,-\epsilon)$ with $\mathbf{c}^{\mathrm{T}}\mathbf{y} \geq \gamma + \epsilon$, then output "YES".
If for all $\mathbf{x} \in S(K,\epsilon)$, $\mathbf{c}^{\mathrm{T}}\mathbf{x} \leq \gamma - \epsilon$, then output "NO".

Definition 4 (Weak Separation (WSEP$_\nu(K)$))**.** Given $\mathbf{z} \in \mathbb{Q}^m$, and error parameter $\nu \in \mathbb{Q}$, such that $\nu > 0$, respond, with respect to the complexity measure $\langle K\rangle + \langle \mathbf{z}\rangle + \langle \nu\rangle$:

If $\mathbf{z} \in S(K,-\nu)$, then output "YES".
If $\mathbf{z} \notin S(K,\nu)$, then return $\mathbf{c} \in \mathbb{Q}^m$, such that $\forall \mathbf{x} \in S(K,-\nu)$, $\mathbf{c}^{\mathrm{T}}\mathbf{x} \leq \mathbf{c}^{\mathrm{T}}\mathbf{z} + \nu$, and $\|\mathbf{c}\|_2 = 1$.

Definition 5 (Weak Membership (WMEM$_\beta(K)$))**.** Given $\mathbf{y} \in \mathbb{Q}^m$, and error parameter $\beta \in \mathbb{Q}$, such that $\beta > 0$, decide, with respect to the complexity measure $\langle K\rangle + \langle \mathbf{y}\rangle + \langle \beta\rangle$:

If $\mathbf{y} \in S(K,-\beta)$, then output "YES".
If $\mathbf{y} \notin S(K,\beta)$, then output "NO".

Note that all three of these problems are stated as promise problems, and can be phrased as follows: $\mathrm{WOPT}_\epsilon(K)$ asks whether there exists a point in the "core" of K that achieves a threshold value slightly higher than γ for a linear function defined by $\mathbf{c}$. $\mathrm{WSEP}_\nu(K)$ asks us to return an "approximately" separating hyperplane from the "core" of K if the given point $\mathbf{z}$ is not "near" K. $\mathrm{WMEM}_\beta(K)$ asks to distinguish whether a given point $\mathbf{y}$ is in the "core" of K or "far away" from K. Note that as per the convention of [48], we define $\mathrm{WSEP}_\nu(K)$ such that $\|\mathbf{c}\|_2 = 1$ for the "NO" case, as opposed to $\|\mathbf{c}\|_\infty = 1$, as in [29]. The latter text uses $\|\mathbf{c}\|_\infty = 1$, since if the error parameter ν is "very small", one must be careful to avoid scaling by an irrational factor when normalizing $\mathbf{c}$ in order to avoid sensitive round-off errors. As mentioned earlier, however, to utilize Liu's reduction for $\mathrm{WOPT}_\epsilon(K) \leq_T \mathrm{WMEM}_\beta(K)$ [48], we must restrict error parameter ϵ to be at least inverse polynomial in the length of the input (from which similar bounds on ν and β follow), and so we can tolerate slight round-off errors [48], allowing us to set $\|\mathbf{c}\|_2 = 1$.

Let us also note that unlike here, in [48] the inputs to $\mathrm{WOPT}_\epsilon(K)$, $\mathrm{WSEP}_\nu(K)$, and $\mathrm{WMEM}_\beta(K)$ are real (as opposed to rational), and specified using $\mathrm{poly}(m)$ bits of precision[3], where $K \subseteq \mathbb{R}^m$. This is because $\mathrm{poly}(m)$ bits of precision suffice if we demand ϵ, ν, and β to be at least inverse polynomial in the input size [48], i.e. if the error parameters themselves are not exponentially small (which we shall also demand here). It is easy to see that we can exactly represent any such $\mathrm{poly}(m)$-bit real numbers as rational numbers in poly-time using $\mathrm{poly}(m)$ bits as well, and hence the case of [48] can be seen as a special case of our definitions here.

Finally, let us formally define the Quantum Separability problem, and discuss how it relates to the problems over convex sets we have just defined.

Definition 6 (Quantum Separability Problem (QUSEP))**.** Let $H_{M,N}$ denote the set of Hermitian operators mapping $\mathbb{C}^M \otimes \mathbb{C}^N \mapsto \mathbb{C}^M \otimes \mathbb{C}^N$, $M, N \geq 2$. Denote the set of separable states as $\mathcal{S}_{M,N} = \mathrm{conv}\{\mathbf{x}\mathbf{x}^\dagger \otimes \mathbf{y}\mathbf{y}^\dagger \mid \mathbf{x} \in \mathbb{C}^M, \mathbf{y} \in \mathbb{C}^N, \|\mathbf{x}\|_2 = \|\mathbf{y}\|_2 = 1\}$, where $\mathrm{conv}\{\mathrm{S}\}$ denotes the convex hull generated by the set S. Then, given quantum state $\rho_{A,B} \in H_{M,N}$, decide:

If $\rho_{AB} \in \mathcal{S}_{M,N}$, output "YES".
Otherwise, output "NO".

As claimed earlier, it is clear from Definition 6 that $\mathcal{S}_{M,N}$ is indeed a convex set. In fact, $\mathcal{S}_{M,N}$ is p-centered and well-bounded, since $\mathcal{S}_{M,N} \subseteq S(\mathbf{0}, R)$ for $\mathbf{0}$ the origin and $R = \sqrt{(MN-1)/MN}$ [6], and $S(p,r) \subseteq \mathcal{S}_{M,N}$ for $p = I/MN$ [6] (where I denotes the identity) and $r = \sqrt{1/MN(MN-1)}$ [32]. Further, $\mathcal{S}_{M,N}$ is compact, since the set of pure product states is closed and bounded, and the convex hull of a convex compact set is also compact [62]. Thus, we can rephrase QUSEP as $\mathrm{WMEM}_\beta(\mathcal{S}_{M,N})$, and investigate hardness of $\mathrm{WMEM}_\beta(\mathcal{S}_{M,N})$ in order to draw conclusions about QUSEP. Thus, we never work with QUSEP directly.

[3] It will be clear later that in our case, we can take m to be polynomial in the encoding size of the original CLIQUE instance in our reduction chain.

Observe now that $\mathrm{WOPT}_\epsilon(K)$, $\mathrm{WSEP}_\nu(K)$, and $\mathrm{WMEM}_\beta(K)$ are phrased over $K \subseteq \mathbb{R}^m$, whereas $\mathcal{S}_{M,N} \subseteq H_{M,N}$. To bridge this gap, recall that $H_{M,N}$ is isomorphic to $\mathbb{R}^{M^2N^2}$, and specifically one can write for any quantum state $\rho \in H_{M,N}$ [45]:

$$\rho = \frac{I}{MN} + \frac{1}{2}\sum_{i=1}^{M^2N^2-1} r_i\lambda_i, \tag{2.5}$$

where we have chosen as a basis for $H_{M,N}$ the identity and the traceless orthogonal Hermitian generators of $SU(MN)$, the latter denoted by λ_i, and the terms $r_i \in \mathbb{R}$ denote the components of the Bloch vector corresponding to ρ, such that component $r_i = \mathrm{Tr}(\rho\lambda_i)$. Thus, by setting $m = M^2N^2 - 1$, we can can work over the space of Bloch vectors in $\mathbb{R}^m$ rather than density matrices in $H_{M,N}$. We remark that there is a scaling factor of $\sqrt{2}$ in Euclidean distance between two states when switching between the two spaces (see Appendix A), which does not affect our analysis.

We can finally formally state what is required to show Theorem 1. From the values of r, R, and p for $\mathcal{S}_{M,N}$, it follows that $\langle \mathcal{S}_{M,N}\rangle = m + \langle R\rangle + \langle r\rangle + \langle \mathbf{p}\rangle \leq \mathrm{poly}(MN)$. Thus, by the complexity measure for Definition 5, our aim is to show NP-hardness of $\mathrm{WMEM}_\beta(\mathcal{S}_{M,N})$ with respect to $MN + \langle \mathbf{y}\rangle + \lceil \frac{1}{\beta}\rceil$, where the $\lceil \frac{1}{\beta}\rceil$ term follows from the fact that we wish to prove *strong* NP-hardness.

2.3 The Reduction

Let us show our main result, Theorem 1. To do so, we demonstrate a Turing reduction from CLIQUE to $\mathrm{WMEM}_\beta(\mathcal{S}_{M,N})$, obtained by a simple Karp reduction combining previous ideas of Gurvits [31], Ioannou [43], and Liu [48]. Specifically, we show the following reduction chain (which is a slightly more detailed version of Equation 2.2):

$$\text{CLIQUE} \leq_K \text{RSDF} \leq_K \mathrm{WOPT}_\epsilon(\mathcal{S}_{M,N}) \leq_T \mathrm{WSEP}_\nu(\mathcal{S}_{M,N}) \leq_T \mathrm{WMEM}_\beta(\mathcal{S}_{M,N}) \tag{2.6}$$

The reduction behind the third link in this chain runs in time polynomial in $1/\epsilon$, implying we must be able to choose $\epsilon \geq 1/\mathrm{poly}(M, N)$ in order for the entire chain to run in polynomial time. We return to this point later.

To begin, the first link in Equation 2.6 is given as follows. Unless otherwise stated, by a poly-time reduction, we mean with respect to the encoding size of the problem instance, as defined in Section 2.2. We use the notation $\Pi = (\textit{input parameters})$ to denote an instance Π of a given problem, with Π specified by the given input parameters.

Lemma 2 (Ioannou [43]). *There exists a poly-time Karp reduction which maps instance $\Pi_1 = (G, n, c)$ of* CLIQUE *to instance $\Pi_2 = (k, l, B_1, \ldots, B_k, \zeta, \eta)$ of* RSDF*, such that $k = n(n-1)/2$, $l = n$, $B_i \in \mathbb{Q}^{n\times n}$ and $\|B_i\|_{\mathrm{F}} \in \Theta(1)$ for all $1 \leq i \leq k$, $\zeta = \Theta(1)$, $\eta \in \Omega(n^{-2})$.*

Proof. The proof hinges on the following theorem relating the maximum clique size to optimization over a square-free quadratic form.

Theorem 3 (Motzkin and Straus [52]). *Denote by $(i,j) \in G$ an edge in graph G between vertices i and j, and let k be the order of the maximal complete graph contained in G. Let Δ_n denote the simplex $\Delta_n := \{\mathbf{x} \in \mathbb{R}^n \mid x_i \geq 0,\ \|\mathbf{x}\|_1 = 1\}$. Then*

$$\max_{\mathbf{x}\in\Delta_n} \sum_{(i,j)\in G} x_i x_j = \frac{1}{2}\left(1 - \frac{1}{k}\right). \tag{2.7}$$

Note that the term on the left side of Equation 2.7 is indeed a square-free form, since by the definition of CLIQUE, G is a simple graph. Proving one direction of Theorem 3 is straightforward — let $1, \ldots, k$ denote the vertices in the maximal clique of size k, and consider vector $\mathbf{x} \in \mathbb{R}^n$, with $x_i = 1/k$ for $1 \leq i \leq k$ and $x_i = 0$ otherwise. Then $\sum_{(i,j)\in G} x_i x_j \geq \binom{k}{2}\frac{1}{k^2} \geq \frac{1}{2}\left(1 - \frac{1}{k}\right)$. Proving the inequality in the other direction is slightly trickier. The reader is referred to [52] for further details.

To proceed with our reduction then, note that we can put the expression on the left side of Equation 2.7 into the form required by RSDF by writing, where A denotes the adjacency matrix of G:

$$\begin{aligned}\max_{\mathbf{x}\in\Delta_n} \sum_{(i,j)\in G} x_i x_j = \max_{\mathbf{x}\in\Delta_n} \mathbf{x}^{\mathrm{T}} A \mathbf{x} &= \max_{\mathbf{y}\in\mathbb{R}^n, \|\mathbf{y}\|_2 = 1} \sum_{i,j=1}^{n} y_i^2 A_{ij} y_j^2 \\ &= \max_{\mathbf{y}\in\mathbb{R}^n, \|\mathbf{y}\|_2 = 1} \frac{1}{2} \sum_{1\leq i<j\leq n} (\mathbf{y}^{\mathrm{T}} D^{ij} \mathbf{y})^2, \end{aligned} \tag{2.8}$$

where the second equality follows from a simple substitution of variables, $x_i = y_i^2$, which we can do over $\mathbb{R}$ since each $x_i \geq 0$. Note that $\|\mathbf{x}\|_1 = 1$ if and only if $\|\mathbf{y}\|_2 = 1$. The third equality is based on straightforward manipulation by defining each $D^{ij} \in \mathbb{Q}^{n\times n}$ to have all zero entries except for $D^{ij}_{i,j} = D^{ij}_{j,i} = A_{i,j}$, giving us a total of $n(n-1)/2$ such matrices D^{ij}.

To complete the reduction, set the parameters of instance Π_2 of RSDF as follows. Let each D^{ij} be a matrix B_i, $k = n(n-1)/2$, $l = n$, $\zeta = 1 - \frac{1}{2c-2} - \frac{1}{2c}$ (i.e. the midpoint of interval $\left[1 - \frac{1}{c-1}, 1 - \frac{1}{c}\right]$), and choose error parameter η such that $1 - \frac{1}{c-1} < \zeta - \eta < \zeta + \eta < 1 - \frac{1}{c}$. Then if the maximum clique size is at least c, Equations 2.7 and 2.8 yield $\max_{\mathbf{y}\in\mathbb{R}^n, \|\mathbf{y}\|_2 = 1} \sum_{1\leq i<j\leq n} (\mathbf{y}^{\mathrm{T}} D^{ij} \mathbf{y})^2 \geq 1 - \frac{1}{c}$, and we have a "YES" instance of RSDF. The "NO" case is analogous. Observing that the length of the interval $\left[1 - \frac{1}{c-1}, 1 - \frac{1}{c}\right] \in \Omega(\frac{1}{c^2}) \in \Omega(\frac{1}{n^2})$ (since $c \leq n$) allows us to conclude $\eta \in \Omega(\frac{1}{n^2})$, completing the proof. □

Lemma 2 gives us the first link in Equation 2.6. In order to prove the next link, RSDF $\leq_K$ WOPT$_\epsilon(\mathcal{S}_{M,N})$, we first require the following lemma relating RSDF to convex optimization over $\mathcal{S}_{M,N}$.

Lemma 4 (Gurvits [31]). *Let $\{\mathbf{b}\}_{i=1}^{M}$ be an orthonormal basis for $\mathbb{C}^M$. Then, given $(M-1)$ real, symmetric matrices $A_1, \ldots, A_{M-1}$ of dimension $N \times N$, define Hermitian block matrix $C \in \mathbb{R}^{MN \times MN}$:*

$$C := \sum_{i=1}^{M-1} \left(\mathbf{b}_1 \mathbf{b}_{i+1}^{\dagger} + \mathbf{b}_{i+1} \mathbf{b}_1^{\dagger} \right) \otimes A_i = \begin{pmatrix} 0 & A_1 & \ldots & A_{M-1} \\ A_1 & 0 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ A_{M-1} & 0 & \ldots & 0 \end{pmatrix}, \tag{2.9}$$

where each block of zeroes is also $N \times N$. Then, for $g(\cdot)$ from Definition 2:

$$\sqrt{g(A_1, \ldots, A_{M-1})} = \max_{\mathbf{x} \in \mathbb{R}^N, \|\mathbf{x}\|_2 = 1} \sqrt{\sum_{i=1}^{M-1} (\mathbf{x}^{\mathrm{T}} A_i \mathbf{x})^2} = \max_{\rho \in \mathcal{S}_{M,N}} \mathrm{Tr}(C\rho). \tag{2.10}$$

Proof. We follow [31] by beginning with the expression on the right side of Equation 2.10 and work backwards, in order to provide a more intuitive approach. Let $f(C) = \max_{\rho \in \mathcal{S}_{M,N}} \mathrm{Tr}(C\rho)$. Since $f(C)$ is simply linear optimization over a convex set, the optimum for $f(C)$ occurs, without loss of generality, at an extreme point of $\mathcal{S}_{M,N}$. If we denote the set of extreme points of $\mathcal{S}_{M,N}$ as $\mathcal{E}_{M,N}$, then by Definition 6, we can write

$$\mathcal{E}_{M,N} = \left\{ \mathbf{y}\mathbf{y}^{\dagger} \otimes \mathbf{x}\mathbf{x}^{\dagger} \mid \mathbf{y} \in \mathbb{C}^M, \mathbf{x} \in \mathbb{C}^N, \|\mathbf{y}\|_2 = \|\mathbf{x}\|_2 = 1 \right\}, \tag{2.11}$$

and by the observation above, $f(C) = \max_{\rho \in \mathcal{E}_{M,N}} \mathrm{Tr}(C\rho)$.

Now consider matrix $H(\mathbf{x}) \in \mathbb{R}^{M \times M}$, such that $\mathbf{x} \in \mathbb{C}^N$, and with entries $h_i = \mathbf{x}^{\dagger} A_i \mathbf{x}$ as follows:

$$H(\mathbf{x}) := \begin{pmatrix} 0 & h_1 & \ldots & h_{M-1} \\ h_1 & 0 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ h_{M-1} & 0 & \ldots & 0 \end{pmatrix}. \tag{2.12}$$

For given $\rho = \mathbf{y}\mathbf{y}^{\dagger} \otimes \mathbf{x}\mathbf{x}^{\dagger} \in \mathcal{E}_{M,N}$, we have via straightforward manipulation that:

$$\begin{aligned} \mathrm{Tr}(C\rho) &= \sum_{i=1}^{M-1} \mathrm{Tr}\left[(\mathbf{b}_1 \mathbf{b}_{i+1}^{\dagger} + \mathbf{b}_{i+1} \mathbf{b}_1^{\dagger}) \mathbf{y}\mathbf{y}^{\dagger} \right] \mathbf{x}^{\dagger} A_i \mathbf{x} & (2.13) \\ &= \mathrm{Tr}\left[\left(\sum_{i=1}^{M-1} h_i (\mathbf{b}_1 \mathbf{b}_{i+1}^{\dagger} + \mathbf{b}_{i+1} \mathbf{b}_1^{\dagger}) \right) \mathbf{y}\mathbf{y}^{\dagger} \right] & (2.14) \\ &= \mathrm{Tr}\left(H(\mathbf{x}) \mathbf{y}^{\dagger} \mathbf{y} \right) & (2.15) \\ &= \mathbf{y}^{\dagger} H(\mathbf{x}) \mathbf{y}. & (2.16) \end{aligned}$$

It follows that the maximum attainable value for $\mathrm{Tr}(C\rho)$ is the largest eigenvalue of $H(\mathbf{x})$, which we denote by $\lambda_{\max}(H(\mathbf{x}))$. We thus have:

$$\max_{\rho \in \mathcal{E}_{M,N}} \mathrm{Tr}(C\rho) = \max_{\mathbf{x} \in \mathbb{C}^N, \|\mathbf{x}\|_2 = 1} \lambda_{\max}\left(H(\mathbf{x}) \right). \tag{2.17}$$

To determine $\lambda_{\max}(H(\mathbf{x}))$, we set up the eigenvector equation for $H(\mathbf{x})$ for some $\mathbf{y} \in \mathbb{C}^M$, such that

$$H(\mathbf{x})\mathbf{y} = \begin{pmatrix} \sum_{k=1}^{M-1} h_k y_{k+1} \\ h_1 y_1 \\ \vdots \\ h_{M-1} y_1 \end{pmatrix} = \lambda \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_M \end{pmatrix}. \tag{2.18}$$

Using straightforward substitution to solve the resulting system of equations, we find that $\lambda^2 = \sum_{k=1}^{M-1} h_k^2$, from which the claim follows. □

With Lemma 4 in hand, we can proceed to prove the second link in Equation 2.6.

Lemma 5. *There exists a polynomial time Karp reduction which maps instance* $\Pi_1 = (k, l, B_1, \ldots, B_k, \zeta, \eta)$ *of* RSDF *to instance* $\Pi_2 = (\mathbf{c}, \gamma)$ *of* $\text{WOPT}_\epsilon(\mathcal{S}_{M,N})$, *where we define for convenience* $\Delta := \sqrt{2 \sum_{i=1}^{k} \|B_i\|_{\text{F}}^2}$, *and such that:*

- $M = k + 1$
- $N = \frac{l(l-1)}{2} + 1$
- $\mathbf{c} = \hat{\mathbf{c}} / \|\hat{\mathbf{c}}\|_2 \in \mathbb{Q}^m$ *with* $\|\hat{\mathbf{c}}\|_2 \in O(m^{1/2}\Delta)$ *and* $m = M^2N^2 - 1$
- $\gamma = \frac{1}{2\|\hat{\mathbf{c}}\|_2}(\sqrt{\zeta + \eta} + \sqrt{\zeta - \eta})$
- $\epsilon \leq \frac{\sqrt{\zeta+\eta} - \sqrt{\zeta-\eta}}{4\|\hat{\mathbf{c}}\|_2(MN-1)+1}$

Proof. Given instance Π_1 of RSDF, we would like to map it to an instance Π_2 of $\text{WOPT}_\epsilon(\mathcal{S}_{M,N})$. Following the argument in [43], we begin by calling Lemma 4 with parameters $M = k + 1$, $N = \frac{l(l-1)}{2} + 1$, and each $A_i \in \mathbb{R}^{N \times N}$ symmetric and all zeroes except for its upper-left $l \times l$-dimensional submatrix, which we set to B_i, giving

$$\sqrt{g(B_1, \ldots, B_k)} = \max_{\rho \in \mathcal{S}_{M,N}} \text{Tr}(C\rho). \tag{2.19}$$

Thus, Π_1 is reduced to maximizing the linear objective function $f(\rho) = \text{Tr}(C\rho)$, defined in Lemma 4, over all $\rho \in \mathcal{S}_{M,N}$. For notational convenience, we define $f_{\max} = \max_{\rho \in \mathcal{S}_{M,N}} f(\rho)$. It is easy to see that $\|C\|_{\text{F}} = \Delta$, as defined in the statement of our claim.

Next, use Equation 2.5 and the fact that $\text{Tr}(C) = 0$ to rephrase $f(\rho)$ in terms of Bloch vectors, as required by $\text{WOPT}_\epsilon(\mathcal{S}_{M,N})$, such that

$$f(\rho) = \frac{1}{2} \sum_{i=1}^{M^2N^2-1} r_i \cdot \text{Tr}(C\sigma_i) = \hat{\mathbf{c}}^{\text{T}}\mathbf{r}, \tag{2.20}$$

for σ_i the generators of $SU(MN)$, $\hat{c}_i = \frac{1}{2}\text{Tr}(C\sigma_i)$, and $\mathbf{r} \in \mathbb{R}^{M^2N^2-1}$ the Bloch vector of ρ. Hence, set $m = M^2N^2 - 1$ and $\mathbf{c} = \hat{\mathbf{c}} / \|\hat{\mathbf{c}}\|_2$. Since $\sigma_i\sigma_j = 2\delta_{ij}$ [45], where δ_{ij}

denotes the Kronecker delta, it follows from Equation 2.20 and the Cauchy-Schwarz inequality that $\|\hat{\mathbf{c}}\|_2 \in O(m^{1/2}\Delta)$. Further, unless C is the zero matrix (i.e. each B_i is a zero matrix), $\|\hat{\mathbf{c}}\|_2 > 0$. To account for the normalization[4] of $\mathbf{c}$, define modified objective functions $\tilde{f}(\mathbf{r}) = \mathbf{c}^{\mathrm{T}}\mathbf{r}$, and $\tilde{f}_{\max} = \max_{\mathbf{r}\in\mathcal{S}_{M,N}} \tilde{f}(\mathbf{r})$, where by $\mathbf{r} \in \mathcal{S}_{M,N}$, we denote a Bloch vector corresponding to some separable state $\rho \in \mathcal{S}_{M,N}$.

It hence remains to show the following (for γ and ϵ to be chosen): If we have $\tilde{f}_{\max} \geq \|\hat{\mathbf{c}}\|_2^{-1}\sqrt{\zeta+\eta}$, then there exists an $\mathbf{r} \in S(\mathcal{S}_{M,N}, -\epsilon)$ such that $\tilde{f}(\mathbf{r}) \geq \gamma + \epsilon$ (i.e. a "YES" instance of RSDF implies a "YES" instance of WOPT$_\epsilon(\mathcal{S}_{M,N})$). If $\tilde{f}_{\max} \leq \|\hat{\mathbf{c}}\|_2^{-1}\sqrt{\zeta-\eta}$, on the other hand, then for all $\mathbf{r} \in S(\mathcal{S}_{M,N}, \epsilon)$, $\tilde{f}(\mathbf{r}) \leq \gamma - \epsilon$ (i.e. a "NO" instance of RSDF implies a "NO" instance of WOPT$_\epsilon(\mathcal{S}_{M,N})$). We proceed case by case. Set $\gamma = \frac{1}{2\|\hat{\mathbf{c}}\|_2}(\sqrt{\zeta+\eta} + \sqrt{\zeta-\eta})$, and let us choose ϵ as needed.

- Case 1: $\tilde{f}_{\max} \geq \frac{1}{\|\hat{\mathbf{c}}\|_2}\sqrt{\zeta+\eta}$.

  Let $\mathbf{r}^* \in \mathcal{S}_{M,N}$ be such that $\tilde{f}(\mathbf{r}^*) = \tilde{f}_{\max}$. To find an $\mathbf{r}' \in S(\mathcal{S}_{M,N}, -\epsilon)$ such that $\tilde{f}(\mathbf{r}') \geq \gamma + \epsilon$, we first use the fact that for any well-bounded origin-centered convex set K, it holds that for all $\mathbf{x} \in K$, there exists a $\mathbf{y} \in S(K, -\epsilon)$ such that $\|\mathbf{x} - \mathbf{y}\|_2 \leq 2\epsilon R/r$ [29]. From the definitions of r and R for $\mathcal{S}_{M,N}$ from Section 2.2, it follows that there exists an $\mathbf{r}' \in S(\mathcal{S}_{M,N}, -\epsilon)$ such that $\|\mathbf{r}' - \mathbf{r}^*\|_2 \leq 2(MN-1)\epsilon$. Since $\tilde{f}$ is linear, we can then write:

  $$\left|\tilde{f}(\mathbf{r}') - \tilde{f}(\mathbf{r}^*)\right| = \left|\mathbf{c}^{\mathrm{T}}(\mathbf{r}' - \mathbf{r}^*)\right| \leq \|\mathbf{c}\|_2 \|\mathbf{r}' - \mathbf{r}^*\|_2 \leq 2(MN-1)\epsilon, \tag{2.21}$$

  where the first inequality follows from the Cauchy-Schwarz inequality. Thus, in order to have $\tilde{f}(\mathbf{r}') \geq \gamma + \epsilon$ as desired, it suffices to have

  $$\tilde{f}(\mathbf{r}') \geq \tilde{f}_{\max} - 2(MN-1)\epsilon \geq \gamma + \epsilon, \tag{2.22}$$

  into which substitution of our values for γ and $\tilde{f}_{\max}$ gives that setting

  $$\epsilon \leq \frac{\sqrt{\zeta+\eta} - \sqrt{\zeta-\eta}}{4\|\hat{\mathbf{c}}\|_2(MN-1)+1} \tag{2.23}$$

  suffices to conclude we have a "YES" instance of WOPT$_\epsilon(\mathcal{S}_{M,N})$.

- CASE 2: $\tilde{f}_{\max} \leq \frac{1}{\|\hat{\mathbf{c}}\|_2}\sqrt{\zeta-\eta}$.

  Let $\mathbf{r}^* \in \mathcal{S}_{M,N}$ be such that $\tilde{f}(\mathbf{r}^*) = \tilde{f}_{\max}$. To see that for all $\mathbf{r}' \in S(\mathcal{S}_{M,N}, \epsilon)$, we have $\tilde{f}(\mathbf{r}') \leq \gamma - \epsilon$, let $\mathbf{r}' \in S(\mathcal{S}_{M,N}, \epsilon)$. Then by the definition of $S(\mathcal{S}_{M,N}, \epsilon)$,

[4] Though normalizing $\hat{\mathbf{c}}$ will slightly complicate our expressions, dropping the normalization requirement in the definition of WOPT$_\epsilon(K)$ will not simplify matters — not only would we have to account for $\|c\|_2$ later in this reduction anyway, we would also require slight modifications to the reduction WOPT$_\epsilon(\mathcal{S}_{M,N}) \leq_C$ WMEM$_\beta(\mathcal{S}_{M,N})$ later, and it would negatively impact parameter scaling for the latter reduction (by a polynomial factor).

there exists some $\mathbf{r} \in \mathcal{S}_{M,N}$ such that $\|\mathbf{r}' - \mathbf{r}\|_2 \leq \epsilon$. By Equation 2.21, it follows that

$$\left|\tilde{f}(\mathbf{r}') - \tilde{f}(\mathbf{r})\right| \leq \epsilon. \tag{2.24}$$

Thus, by considering $\mathbf{r} = \mathbf{r}^*$, we can conclude that $\tilde{f}(\mathbf{r}') \leq \tilde{f}_{\max} + \epsilon$ for any $\mathbf{r}' \in S(\mathcal{S}_{M,N}, \epsilon)$. To achieve $\tilde{f}(\mathbf{r}') \leq \gamma - \epsilon$ then, we set $\tilde{f}(\mathbf{r}') \leq \tilde{f}_{\max} + \epsilon \leq \gamma - \epsilon$, into which substitution of our values for γ and $\tilde{f}_{\max}$ yields that setting

$$\epsilon \leq \frac{\sqrt{\zeta + \eta} - \sqrt{\zeta - \eta}}{2\,\|\hat{\mathbf{c}}\|_2 + 2} \tag{2.25}$$

suffices to conclude we have a "NO" instance of $\mathrm{WOPT}_\epsilon(\mathcal{S}_{M,N})$.

□

Lemma 5 gives us the second link in Equation 2.6, RSDF $\leq_K$ $\mathrm{WOPT}_\epsilon(\mathcal{S}_{M,N})$. Observe that combining Lemma 2 and Lemma 5 gives $M = N = \frac{n(n-1)}{2} + 1$, and by an argument of Ioannou [43], one can in fact have Lemma 5 holding for $M \geq N$ [43] by padding the matrix C (from the proof of Lemma 5) with extra $N \times N$-dimensional zero matrices. Thus, the hardness result we will show for $\mathrm{WMEM}_\beta(\mathcal{S}_{M,N})$ by building on this link will be valid in general for $M \geq N$.

Let us also note that we have not lower bounded $\|\hat{\mathbf{c}}\|_2$ in Lemma 5, which appears in the expressions for γ and ϵ. This is not a problem, as we are interested in a lower bound for ϵ, since we need $\epsilon \geq 1/\operatorname{poly}(M, N)$. We return to this point later, but the curious reader is referred to Appendix 4, where we derive an exact expression for $\|\hat{\mathbf{c}}\|_2$ in the context of our reduction from CLIQUE, from which it is clear $\|\hat{\mathbf{c}}\|_2 \geq \sqrt{2}$ (except in the trivial case of an input graph for CLIQUE with no edges, in which case $\|\hat{\mathbf{c}}\|_2 = 0$).

To show the last two links of Equation 2.6, we now follow the non-ellipsoidal Turing reduction of Liu [48] of $\mathrm{WOPT}_\epsilon(K) \leq_T \mathrm{WSEP}_\nu(K) \leq_T \mathrm{WMEM}_\beta(K)$, which holds for arbitrary p-centered well-bounded compact convex set $K \subseteq \mathbb{R}^m$, specified using parameters R, r, and $\mathbf{p}$ (as defined in Section 2.2). The first of these two links is given as follows.

Lemma 6 (Liu [48]). *Given instance* $\Pi = (\mathbf{c}, \gamma)$ *of* $\mathrm{WOPT}_\epsilon(K)$*, for* K *with associated parameters* (m, R, r, p)*,* $0 < \epsilon < 1$*, and* $\|\mathbf{c}\|_2 = 1$*, there exists an algorithm which runs in time* $\operatorname{poly}(\langle K\rangle, R, \lceil 1/\epsilon\rceil)$*, and solves* Π *using an oracle for* $\mathrm{WSEP}_\nu(K)$ *with* $\nu = \epsilon/3$*.*

Proof. We follow [48]. Define convex set $\hat{K} := \left\{\mathbf{y} \in K \mid \mathbf{c}^{\mathrm{T}}\mathbf{y} \geq \gamma\right\}$. Note we have not used the error parameter ϵ in defining $\hat{K}$. We first show that one can build an oracle for $\mathrm{WSEP}_\nu(\hat{K})$ using a given oracle for $\mathrm{WSEP}_\nu(K)$, and then use the former to state the desired algorithm.

Consider the following algorithm for constructing an oracle for $\mathrm{WSEP}_\nu(\hat{K})$.

Algorithm 1 (WSEP$_\nu(\hat{K})$)**.** Given input $\mathbf{z} \in \mathbb{Q}^m$, rational $\nu > 0$, and an oracle Q for WSEP$_\nu(K)$, run Q on input $\mathbf{z}$ and ν, and respond as follows:

1. If Q outputs "YES":

    (a) If $\mathbf{c}^{\mathrm{T}}\mathbf{z} < \gamma$, return $\hat{\mathbf{c}} = -\mathbf{c}$.

    (b) Otherwise, output "YES".

2. If Q returns (approximately) separating hyperplane $\hat{\mathbf{c}}$, return $\hat{\mathbf{c}}$.

Analysis. If Q answers "YES" (Case 1), then $\mathbf{z} \in S(K, -\nu)$, but not necessarily $\mathbf{z} \in S(\hat{K}, -\nu)$. If $\mathbf{c}^{\mathrm{T}}\mathbf{z} < \gamma$ (Case (1a)), in fact, $\mathbf{c}$ is a separating hyperplane for $\hat{K}$. To see this, recall that by the definition of WSEP$_\nu(\hat{K})$, we want to return $\hat{\mathbf{c}}$ such that for all $\mathbf{x} \in S(\hat{K}, -\nu)$, $\hat{\mathbf{c}}^{\mathrm{T}}\mathbf{x} - \hat{\mathbf{c}}^{\mathrm{T}}\mathbf{z} \leq \nu$. Choosing $\hat{\mathbf{c}} = -\mathbf{c}$ hence gives $\hat{\mathbf{c}}^{\mathrm{T}}\mathbf{x} - \hat{\mathbf{c}}^{\mathrm{T}}\mathbf{z} < 0$ for any $\mathbf{x} \in \hat{K}$, as desired. If, on the other hand, $\mathbf{c}^{\mathrm{T}}\mathbf{z} \geq \gamma$ (Case (1b)), then $\mathbf{z} \in \hat{K}$, and we can simply output "YES", since if $\mathbf{z} \in S(\hat{K}, -\nu)$, then "YES" is the correct answer, and otherwise we are in an "error region" where the oracle is allowed to answer either way.

If we are in Case 2, on the other hand, meaning Q returns an (approximately) separating hyperplane $\hat{\mathbf{c}}$ for $S(K, -\nu)$, then $\mathbf{z} \notin S(K, \nu)$, implying $\mathbf{z} \notin S(\hat{K}, \nu)$. Thus, since $\hat{K} \subseteq K$, it holds that $S(\hat{K}, -\nu) \subseteq S(K, -\nu)$, and $\hat{\mathbf{c}}$ is also an (approximately) separating hyperplane for $S(\hat{K}, -\nu)$. □

With Algorithm 1 in hand, we can state the desired algorithm for WOPT$_\epsilon(K)$.

Algorithm 2. Given input $\mathbf{c} \in \mathbb{Q}^m$ with $\|\mathbf{c}\|_2 = 1$, $\gamma \in \mathbb{Q}$, rational $\epsilon > 0$, and an oracle Q for WSEP$_\nu(\hat{K})$, choose rational $0 < \nu < \epsilon$ and proceed as follows:

1. Let $\mathbf{a}$ initially be the origin in $\mathbb{Q}^m$ (i.e. $\mathbf{a} = (0, \ldots, 0)$).

2. Call Q on input $\mathbf{a}$.

    (a) If Q returns "YES", output "YES" and stop.

    (b) Otherwise, Q returns (approximately) separating hyperplane $\hat{\mathbf{c}}$. Update $\mathbf{a} = \mathbf{a} - (\epsilon - 2\nu)\hat{\mathbf{c}}$.

3. Go back to line 2 if number of completed iterations is less than $\left\lceil \frac{R^2}{(\epsilon - 2\nu)^2} \right\rceil$.

4. Output "NO".

Let us first demonstrate the correctness of Algorithm 2. Consider first the case of Π a "NO" instance of WOPT$_\epsilon(K)$, meaning $\mathbf{c}^{\mathrm{T}}\mathbf{x} \leq \gamma - \epsilon$ for all $\mathbf{x} \in S(K, \epsilon)$. We show by contradiction that WSEP$_\nu(\hat{K})$ always outputs "NO". Suppose that on input $\mathbf{a}$, Q returns "YES". This implies $\mathbf{a} \in S(\hat{K}, \nu)$, since WSEP$_\nu(\hat{K})$ can answer "YES" if $\mathbf{a} \in S(\hat{K}, -\nu)$ or if $\mathbf{a}$ is in the "error region" $S(\hat{K}, \nu) \backslash S(\hat{K}, -\nu)$. Since

$\|\mathbf{c}\|_2 = 1$, however, this gives $\mathbf{c}^{\mathrm{T}}\mathbf{a} \geq \gamma - \nu$, which is a contradiction since $\nu < \epsilon$. Thus, Algorithm 2 will return “NO”, as desired.

Now suppose Π is a “YES” instance of $\mathrm{WOPT}_\epsilon(K)$, meaning there exists $\mathbf{y} \in S(K, -\epsilon)$ with $\mathbf{c}^{\mathrm{T}}\mathbf{y} \geq \gamma + \epsilon$. We show that with each iteration of line 2 of Algorithm 2, $\|\mathbf{y} - \mathbf{a}\|_2$ decreases, and so we will eventually output “YES”. To do so, let $\mathbf{a}^{(i)}$ be $\mathbf{a}$ after i iterations, and examine iteration $i+1$. If Q returns “YES”, then by our argument for the “NO” case, we can safely output “YES” for $\mathrm{WOPT}_\epsilon(K)$, as desired. If, on the other hand, Q returns hyperplane $\hat{\mathbf{c}}$, we have by the definition of $\mathrm{WSEP}_\nu(\hat{K})$ that for all $\mathbf{x} \in (\hat{K}, -\nu)$, $\hat{\mathbf{c}}^{\mathrm{T}}\mathbf{x} \leq \hat{\mathbf{c}}^{\mathrm{T}}\mathbf{a}^{(i)} + \nu$. Consider then the following, which follows from straightforward manipulation and the fact that $\|\hat{\mathbf{c}}\|_2 = 1$:

$$\left\|\mathbf{a}^{(i+1)} - \mathbf{y}\right\|_2^2 = \left\|\mathbf{a}^{(i)} - \mathbf{y}\right\|_2^2 + (\epsilon - 2\nu)^2 + 2(\epsilon - 2\nu)\left[\hat{\mathbf{c}}^{\mathrm{T}}(\mathbf{y} - \mathbf{a}^{(i)})\right]. \tag{2.26}$$

If we can show $\left[\hat{\mathbf{c}}^{\mathrm{T}}(\mathbf{y} - \mathbf{a}^{(i)})\right] \leq -(\epsilon - 2\nu)$, then we have

$$\left\|\mathbf{a}^{(i+1)} - \mathbf{y}\right\|_2^2 \leq \left\|\mathbf{a}^{(i)} - \mathbf{y}\right\|_2^2 - (\epsilon - 2\nu)^2, \tag{2.27}$$

so each iteration brings us closer to $\mathbf{y}$, as desired. To show this, consider the point $\mathbf{x} = \mathbf{y} + (\epsilon - \nu)\hat{\mathbf{c}}$. Since by assumption we have $\mathbf{c}^{\mathrm{T}}\mathbf{y} \geq \gamma + \epsilon$ and $\|\mathbf{c}\|_2 = 1$, it is easy to see that $S(\mathbf{y}, \epsilon) \subseteq \hat{K}$, and so it follows that $\mathbf{x} \in S(\hat{K}, -\nu)$. By the definition of $\hat{\mathbf{c}}$, we thus have through straightforward manipulation that $\left[\hat{\mathbf{c}}^{\mathrm{T}}(\mathbf{y} - \mathbf{a}^{(i)})\right] \leq -(\epsilon - 2\nu)$, as desired. Finally, recalling that (by assumption) K must be contained in an origin-centered ball of radius R, we know that $\left\|\mathbf{a}^{(0)} - \mathbf{y}\right\|_2^2 \leq R^2$, and so at most $\left\lceil \frac{R^2}{(\epsilon - 2\nu)^2} \right\rceil$ will suffice to return “YES” (eventually we will fall into $S(\mathbf{y}, \epsilon) \subseteq \hat{K}$, at which point the oracle Q must return “YES”). This completes the correctness argument.

We close by remarking that one can set, for example [48], $\nu = \epsilon/3$ for Algorithm 2, giving a runtime of $\mathrm{poly}(m, R, \lceil 1/\epsilon \rceil)$. □

Recall now our claim at the outset of this section that one must choose ϵ for $\mathrm{WOPT}_\epsilon(K)$ such that $\epsilon \geq 1/\mathrm{poly}(M, N)$ in order for the entire reduction chain of Equation 2.6 to run in polynomial time — the justification for this is now clear due to the dependence on $1/\epsilon$ in the runtime of Lemma 6. Let us observe that we can choose ϵ so. Specifically, by Lemma 5, we can set

$$\epsilon = \frac{\sqrt{\zeta + \eta} - \sqrt{\zeta - \eta}}{4\left\|\hat{\mathbf{c}}\right\|_2 (MN - 1) + 1}. \tag{2.28}$$

Piecing together Lemma 2 and Lemma 5, we have $\zeta \in \Theta(1)$ and $\eta \in \Omega(1/N)$. It follows that $\sqrt{\zeta + \eta} - \sqrt{\zeta - \eta} \in \Omega(1/N)$. To see this, recall that for given functions $f(x)$ and $g(x)$, $f(x) \in \Omega(g(x))$ if $f(x) \geq cg(x) \geq 0$ for all $x > x_0$ and positive constants c and x_0 [18]. Then one has, for positive constants c_1, c_2, N_1, and N_2:

$$\begin{aligned}
\sqrt{\zeta+\eta}-\sqrt{\zeta-\eta} &\geq \sqrt{c_1+\frac{c_2}{N}}-\sqrt{c_1-\frac{c_2}{N}} \quad \forall\ N \geq \max(N_1,N_2) && (2.29)\\
&= \frac{(c_1+\frac{c_2}{N})-(c_1-\frac{c_2}{N})}{\sqrt{c_1+\frac{c_2}{N}}+\sqrt{c_1-\frac{c_2}{N}}} && (2.30)\\
&= \frac{2c_2}{N\sqrt{c_1+\frac{c_2}{N}}+\sqrt{c_1-\frac{c_2}{N}}} && (2.31)\\
&\geq \frac{2c_2}{N(\sqrt{c_1+c_2}+\sqrt{c_1})}, && (2.32)
\end{aligned}$$

where in the first inequality we have used the identity $x-y=(x^2-y^2)/(x+y)$. The claim immediately follows. With a little thought, we also have $\|\hat{\mathbf{c}}\|_2 \in O(\sqrt{N})$ (see Appendix 4). Thus, we can always solve an instance of CLIQUE by choosing some $\epsilon \in \Omega(M^{-1}N^{-5/2})$ for $\mathrm{WOPT}_\epsilon(\mathcal{S}_{M,N})$, giving a polynomial runtime for the reduction of Lemma 6, as desired.

We can now move on to the final link, $\mathrm{WSEP}_\nu(\mathcal{S}_{M,N}) \leq_T \mathrm{WMEM}_\beta(\mathcal{S}_{M,N})$, which is again given by [48], and is in fact based on Lemmas (4.3.3) and (4.3.4) of [29], with a minor preprocessing step added at the beginning. Specifically, this last link in Equation 2.6 is more accurately broken up into three sub-links:

$$\mathrm{WSEP}_\nu(\mathcal{S}_{M,N}) \leq_C \mathrm{WSEP}^\delta_{\nu'}(\mathcal{S}_{M,N}) \leq_T \mathrm{WMEM}^1_{\beta'}(\mathcal{S}_{M,N}) \leq_C \mathrm{WMEM}_\beta(\mathcal{S}_{M,N}). \tag{2.33}$$

Here, $\mathrm{WSEP}^\delta_{\nu'}(\mathcal{S}_{M,N})$ and $\mathrm{WMEM}^1_{\beta'}(\mathcal{S}_{M,N})$ are variants of problems $\mathrm{WSEP}_\nu(\mathcal{S}_{M,N})$ and $\mathrm{WMEM}_\beta(\mathcal{S}_{M,N})$, respectively, defined over an arbitrary convex set K as follows.

Definition 7 (Weak(er) Separation ($\mathrm{WSEP}^\delta_{\nu'}(K)$))**.** Given $\mathbf{z}' \in \mathbb{Q}^m$, and parameters $\nu', \delta \in \mathbb{Q}$, such that $0 < \nu', \delta < 1$, respond, with respect to the complexity measure $(\langle K\rangle + \langle \mathbf{z}'\rangle + \langle \nu'\rangle + \langle \delta\rangle)$:

If $\mathbf{z}' \in S(K,-\nu')$, then output "YES".
If $\mathbf{z}' \notin S(K,\nu')$, then return $\mathbf{c} \in \mathbb{Q}^m$, such that for all $\mathbf{x} \in K$,
$\mathbf{c}^{\mathrm{T}}\mathbf{x} \leq \mathbf{c}^{\mathrm{T}}\mathbf{z}' + \nu' + \delta\,\|\mathbf{x}-\mathbf{z}'\|_2$, and $\|\mathbf{c}\|_2 = 1$.

Definition 8 (One-Sided Weak Membership ($\mathrm{WMEM}^1_{\beta'}(K)$))**.** Given $\mathbf{y}' \in \mathbb{Q}^m$, and rational parameter $\beta' \in \mathbb{Q}$, such that $\beta' > 0$, decide, with respect to the complexity measure $(\langle K\rangle + \langle \mathbf{y}'\rangle + \langle \beta'\rangle)$:

If $\mathbf{y}' \in K$, then output "YES".
If $\mathbf{y}' \notin S(K,\beta')$, then output "NO".

$\mathrm{WSEP}^\delta_{\nu'}(K)$ is a weaker form of $\mathrm{WSEP}_\nu(K)$ — namely, the (approximately) separating hyperplane $\mathbf{c}$ returned in the "NO" case allows a larger margin of error for points in K which are farther away from the input $\mathbf{z}'$. $\mathrm{WMEM}^1_{\beta'}(\mathcal{S}_{M,N})$, on the other hand, appears at first glance like a stronger version of $\mathrm{WMEM}_\beta(\mathcal{S}_{M,N})$, in which the error region has been reduced from $S(K,\beta)\backslash S(K,-\beta)$ to $S(K,\beta)\backslash K$. As will be shown later via the last link of Equation 2.33, however, $\mathrm{WMEM}_{\beta}(\mathcal{S}_{M,N})$ and $\mathrm{WMEM}^{1}_{\beta'}(\mathcal{S}_{M,N})$ are actually equivalent.

The first and third links in Equation 2.33 are fairly straightforward, and will be discussed next. The second link, however, is quite involved, and is well-established and detailed in [29]. Hence, we will suffice with stating its result and referring the reader to [29] for further details. Let us now discuss the first link of Equation 2.33, $\mathrm{WSEP}_{\nu}(\mathcal{S}_{M,N}) \leq_T \mathrm{WSEP}^{\delta}_{\nu'}(\mathcal{S}_{M,N})$.

Lemma 7 (Liu [48])**.** *Given instance* $\Pi = (\mathbf{z})$ *of* $\mathrm{WSEP}_{\nu}(K)$*, for* K *with associated parameters* (m,R,r,p)*, there exists a poly-time algorithm which solves* Π *using an oracle for* $\mathrm{WSEP}^{\delta}_{\nu'}(K)$ *with input* $\mathbf{z}'=\mathbf{z}$*,* $\nu'=\frac{\nu}{2}$*, and* $\delta=\frac{\nu}{4R}$*.*

Proof. We follow [48]. The algorithm is stated as follows.

Algorithm 3. Let $\delta = \frac{\nu}{4R}$ and $\nu' = \nu/2$. Given an oracle Q for $\mathrm{WSEP}^{\delta}_{\nu'}(K)$, proceed as follows:

1. If $\|\mathbf{z}\|_2 > R$, return (approximately) separating hyperplane $\mathbf{c} = \frac{\mathbf{z}}{\|\mathbf{z}\|_2}$.

2. Otherwise, call Q on input $\mathbf{z}'=\mathbf{z}$, and return the result.

Let us show correctness. Note first that if $\|\mathbf{z}\|_2 > R$, then clearly $\mathbf{z} \notin S(K,-\nu)$, since K is contained in an origin-centered ball of radius R. The latter point also implies that $\|\mathbf{x}\|_2 \leq R$ for all $\mathbf{x}\in K$. Thus, $\mathbf{c} = \mathbf{z}/\|\mathbf{z}\|_2$ is in fact a true (i.e. not approximate) separating hyperplane, since $\mathbf{c}^{\mathrm{T}}\mathbf{z} > R$, while $\mathbf{c}^{\mathrm{T}}\mathbf{x} \leq R$ for all $\mathbf{x}\in K$, and we can return it, as required.

Suppose then that $\|\mathbf{z}\|_2 \leq R$. If Q returns "YES", we have $\mathbf{z}\in S(K,\frac{\nu}{2})$ (including the "error region" $S(K,\frac{\nu}{2})\backslash S(K,-\frac{\nu}{2})$), and so clearly $\mathbf{z}\in S(K,\nu)$ and we can similarly answer "YES". If, on the other hand, Q returns some $\mathbf{c}$ such that for all $\mathbf{x}\in K$, $\mathbf{c}^{\mathrm{T}}\mathbf{x} \leq \mathbf{c}^{\mathrm{T}}\mathbf{z}' + \nu' + \delta\|\mathbf{x}-\mathbf{z}'\|_2$, then observing again that $\|\mathbf{x}-\mathbf{z}'\|_2 \leq 2R$ gives

$$\mathbf{c}^{\mathrm{T}}\mathbf{x} \leq \mathbf{c}^{\mathrm{T}}\mathbf{z}' + \nu' + \delta\|\mathbf{x}-\mathbf{z}'\|_2 \leq \mathbf{c}^{\mathrm{T}}\mathbf{z} + \frac{\nu}{2} + (\frac{\nu}{4R})(2R) \leq \mathbf{c}^{\mathrm{T}}\mathbf{z} + \nu, \qquad (2.34)$$

and we can correctly return $\mathbf{c}$ as an (approximately) separating hyperplane. □

Lemma 7 gives us the link $\mathrm{WSEP}_{\nu}(K) \leq \mathrm{WSEP}^{\delta}_{\nu'}(K)$ of Equation 2.33. As previously discussed, the next link, $\mathrm{WSEP}^{\delta}_{\nu'}(\mathcal{S}_{M,N}) \leq_T \mathrm{WMEM}^{1}_{\beta'}(\mathcal{S}_{M,N})$, is given by a non-ellipsoidal Turing reduction, and is stated as follows, the details of which can be found in [29] (also outlined in [48]). Note the runtime dependence on parameter $1/\delta$. In our case, combining Lemmas 6 and 7, we have $\delta = \frac{\epsilon}{12R}$, for ϵ the error parameter in Lemma 6 and R the radius of the origin-centered ball in which K sits. Since we chose $\epsilon \geq 1/\operatorname{poly}(M,N)$, and $R\in\Theta(1)$ for $K=\mathcal{S}_{M,N}$, the following reduction will also run in polynomial time.

Lemma 8 (Grötschel et al. [29, 48])**.** *Given instance* $\Pi = (\mathbf{z}')$ *of* $\text{WSEP}^{\delta}_{\nu'}(K)$*, for* K *with associated parameters* (m, R, r, p)*, and* $0 < \nu', \delta < 1$*, there exists an algorithm which runs in time* $\text{poly}(\langle K \rangle, \langle \mathbf{z}' \rangle, \langle \nu' \rangle, \lceil 1/\delta \rceil)$*, and solves* Π *using an oracle for* $\text{WMEM}^{1}_{\beta'}(K)$ *with error parameter* $\beta' = \frac{\delta^2 r^2 \nu'}{2^6 m^5 R(R+r)}$*.*

This leaves us with the last link of Equation 2.33, which is also given in [29], and is written $\text{WMEM}^{1}_{\beta'}(\mathcal{S}_{M,N}) \leq_T \text{WMEM}_{\beta}(\mathcal{S}_{M,N})$. Showing this reduction will complete the desired chain of Equation 2.6.

Lemma 9 (Grötschel et al. [29])**.** *Given instance* $\Pi = (\mathbf{y}')$ *of* $\text{WMEM}^{1}_{\beta'}(K)$*, for* K *with associated parameters* (m, R, r, p) *and* $\beta' > 0$*, there exists a poly-time algorithm which solves* Π *using an oracle for* $\text{WMEM}_{\beta}(K)$ *with* $\beta = \frac{r\beta'}{4R}$*.*

Proof.

Algorithm 4. Let $\mathbf{y} = (1 - \frac{\beta'}{4R})\mathbf{y}' + \frac{\beta'}{4R}\mathbf{p}$ and $\beta = \frac{r\beta'}{4R}$. Given an oracle Q for $\text{WMEM}_{\beta}(K)$, proceed as follows:

1. If $\|\mathbf{y}' - \mathbf{p}\|_2 \geq 2R$, output "NO".
2. Otherwise, call Q on input $\mathbf{y}$, and return the result.

To show correctness, suppose first that we have Case 1, $\|\mathbf{y}' - \mathbf{p}\|_2 \geq 2R$. Then clearly $\mathbf{y}' \notin K$ (since by assumption K is circumscribed in a ball of radius R), and we can safely answer "NO" (including $\mathbf{y}'$ possibly in the error region).

We are left with Case 2. We must show that if $\mathbf{y}' \in K$, then $\mathbf{y} \in S(K, -\beta)$ (a "YES" instance of $\text{WMEM}^{1}_{\beta'}(K)$ implies a "YES" instance of $\text{WMEM}_{\beta}(K)$), and if otherwise we have $\mathbf{y}' \notin S(K, \beta')$, then $\mathbf{y} \notin S(K, \beta)$ (a "NO" instance of $\text{WMEM}^{1}_{\beta'}(K)$ implies a "NO" instance of $\text{WMEM}_{\beta}(K)$).

To show the first, suppose $\mathbf{y}' \in K$. Recall that by assumption, $S(\mathbf{p}, r) \subseteq K$. Since K is convex, and since $\mathbf{y}$ is a convex combination of $\mathbf{y}'$ and $\mathbf{p}$ with weight $\frac{\beta'}{4R}$ on $\mathbf{p}$, there is a scaled down ball of radius $r(\frac{\beta'}{4R}) = \beta$ around $\mathbf{y}$ and contained in K, as depicted by Figure 3.2. Thus, $\mathbf{y} \in S(K, -\beta)$, and we answer "YES", as required.

For the "NO" case, suppose by contrapositive that $\mathbf{y} \in S(K, \beta)$. Then there exists $\mathbf{x} \in K$ such that $\|\mathbf{x} - \mathbf{y}\|_2 \leq \beta$. Observe now that

$$\|\mathbf{y}' - \mathbf{y}\|_2 = \frac{\beta'}{4R} \|\mathbf{y}' - \mathbf{p}\|_2 < \frac{\beta'}{4R}(2R) = \frac{\beta'}{2}, \tag{2.35}$$

since by Line 1 of Algorithm 4, we can assume $\|\mathbf{y}' - \mathbf{p}\|_2 < 2R$. Therefore, there exists $\mathbf{x} \in K$ such that $\|\mathbf{x} - \mathbf{y}'\|_2 \leq \beta + \frac{\beta'}{2} < \beta'$, meaning $\mathbf{y}' \in S(K, \beta')$, which is a contradiction. Thus, a "NO" instance of $\text{WMEM}^{1}_{\beta'}(K)$ implies a "NO" instance of $\text{WMEM}_{\beta}(K)$, as required. □

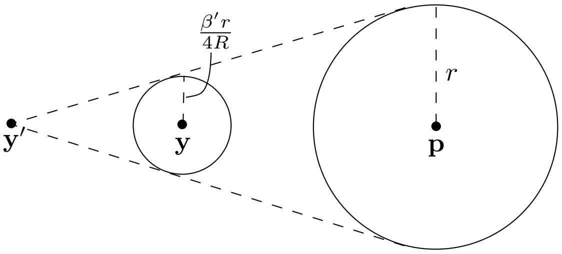

Figure 2.2: Depiction of points $\mathbf{y}, \mathbf{y}', \mathbf{p} \in K$, the assumed ball of radius r in K centered at $\mathbf{p}$, and by convexity, the resulting ball of radius $\frac{\beta' r}{4R}$ centered about $\mathbf{y}$.

We have now seen all the links in the reduction chains given by Equation 2.33, and in turn Equation 2.6. Note that a number of these reductions (specifically, all the Turing reductions) were shown for arbitrary well-bounded p-centered convex set K. Setting $K = \mathcal{S}_{M,N}$ for these reductions gives us exactly what we want — a Turing reduction from CLIQUE to $\text{WMEM}_\beta(\mathcal{S}_{M,N})$ where error parameter $\beta \geq 1/\operatorname{poly}(M,N)$, the latter following immediately from our choice of $\epsilon \geq 1/\operatorname{poly}(M,N)$ in Lemma 6. Specifically, we have the following result, from which Theorem 1 immediately follows:

Theorem 10. *Given instance $\Pi = (G, n, c)$ of* CLIQUE*, there exists an algorithm which solves Π in time* $\operatorname{poly}(n)$ *using an oracle for* $\text{WMEM}_\beta(\mathcal{S}_{M,N})$ *with parameters $M = N = \frac{n(n-1)}{2} + 1$ and some $\beta \in \Omega(n^{-73})$. More generally, for $N = \frac{n(n-1)}{2} + 1$ and any choice of $M \geq N$, the result holds for some $\beta \in \Omega(M^{-16}N^{-20.5})$.*

Theorem 1 (*revised*)**.** $\text{WMEM}_\beta(\mathcal{S}_{M,N})$ *is NP-hard for $\beta \leq \operatorname{poly}(M^{-16}N^{-20.5})$ and $M \geq N$, or equivalently, is strongly NP-hard.*

We stress the phrase "*some* $\beta \in \Omega(n^{-73})$" in the statement of Theorem 10 above — specifically, we cannot have $\beta \in O(1)$ in our reduction, due, for example, to the expression for ϵ in Lemma 5. Moreover, note that after the second link of Equation 2.6, we had $\epsilon \in \Omega(M^{-1}N^{-5/2})$, whereas our final estimate in Theorem 10 is $\beta \in \Omega(M^{-16}N^{-20.5})$. The main contributor to this parameter scaling is the reduction of Lemma 8, where we had for error parameter β' of $\text{WMEM}^1_{\beta'}(K)$ that $\beta' = \frac{\delta^2 r^2 \nu'}{2^6 m^5 R(R+r)}$. It would be interesting to determine whether the bound on β given by Theorem 10 could be improved, perhaps to $\beta \in O(1)$, which we discuss shortly in Section 2.5.

2.4 Applications

We now observe two applications of Theorem 10. First, one immediately has a lower bound on the maximum distance a bound entangled state can have from $\mathcal{S}_{M,N}$. To see this, recall that bound entangled states are mixed entangled quantum states from which no pure (state) entanglement can be distilled [36], and are the only entangled states whose entanglement is not detected by the Peres-Horodecki Positive Partial Transpose (PPT) [55, 35] criterion (the converse question of whether there exist bound entangled states detectable by the PPT test is, however, a major open question [13]). Now, Theorem 1 implies that unless $P \neq NP$, any test of membership for $\mathcal{S}_{M,N}$ must be unable to efficiently resolve $\mathcal{S}_{M,N}$ within distance $\beta \in \Omega(M^{-16}N^{-20.5})$ of its border in the general case. It follows that unless $P = NP$, there must exist bound entangled state(s) ρ_{be} such that for any separable state ρ_{sep}, $\|\rho_{be} - \rho_{sep}\|_{\mathrm{F}} \in \Omega(M^{-16}N^{-20.5})$ — if not, one could determine the separability of any quantum state within this region using the PPT test, contradicting Theorem 1. For this reason, it would be of interest to determine precisely how large one can make β before NP-hardness of $\mathrm{WMEM}_\beta(\mathcal{S}_{M,N})$ ceases to hold.

We next use the known duality between linear operators and linear maps in order to obtain immediate hardness results involving entanglement breaking maps. Specifically, let $\mathcal{D}^N$ denote the (convex) set of density operators corresponding to quantum states of dimension N. Then, for arbitrary linear map $\Phi : \mathcal{D}^N \mapsto \mathcal{D}^N$, one can associate a unique operator ρ_Φ acting in the larger Hilbert space $\mathcal{H}^N \otimes \mathcal{H}^N$ via the Jamiołkowski isomorphism[44]:

$$\Phi : \mathcal{D}^N \mapsto \mathcal{D}^N \quad \longleftrightarrow \quad \rho_\Phi \equiv [\Phi \otimes I] \left(|\phi^+\rangle\langle\phi^+| \right), \tag{2.36}$$

where $|\phi^+\rangle$ is the maximally entangled state $|\phi^+\rangle = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} |k\rangle \otimes |k\rangle$. In particular, if (and only if) Φ is completely positive, then ρ_Φ is positive semidefinite [14], and if (and only if) Φ is trace-preserving, then $\mathrm{Tr}_A(\rho_\Phi) = I/N$ [6]. Leveraging this isomorphism then, it is known that the (convex) set of separable quantum states $S_{N,N}$ is isomorphic to the (convex) set of *super-positive* maps [2, 6], also known as *entanglement-breaking* maps, acting on the space $\mathcal{D}^N$. Intuitively, such maps have the property that for any input state $\rho \in \mathcal{H}^N \otimes \mathcal{H}^N$, $(\Phi \otimes I)(\rho)$ is separable, but formally, they can be defined in any of the following equivalent ways:

Definition 9 (Entanglement-Breaking Maps [37]). Given completely positive (but not necessarily trace-preserving) linear map $\Phi : \mathcal{D}^N \mapsto \mathcal{D}^N$, the following are equivalent:

1. Φ is entanglement-breaking (EB).

2. ρ_Φ is separable.

3. Φ can be written $\Phi(\rho) = \sum_k \sigma_k \mathrm{Tr}(F_k \rho)$, for σ_k density matrices and F_k positive semidefinite[5].

4. Φ can be expressed in operator sum form with Kraus operators [46] strictly of rank 1.

5. $\Gamma \circ \Phi$ and $\Phi \circ \Gamma$ are completely positive for all positive maps Γ.

[5] This is sometimes dubbed the Holevo form, as introduced by Holevo [34].

6. $\mathrm{Tr}(\rho_\Phi \rho_\Gamma) \geq 0$ for all positive maps Γ.

Point (2) of Definition 9 and Equation 2.36 together say the following — if we assume map Φ is given to us in its Jamiołkowski representation (i.e. in terms of ρ_Φ), then checking whether Φ is entanglement-breaking is equivalent to solving the quantum separability problem. More formally, if we denote the (convex) set of entanglement breaking maps $\Phi : \mathcal{D}^N \mapsto \mathcal{D}^N$ as B_N, then one could consider the Weak Membership problem over B_N, denoted $WMEM_{\beta'}(B_N)$, where the input Φ is given in terms of ρ_Φ. Then, Equation 2.36 immediately implies the (now trivial) Karp reduction

$$WMEM_\beta(S_{N,N}) \leq_K WMEM_{\beta'}(B_N), \tag{2.37}$$

where we can clearly take $\beta = \beta'$. Thus, we have the following immediate corollary to Theorem 1.

Corollary 11. *$WMEM_{\beta'}(B_N)$ is strongly NP-hard.*

Observe that Corollary 11 in turn implies that each of the other tasks listed as equivalent definitions of entanglement-breaking maps in Definition 9 must also be NP-hard (with respect to their appropriate formal definitions).

We remark that any entanglement-breaking map Φ is completely positive [6], but by Definition 9, Φ is not necessarily trace-preserving. An analogous version of Definition 9 for trace-preserving Φ, however, can be made to hold [37] simply by enforcing $\{F_k\}$ to be a Positive Operator Valued Measure (POVM) in point (3), $\mathrm{Tr}_A(\rho_\Phi) = I/N$ in point (2), the Kraus operators K_i in point (4) to satisfy $\sum_i K_i^\dagger K_i = I$, and Γ in point (5) to be trace-preserving. An interesting open question is thus whether an efficient algorithm for determining whether a map is entanglement-breaking exists for the special case of Φ both completely positive and trace-preserving, the latter two properties defining what are known as quantum channels or stochastic maps [6]. Specifically, quantum channels are an important case of interest, as they correspond to physically realizable processes. Observing that this question reduces to solving QUSEP for the special case of input ρ with a maximally mixed reduced state, one possible approach may be to consider local filters [28], through which any positive *definite* density matrix $\rho \in \mathcal{H}^M \otimes \mathcal{H}^N$ can be transformed into the following (un-normalized) form [47] via invertible operators C and D:

$$\rho' = C \otimes D \rho C^\dagger \otimes D^\dagger = \frac{1}{MN}\left(I^A \otimes I^B + \sum_i^{M^2-1} t_i \lambda_i^A \otimes \lambda_i^B \right), \tag{2.38}$$

with λ_i as in Equation 2.5. Since C and D are invertible, it follows that ρ' is separable if and only if ρ is separable, and so QUSEP for full rank ρ is reduced to QUSEP for ρ' with maximally mixed subsystems, as desired. Of course, this prompts the question — is QUSEP still NP-hard when restricted to full rank ρ?

Finally, we remark that although we have assumed here that the input map Φ is specified in Jamiołkowski form, it is typically straightforward to move to another representation, such as the operator sum or Kraus representation [46]. Briefly, the operator sum representation for a map Φ is given as [6] a set of operators K_i acting on the Hilbert space of the input quantum state ρ, such that $\Phi(\rho) = \sum_{i=1}^{m} K_i \rho K_i^\dagger$, with $\sum_i^m K_i^\dagger K_i \leq I$. Then, the canonical[6] way to determine a Kraus representation for Φ given the Jamiołkowski form ρ_Φ is to reshuffle the eigenvectors of ρ_Φ into matrices K_i. We refer the reader to [6] for further details.

2.5 A Brief Conclusion and Open Problems

We have seen that the problem of Weak Membership over the set of separable quantum states $\mathcal{S}_{M,N}$ is strongly NP-hard, meaning it is NP-hard even if the error parameter $\beta \leq 1/\operatorname{poly}(M,N)$ (i.e. if β is "moderately" large). As seen in Section 2.4, this hardness result immediately also translates into NP-hardness of Weak Membership over the set of entanglement-breaking maps.

Admittedly, our value of $\beta \in \Omega(M^{-16}N^{-20.5})$ in Theorem 10 is a rather modest one. It would be interesting if this value could be improved, perhaps even to $\beta \in O(1)$. As discussed in Section 2.4, such a result would have immediate implications to the study of bound entanglement. One approach, as suggested by Liu [48], may include replacing the algorithm for the reduction $\text{WOPT}_\epsilon(\mathcal{S}_{M,N}) \leq_T \text{WMEM}_\beta(\mathcal{S}_{M,N})$ in Equation 2.6 used here with an algorithm based on random walks (although this alone will not suffice, since an inverse polynomial dependence on the dimension is already introduced in the first reduction of Lemma 2). Specifically, there is a recent iterative algorithm of Bertsimas and Vempala [11] which solves convex problems using random walks, given an initial point in the convex set of interest K. At a high level, if we denote by K' the portion of K which the algorithm believes contains the optimum solution after its most recent iteration, then given a (strong) membership oracle for K, the algorithm proceeds in each iteration by sampling random points from K' via random walk, computing an approximate centroid for K', and subsequently returning an optimal solution, or ruling out a constant fraction of the volume of K' for the next iteration. Such an approach may prove advantageous when the given membership oracle is weak for the following reasons. Since the centroid is estimated using random samples of K', it is plausible that, for an appropriately chosen random sampling technique, the algorithm is unlikely in each iteration to choose many points from the error region surrounding K (especially if the error region is "sufficiently small"). Further, it is plausible that a small number of sampling errors per iteration would not drastically impact the estimate of the centroid in that iteration. It is hence possible that an in-depth analysis of the robustness of the Bertsimas-Vempala algorithm when given a weak membership oracle may provide better parameter scaling for β than achieved here. We remark, however, that this alone would not improve our estimate to $\beta \in O(1)$

[6] We use the term canonical here, as unlike the Jamiołkowski representation, the operator sum representation is not unique.

using the reduction chain of Equation 2.6, as an inverse dependence on the dimension is already introduced in reducing CLIQUE to RSDF (i.e. the first link of the chain).

Chapter 3

The Global Power of Locally Invariant Unitary Operations

3.1 Overview

Having discussed NP-hardness results for the quantum separability problem in Chapter 2, we now switch roles and investigate one attempt at solving the problem, based on what we call locally invariant unitary operations. Specifically, let ρ be a bipartite quantum state in Hilbert space $\mathcal{H}^M \otimes \mathcal{H}^N$ with subsystems denoted by A and B, such that $\rho_B = \mathrm{Tr}_A(\rho)$. Then, we define a *locally invariant* or *cyclic* unitary operation U^B acting on B as one with the property $U^B \rho_B U^{B^\dagger} = \rho_B$. The study of such unitaries within the context of entanglement detection was first proposed by Fu in 2006 [25], who posed the question—for given ρ, cyclic unitary U^B, and for I^A denoting the identity acting on A, what is

$$\mathrm{d}(\rho, U^B) := \frac{1}{\sqrt{2}} \left\| \rho - (I^A \otimes U^B)\rho(I^A \otimes U^B)^\dagger \right\|_{\mathrm{F}}, \tag{3.1}$$

and can it be used to detect entanglement (where $\| \; \|_{\mathrm{F}}$ denotes the Frobenius norm for matrices)? Note that the principle underlying this question has long been implicitly harnessed, for example, in superdense coding [8], where applying a Pauli operator to half of a Bell state gives rise to an orthogonal Bell state.

In this chapter we give a partial answer to Fu's question by considering the quantity $\mathrm{d}_{\max}(\rho) := \max_{\text{cyclic } U^B} \mathrm{d}(\rho, U^B)$. Specifically, we first derive a closed formula for $\mathrm{d}_{\max}(\rho)$ for the case of (pseudo)pure bipartite quantum states ρ of the form $\rho = \epsilon\sigma + \frac{1-\epsilon}{d} I$, for σ a pure quantum state of dimension d, and $0 \leq \epsilon \leq 1$. We place parentheses around the term *pseudo* here to emphasize the inclusion of pure states for $\epsilon = 1$. We next give a closed formula for the case of Werner states [63] of arbitrary dimension. Third, we demonstrate an equivalence between violation of the CHSH inequality [16] and $\mathrm{d}_{\max}(\rho)$ for certain classes of two-qubit states, as well as show that $\mathrm{d}_{\max}(\rho)$ is generally a weaker entanglement detection criterion than the CHSH inequality, both accomplished by deriving a closed formula for $\mathrm{d}_{\max}(\rho)$ for the two-qubit case. Finally, we investigate $\mathrm{d}_{\max}(\rho)$ for three distinct constructions of two-qutrit bound entangled states. Let us now elaborate on these points briefly.

It turns out that $\mathrm{d}_{\max}(\rho)$ cannot be used to define an entanglement measure (with respect to the set of criteria generally required of an entanglement measure, as outlined for example in [6]), nor is it quite a non-locality measure. Our closed formula for (pseudo)pure states immediately reveals the existence of non-maximally entangled (NME) states which are capable of producing in higher dimensions as large a shift as maximally entangled (ME) states, hence ruling out the former possiblity. One may notice, however, that this anomaly is similar to a growing list of arguments claiming non-locality and entanglement as distinct resources [21, 1, 12, 50]—for generalized Bell inequalities for qutrits, or the ability to close the detection loophole, for example, NME states *outperform* ME states. One may then be tempted to claim $\mathrm{d}_{\max}(\rho)$ is the first example of a non-locality measure in which NME states perform as well as, but not better than ME ones (as suggested and subsequently dismissed by Méthot [49]). Unfortunately, it is known that classically correlated states ρ_{cc}, which are separable by definition and hence local, can achieve $\mathrm{d}(\rho_{\mathrm{cc}}, U^B) > 0$ [25]. Here, we define a classically correlated state [63] $\rho_{\mathrm{cc}} \in \mathcal{H}^M \otimes \mathcal{H}^N$ as $\rho_{\mathrm{cc}} = \sum_{k=1}^{n} p_k |a_k\rangle\langle a_k| \otimes |b_k\rangle\langle b_k|$, for $p_k \in \mathbb{R}^+$, $\sum_{k=1}^{n} p_k = 1$, and $n > 1$. Thus, $\mathrm{d}_{\max}(\rho)$ is not a non-locality measure either, disallowing such a claim. For this reason, we also dub the effects observed by $\mathrm{d}(\rho, U^B)$ "global" effects, as opposed to "non-local", the latter being the term originally used by Fu [25].

In light of the fact that $\mathrm{d}_{\max}(\rho_{\mathrm{cc}}) > 0$ is possible for classically correlated states ρ_{cc}, the application proposed by Fu was hence to detect entanglement by distinguishing between classical and quantum correlations. Specifically, following an argument of Fu [25], we have $\mathrm{d}_{\max}(\rho_{\mathrm{cc}}) \leq \sqrt{2(M-1)(N-1)/(MN)}$. Unfortunately, this bound is meaningless for $\max(M, N) \geq 4$, but gives a tight bound of $\mathrm{d}_{\max}(\rho_{\mathrm{cc}}) \leq 1/\sqrt{2}$ for $M = N = 2$, achieved, for example, by the state $\rho_{cc} = \frac{1}{2}|00\rangle\langle 00| + \frac{1}{2}|11\rangle\langle 11|$ and $U^B = \sigma_x$, the Pauli X operator. For $M = N = 3$, this gives a (likely loose) bound of $\mathrm{d}_{\max}(\rho_{\mathrm{cc}}) \leq \sqrt{8/9}$. Determining a tight upper bound for arbitrary dimensions remains at present an intriguing open problem. Here, we derive a closed formula for $\mathrm{d}_{\max}(\rho)$ for Werner states ρ_w of arbitrary dimension, which shows that at least for Werner states, as the dimension grows, the ability to distinguish between classical and quantum correlations using $\mathrm{d}_{\max}(\rho)$ becomes negligible.

We next demonstrate for certain classes of two-qubit states an equivalence between violating the CHSH inequality [16] and obtaining $\mathrm{d}_{\max}(\rho) > 1/\sqrt{2}$, with $1/\sqrt{2}$ the bound on $\mathrm{d}_{\max}(\rho_{\mathrm{cc}})$ for classically-correlated two-qubit states stated above. For the general case of two-qubit states, we show that only one direction of this equivalence holds, implying $\mathrm{d}_{\max}(\rho)$ is a weaker entanglement detection criterion than the CHSH inequality. These results are achieved by first deriving a closed formula for $\mathrm{d}_{\max}(\rho)$ for arbitrary two-qubit states.

Finally, we investigate $\mathrm{d}_{\max}(\rho)$ for three distinct constructions of bound entangled two-qutrit states ρ, in an attempt to discern whether one can detect bound entanglement using $\mathrm{d}_{\max}(\rho)$. In all three cases, we find conclusive or suggestive evidence that $\mathrm{d}_{\max}(\rho)$ cannot surpass the bound of $\mathrm{d}_{\max}(\rho_{\mathrm{cc}}) \leq \sqrt{8/9}$ for classically correlated two-qutrit states discussed above. Thus, with no better bound on $\mathrm{d}_{\max}(\rho_{\mathrm{cc}})$ currently availble, we cannot conclude that $\mathrm{d}_{\max}(\rho)$ can be used to detect bound entanglement.

Throughout this chapter, we will often use the term Fu shift interchangeably with $\mathrm{d}(\rho, U^B)$. We proceed as follows. In Section 3.2, we briefly define the Fu shift formally, and discuss its properties. Section 3.3.1 presents our closed formula for $\mathrm{d}_{\max}(\rho)$ for pseudopure quantum states. Section 3.3.2 gives our formula for $\mathrm{d}_{\max}(\rho)$ for Werner states of arbitrary dimension. In Section 3.4, we investigate connections between the CHSH inequality and $\mathrm{d}_{\max}(\rho)$ for two-qubit systems. Section 3.6 briefly discusses failed attempts in considering variants of $\mathrm{d}(\rho, U^B)$ which constitute true non-locality measures (i.e. $\mathrm{d}(\rho_{\mathrm{cc}}, U^B) = 0$ for all classically correlated ρ_{cc}). Finally, we conclude in Section 3.7 and pose open questions. This chapter is based on joint work with Dagmar Bruß and Hermann Kampermann of the Heinrich-Heine-Universität Düsseldorf.

3.2 The Fu Shift

Throughout our discussion, we denote by ρ a bipartite quantum state living in Hilbert space $\mathcal{H}^M \otimes \mathcal{H}^N$, with subsystems $\rho_A = \mathrm{Tr}_B(\rho)$ and $\rho_B = \mathrm{Tr}_A(\rho)$ of dimensions M and N, respectively. We denote by $\|\mathbf{x}\|_2$ the Euclidean norm for a vector $\mathbf{x}$, and by $\|A\|_{\mathrm{F}} := \sqrt{A^\dagger A}$ the Frobenius norm for a matrix A.

Let us now discuss the definition of the Fu shift from Equation 3.1 in detail. Consider a locally invariant, or cyclic, unitary operation U^B, such that one has $U^B \rho_B U^{B\dagger} = \rho_B$. This is equivalent to the condition:

$$[\rho_B, U^B] = 0. \tag{3.2}$$

To see this, note that $U^B \rho_B = \rho_B U^B \Leftrightarrow U^B \rho_B U^{B\dagger} = \rho_B$, by the unitarity of U^B. For convenience, define $\rho_f = (I \otimes U^B)\rho(I \otimes U^{B\dagger})$. Then, as per Equation 3.1, the Fu shift $\mathrm{d}(\rho, U^B)$ is defined as the (scaled) Euclidean distance between ρ and ρ_f [25], such that $\mathrm{d}(\rho, U^B) = \frac{1}{\sqrt{2}} \|\rho - \rho_f\|_{\mathrm{F}}$. This can straightforwardly be rewritten in the following useful form [25]:

$$d(\rho, U^B) = \sqrt{\mathrm{Tr}(\rho^2) - \mathrm{Tr}(\rho\rho_f)}. \tag{3.3}$$

As stated earlier, we denote by $\mathrm{d}_{\max}(\rho)$ the maximum Fu shift attainable for *any* choice of cyclic U^B on state ρ, i.e.

$$\mathrm{d}_{\max}(\rho) := \max_{\text{cyclic } U^B} \mathrm{d}(\rho, U^B). \tag{3.4}$$

It is easy to see from Equation 3.3 that $\mathrm{d}(\rho, U^B)$ has range $0 \leq \mathrm{d}(\rho, U^B) \leq 1$, with the latter inequality saturated if and only if ρ is pure and orthogonal to ρ_f. It is not known whether $\mathrm{d}_{\max}(\rho) > 0$ for all entangled ρ, though our results will show that for (pseudo)pure states and Werner states, this is indeed the case. It is easy to show that for any product state ρ, $\mathrm{d}_{\max}(\rho) = 0$ [25].

We close this section by remarking that one could instead consider defining $\mathrm{d}_{\min}(\rho) := \min_{U^B} \mathrm{d}(\rho, U^B)$, as opposed to $\mathrm{d}_{\max}(\rho)$. We investigate this notion briefly in Section 3.6 in the (dashed) hope of eliminating the issue of classically correlated ρ_{cc} achieving $\mathrm{d}_{\max}(\rho_{\mathrm{cc}}) > 0$.

3.3 Maximizing the Fu Shift

3.3.1 Pseudopure States

We now derive a closed formula for $\mathrm{d}_{\max}(\rho)$ for pseudopure states, and follow with a discussion of its implications. Specifically, consider bipartite pseudopure state ρ of the form

$$\rho = \epsilon\sigma + \frac{1-\epsilon}{MN} I, \tag{3.5}$$

with σ a pure bipartite quantum state of dimension MN, and $0 \leq \epsilon \leq 1$. For simplicity of exposition, we assume $M \geq N$, although the results shown straightforwardly hold for arbitrary M and N. By the Schmidt decomposition [53], we can assume without loss of generality that $\sigma = |\psi\rangle\langle\psi|$, where $|\psi\rangle$ is of the form $|\psi\rangle = \sum_{k=0}^{N-1} a_k |kk\rangle$, with $\sum_{k=0}^{N-1} |a_k|^2 = 1$, and where $\{|k\rangle\}_{k=0}^{N-1}$ are elements of the respective Schmidt bases for each subsystem[1].

To begin, we first need the following two lemmas, the first of which follows from straightforward verification via the cosine law for triangles, and whose proof we omit.

Lemma 12. *Let a, b, and c be the lengths of sides of a triangle, with θ_b and θ_c the angles opposite sides b and c, respectively. Then, setting $\gamma_b = \pi - \theta_c$ and $\gamma_c = \pi - \theta_b$ gives $a + be^{i\gamma_b} + ce^{i\gamma_c} = 0$.*

Lemma 13. *Let ρ be a pseudopure quantum state with $\epsilon > 0$. Then, for any $0 \leq k \leq N-1$ such that $|a_k| \neq |a_j|\ \forall\ j \neq k$ (i.e. $|a_k|$ is an amplitude of unique absolute value) and unitary U^B,*

$$[\rho_B, U^B] = 0 \quad \Rightarrow \quad \left|U^B_{k,k}\right| = 1. \tag{3.6}$$

[1] For notational simplicity, we denote elements of the Schmidt bases for both subsystems by $\{|k\rangle\}_{k=0}^{N-1}$, although both bases can in general be different (which does not affect our analysis). We remark that by Lemma 16 seen later, one can in fact take both Schmidt bases here to coincide if desired.

Proof. Let U^B be an arbitrary $N \times N$ matrix, given by $U^B = \sum_{m,n=0}^{N-1} \langle m|U^B|n\rangle |m\rangle\langle n|$, for $\{|m\rangle\}_{m=0}^{N-1}$ the Schmidt basis for subsystem B above. Observe that:

$$[\rho_B, U^B] = \epsilon \sum_{m,n=0}^{N-1} (|a_m|^2 - |a_n|^2) \langle m|U^B|n\rangle |m\rangle\langle n|. \tag{3.7}$$

Hence, if two amplitudes of $|\psi\rangle$ do not agree in absolute value, the corresponding entry in U^B must be 0 in order for ρ_B and U^B to commute. If $|a_k|$ is distinct then, it follows that row k and column k of U^B must be all zeroes, except for position ${U^B}_{k,k}$. Since U^B is unitary, we thus have $\left|{U^B}_{k,k}\right| = 1$. □

We now show the main result of this section. For the remainder of our discussion, let $a_m = \max_i |a_i|$, for a_i the amplitudes of $|\psi\rangle$ from the pseudopure state ρ of Equation 3.5.

Theorem 14. *Let $\rho \in \mathcal{H}^M \otimes \mathcal{H}^N$ be a pseudopure quantum state. Then,*

$$\mathrm{d}_{\max}(\rho) = \begin{cases} \epsilon & \text{if } |a_m|^2 \leq \sum_{k \neq m} |a_k|^2, \\ 2\epsilon |a_m| \sqrt{1 - |a_m|^2} & \text{otherwise,} \end{cases} \tag{3.8}$$

where for $\epsilon = 0$, we define $a_k = 0$ for $0 \leq k \leq N-1$.

Proof. Observe first that substituting ρ into Equation 3.3 gives for arbitrary U^B (not necessarily cyclic):

$$\mathrm{d}(\rho, U^B) = \epsilon \sqrt{1 - \left| \sum_{k=0}^{N-1} |a_k|^2 \langle k|U^B|k\rangle \right|^2} \tag{3.9}$$

Thus, $\mathrm{d}(\rho, U^B)$ depends only on the diagonal entries of U^B. We hence first consider the case of diagonal U^B, and subsequently show that choosing U^B diagonal is in fact always optimal. We remark that for $\epsilon = 0$, we clearly have $\mathrm{d}_{\max}(\rho) = 0$, agreeing with Equation 3.8, and so we henceforth assume $\epsilon > 0$.

Assume then that U^B is a diagonal unitary matrix with eigenvalue $e^{i\theta_k}$ on row k. Then by Equation 3.9, maximizing $d(\rho, U^B)$ reduces to minimizing $\left| \sum_{k=0}^{N-1} |a_k|^2 e^{i\theta_k} \right|$. Since ρ_B is diagonal, any choice of θ_k's constitutes a commuting unitary operation U^B, and so this minimization problem has a simple geometric solution as follows.

If, for a_m the amplitude of largest absolute value, we have $|a_m|^2 \leq \sum_{k \neq m} |a_k|^2$, then for $N \geq 3$, we can always construct a triangle using each of the lengths $|a_k|^2$ exactly once, due to the well known fact that three positive numbers a, b, and c constitute the lengths of sides of a triangle if and only if $a \leq b + c$, $b \leq a + c$, and $c \leq a + b$ [51]. Specifically, one can sort the $|a_k|^2$ in decreasing order (such that

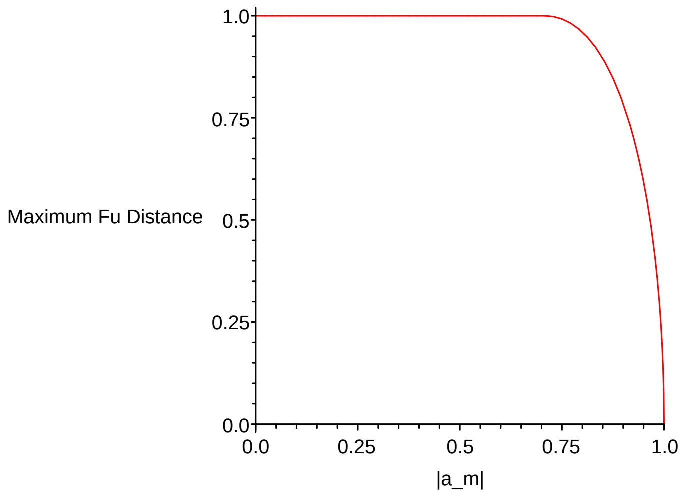

Figure 3.1: A plot of $\mathrm{d}_{\max}(\rho)$ for pure states as a function of $|a_m|$, as given by Equation 3.8.

$a_0 = a_m$), and set the longest side of the triangle to be of length $\sum_{k=0}^{i} |a_k|^2$ for $1 \leq i \leq N-3$, such that

$$\sum_{k=0}^{i} |a_k|^2 \leq \sum_{k=i+1}^{N-1} |a_k|^2 \qquad \text{and} \qquad \sum_{k=0}^{i+1} |a_k| > \sum_{k=i+2}^{N-1} |a_k|. \tag{3.10}$$

Then, applying Lemma 12 gives $\mathrm{d}_{\max}(\rho) = \epsilon$. Note that if $N = 2$, setting $\theta_0 = 0$ and $\theta_1 = \pi$ achieves the same result.

If, on the other hand, we have $|a_m|^2 > \sum_{k \neq m} |a_k|^2$, then clearly no triangle can be constructed, since the length of the longest side cannot be matched by any combination of shorter sides into a single line segment. Thus, the best minimization strategy is simply to set $\theta_m = 0$ and $\theta_k = \pi$, for all $k \neq m$. Then, substituting into our simplified expression for Equation 3.9 and repeated use of the identity $\sum_{k=0}^{N-1} |a_k|^2 = 1$ gives $\mathrm{d}_{\max}(\rho) = 2\epsilon |a_m| \sqrt{1 - |a_m|^2}$.

Finally, to see that choosing U^B diagonal is always optimal, note first that if $|a_m|^2 \leq \sum_{k \neq m} |a_k|^2$, then by our argument above, choosing diagonal U^B always gives the maximum possible value for Equation 3.9. If $|a_m|^2 > \sum_{k \neq m} |a_k|^2$, on the other hand, clearly $|a_m| \neq |a_k|$ for all $k \neq m$, and so it follows from Lemma 13 that $\left|\langle m|U^B|m\rangle\right| = 1$ in Equation 3.9. Thus, the best minimization strategy is to set $\theta_m = 0$, and for all $k \neq m$, set $\left|\langle k|U^B|k\rangle\right| = 1$ (i.e. make U^B diagonal) and $\theta_k = \pi$, as we did for the second case above. Recalling that diagonal U^B will always commute with ρ_B for any ρ of the given form concludes the proof. □

We have plotted Equation 3.8 for pure states in Figure 3.1 to highlight the interesting point Theorem 14 makes—namely, aside from ϵ, the strength of global effect producible for pseudopure ρ is dependent only on the relation between $|a_m|$ and the remaining amplitudes. Specifically, if we let $\epsilon = 1$, it gives the exact point at which ρ fails to achieve a maximal Fu shift of $\mathrm{d}_{\max}(\rho) = 1$ (namely, when $|a_m| > \frac{1}{\sqrt{2}}$), immediately implying that ρ need not be maximally entangled in order to achieve $\mathrm{d}_{\max}(\rho) = 1$. As mentioned in Section 3.1, this is similar to an intriguing phenomenon seen also in the context of non-locality measures, such as generalized Bell inequalities in higher dimensions [1]. Due to this property, the Fu shift cannot be used to define an entanglement measure. From Theorem 14, we remark that it is also clear that $\mathrm{d}_{\max}(\rho) > 0$ for all entangled pseudopure ρ.

We further remark that for the two-qubit pure state case, it is known that $\mathrm{d}_{\max}(\rho) = C(\rho) = 2|a_0a_1|$ [25], where $C(\rho)$ denotes the *concurrence* of ρ [33], to which Equation 3.8 here also reduces. Used by Wootters [65] to derive an analytic formula for the entanglement of formation for arbitrary two-qubit states, and an entanglement measure in its own right, the concurrence has since been generalized to higher dimensional systems using various approaches [59, 56, 3, 66, 4]. Thus, we ask, does Equation 3.8 also reduce to any of these generalized notions of concurrence for pure states in higher dimensions? Two of the generalizations have simple closed forms for the case of pure states, which we analyze here.

Let us assume as before that ρ is of the form given by Equation 3.5 with $\epsilon = 1$, and let $d = \min\{M, N\}$. Then, Rungta et al. [56] define the concurrence $C_R(\rho)$ for pure states ρ as $C_R(\rho) = \sqrt{1 - \mathrm{Tr}(\rho_A^2)}$, where $0 \leq C_R(\rho) \leq \sqrt{2(d-1)/d}$. We can rewrite this expression as

$$C_R(\rho) = \sqrt{2\left(1 - \sum_{k=0}^{d-1} |a_k|^4\right)}, \tag{3.11}$$

from which it is clear that our expression for $\mathrm{d}_{\max}(\rho)$ of Equation 3.8 does not reduce to (a normalized version of) $C_R(\rho)$. For example, take $|\psi\rangle = \frac{1}{\sqrt{2}}|00\rangle + \frac{1}{2}|11\rangle + \frac{1}{2}|22\rangle$, for which $\mathrm{d}_{\max}(\rho) = 1$, but $C_R(\rho) \approx 0.9682$, where we have normalized the latter by the maximum value possible for qutrits, $2/\sqrt{3}$. Next, consider the generalization of Audenaert et al. [3], which states that for pure ρ, we have $C_A(\rho) = 2s_1s_2$, where s_1 and s_2 are the first and second largest Schmidt coefficients in the Schmidt decomposition of $|\psi\rangle$. Again, it is clear that our expression from Equation 3.8 for $\mathrm{d}_{\max}(\rho)$ does not reduce to this definition $C_A(\rho)$ either. As a counterexample, consider the maximally entangled two qutrit state $|\psi\rangle = \frac{1}{\sqrt{3}}(|00\rangle + |11\rangle + |22\rangle)$, for which $\mathrm{d}_{\max}(\rho) = 1$, but $C_A(\rho) = 2/3$. Thus, $\mathrm{d}_{\max}(\rho)$ does not reduce to either of these two generalized notions of concurrence in higher dimensions.

3.3.2 Werner States

We now turn our attention to bipartite Werner states ρ of arbitrary dimension, for which we derive a closed formula for $\mathrm{d}_{\max}(\rho)$. Letting $\{|j\rangle\}_{j=0}^{d-1}$ denote an arbitrary orthonormal basis for $\mathbb{C}^d$, where $d \geq 2$, the Werner state $\rho \in \mathcal{H}^d \otimes \mathcal{H}^d$ can be defined using the following terms [63]:

$$\begin{aligned} P &= \sum_{j,k} |j\rangle\langle k| \otimes |k\rangle\langle j| & (3.12) \\ P_{sym} &= \frac{1}{2}(I_{d^2} + P) & (3.13) \\ P_{as} &= \frac{1}{2}(I_{d^2} - P) & (3.14) \\ \rho &= p\frac{2}{d^2+d}P_{sym} + (1-p)\frac{2}{d^2-d}P_{as}, & (3.15) \end{aligned}$$

where I_{d^2} is the d^2-dimensional identity matrix, $0 \leq p \leq 1$, and ρ is a Werner state. This state has the following interesting properties. First, it is invariant under operation $U \otimes U$, for any choice of unitary U. Second, it is entangled for $p < 1/2$, and separable otherwise. Investigating in terms of Fu shift, we find the following result.

Theorem 15. *Let $\rho \in \mathcal{H}^d \otimes \mathcal{H}^d$ be a Werner state. Then we have*

$$\mathrm{d}_{\max}(\rho) = \frac{|2pd - d - 1|}{d^2 - 1}, \tag{3.16}$$

obtained using any traceless $d \times d$ choice of unitary U^B.

Proof. We begin by setting up Equation 3.3 using the given definition of ρ and arbitrary U^B. Observing that $\mathrm{Tr}(P) = d$, $\mathrm{Tr}(P^2) = d^2$, and defining for convenience $\beta := \mathrm{Tr}(P(I \otimes U^B)P(I \otimes U^{B^\dagger})) = \mathrm{Tr}(U^B)\mathrm{Tr}(U^{B^\dagger})$, straightforward manipulation leads us to

$$d(\rho, U^B) = \frac{\sqrt{(2pd - d - 1)^2(d^2 - \beta)}}{d(d^2 - 1)}. \tag{3.17}$$

Examining the boundary and critical points of the first derivative of Equation 3.17, we find the two cases of interest are $\beta = -d$ and $\beta = 0$. Note, however, that $\beta = -d$ implies $\mathrm{Tr}(U^B)\mathrm{Tr}(U^{B^\dagger}) = -d$, which is impossible, since $aa^* \geq 0$ for all $a \in \mathbb{C}$. Hence, the maximum Fu value is achieved when $\beta = 0$, implying that U^B is traceless, giving the desired result. □

We now direct the reader's attention to Figure 3.2, which graphically depicts $\mathrm{d}_{\max}(\rho)$ for various dimensions, d. For $d = 2$, recall from Section 3.1 that for all classically correlated states ρ_{cc}, we have $\mathrm{d}_{\max}(\rho_{\mathrm{cc}}) \leq 1/\sqrt{2}$, and hence there are certain values of p for which the Fu shift can be used to reliably detect entanglement in Werner states. For arbitrary d, if one is promised that ρ is a Werner state, but not given the value of p, then by Figure 3.2, attaining $\mathrm{d}_{\max}(\rho) > 1/(d + 1)$ is sufficient to conclude ρ is entangled (seen by setting $p = 1$).

We remark that examining the critical points of the first derivative of Equation 3.16 as $d \to \infty$ suggests there is a "kink" in the graph at $p = 1/2$. This is

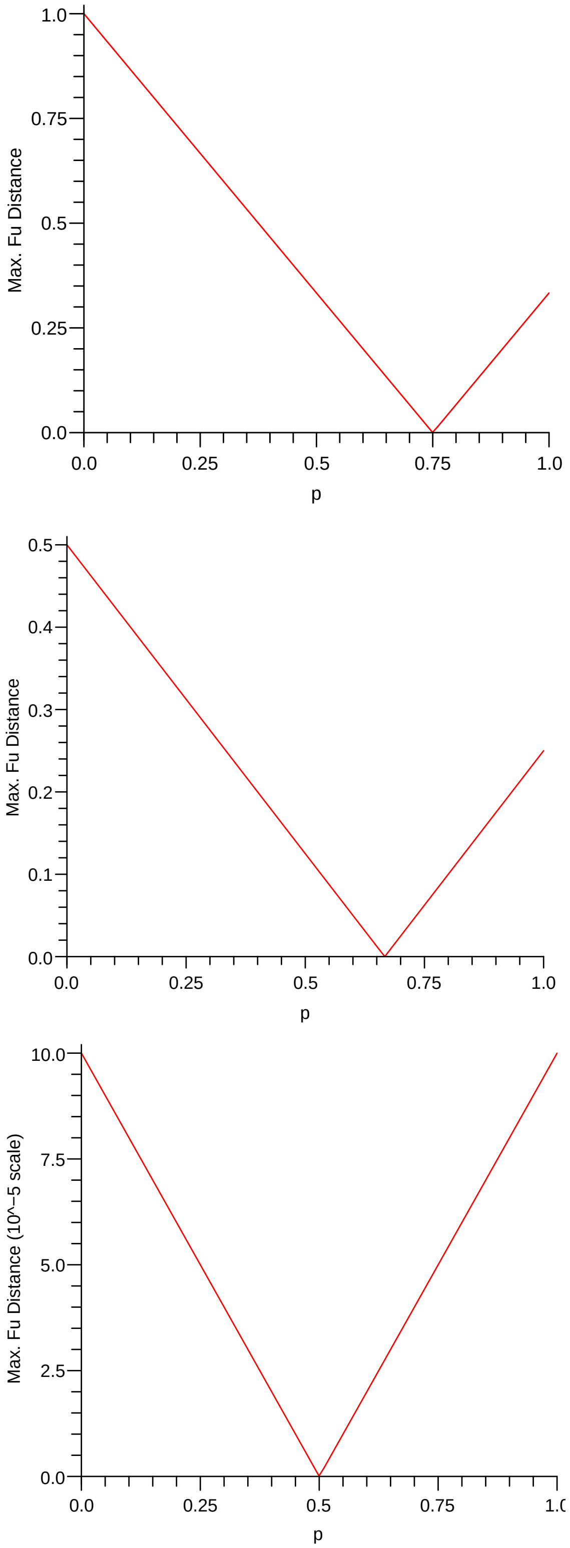

Figure 3.2: A plot of $\mathrm{d}_{\max}(\rho)$ for Werner states (as defined by Equation 3.12) of subsystem dimensions $d = 2, 3, 10000$, respectively, as given by Equation 3.16. Note the scale of 10^{-5} for the vertical axis for the case of $d = 10000$.

interesting, since $p = 1/2$ is exactly the boundary between entangled and separable Werner states. It also suggests that as $d \to \infty$, the Fu shifts attainable by quantum and classical correlations for Werner states tend towards becoming identical, and coupled with the observation that $\mathrm{d}_{\max}(\rho) \to 0$ as $d \to \infty$, that the possibility of distinguishing between the two if d is large enough vanishes. We close this section by noting that Equation 3.16 gives $\mathrm{d}_{\max}(\rho) > 0$ for all entangled Werner states.

3.4 Connections to the CHSH Inequality

In Section 3.3.1, we derived a closed formula for $\mathrm{d}_{\max}(\rho)$ for any pseudopure state ρ. Consider then the (pseudopure) two-qubit Werner state $\rho = p|\phi^+\rangle\langle\phi^+| + \frac{1-p}{4}I$, where $|\phi^+\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$ [13]. Applying our formula to ρ immediately gives us that since $\mathrm{d}_{\max}(\rho_{\mathrm{cc}}) \leq 1/\sqrt{2}$ for two-qubit classically correlated states ρ_{cc}, we require $p > 1/\sqrt{2}$ in order to reliably detect entanglement in ρ using the Fu shift. Strangely, this is the exact same bound we require on p in order to violate the original CHSH inequality [16, 41]. Hence, we ask, is there a connection between violation of the CHSH inequality and the ability to reliably detect entanglement using the Fu shift in the two-qubit case?

In order to answer this question, we need three preliminary building blocks. The first is the fact that any bipartite state ρ can be written in what is known as the Fano form [24]:

$$\rho = \frac{1}{MN}\left(I^A \otimes I^B + \vec{r}^A \cdot \vec{\sigma}^A \otimes I^B + I^A \otimes \vec{r}^B \cdot \vec{\sigma}^B + \sum_{i=0}^{M^2-1}\sum_{j=0}^{N^2-1} T_{ij}\sigma_i^A \otimes \sigma_j^B\right), \tag{3.18}$$

where M denotes the dimension of subsystem A, I denotes the identity matrix, $\vec{r}^A$ denotes the (M^2-1)-dimensional Bloch vector for subsystem A such that entry $r_i^A = \frac{M}{2}\mathrm{Tr}(\sigma_i^A \rho_A)$, $\vec{\sigma}^A$ denotes the (M^2-1)-component vector of traceless Hermitian generators for $SU(M)$, and the matrix T is a real matrix known as the *correlation matrix*, whose entries are given by $T_{ij} = \frac{MN}{4}\mathrm{Tr}(\sigma_i \otimes \sigma_j \rho)$. The definitions for subsystem B are analogous. We stress that the form above is normalized differently than that used in [25]. For qubits, however, the two expressions coincide.

With the Fano form in hand, we can express our second building block, a known necessary and sufficient condition for violation of the CHSH inequality [41]. For a given input state ρ, define

$$M(\rho) := \lambda_1(T^T T) + \lambda_2(T^T T), \tag{3.19}$$

for T the correlation matrix from the Fano form, T^T its transpose, and $\lambda_1(T^T T)$ and $\lambda_2(T^T T)$ the first and second largest eigenvalues of $T^T T$, respectively. Then, ρ can violate the CHSH inequality if and only if

$$M(\rho) > 1. \tag{3.20}$$

Let us demonstrate the use of this condition on the two-qubit Werner state discussed earlier, $\rho = p|\phi^+\rangle\langle\phi^+| + \frac{1-p}{4}I$. The correlation matrix, T, for ρ is just a diagonal matrix with entries p, $-p$, and p appearing on its diagonal, respectively. Thus, $T^T T$ is also diagonal, and has eigenvalue p^2 with multiplicity 3. By Equation 3.20, ρ hence violates the CHSH inequality if and only if $M(\rho) = 2p^2 > 1$, or when $p > 1/\sqrt{2}$, as stated earlier.

Finally, in order to tie violation of the CHSH inequality to reliable entanglement detection using the Fu shift, we note that the value of the Fu shift can be rewritten in terms of the correlation matrix T as follows [25]:

$$d(\rho, U^B) = \frac{2}{MN}\sqrt{\sum_{i,j} T_{ij}^2 - \sum_{i,j} T_{ij}T_{ij}^f} = \frac{1}{2}\sqrt{\sum_{i,j} T_{ij}^2 - \sum_{i,j} T_{ij}T_{ij}^f}, \tag{3.21}$$

where T_{ij}^f is the entry at position (i, j) in the final version of T after our unitary U^B has been applied to the second subsystem of our state. We can now proceed to investigate how the CHSH inequality and $\mathrm{d}_{\max}(\rho)$ are related.

3.4.1 Diagonal T and Connections Between CHSH and Fu Shift

We now investigate connections between violation of the CHSH inequality and the ability to reliably detect entanglement using the Fu shift. Formally, we would like to consider the plausibility of the statement:

$$M(\rho) > 1 \qquad \Longleftrightarrow \qquad \mathrm{d}_{\max}(\rho) > \frac{1}{\sqrt{2}}, \tag{3.22}$$

with $M(\rho)$ defined as in Equation 3.19. In order to do so, we first show that, without loss of generality, one can take the correlation matrix T of Equation 3.18 to be diagonal. We then derive a closed formula for $\mathrm{d}_{\max}(\rho)$ for any two-qubit state ρ with diagonal T. Finally, using this formula, we show our desired results by comparing $\mathrm{d}_{\max}(\rho)$ and $M(\rho)$.

First, to show one can take T to be diagonal, we require the following lemma.

Lemma 16. *For bipartite state ρ of arbitrary dimension, $\mathrm{d}_{\max}(\rho)$ is invariant under local unitary operations.*

Proof. Let U_1 and U_2 be local unitary operations applied to subsystems A and B of ρ, respectively. Then, straightforward manipulation of Equation 3.3 yields that $\mathrm{d}(U_1 \otimes U_2 \rho U_1^\dagger \otimes U_2^\dagger, U^B) = \mathrm{d}(\rho, U_2^\dagger U^B U_2)$. It is easy to see that $[U_2 \rho_B U_2^\dagger, U^B] = 0$ if and only if $[\rho_B, U_2^\dagger U^B U_2] = 0$. Thus, since the set of unitary matrices (of fixed dimension) forms a group, cycling through all possible choices of U^B in $\mathrm{d}(\rho, U_2^\dagger U^B U_2)$ gives the desired result. □

In order to use Lemma 16 to show that we can assume T is diagonal, we follow [40] and note that applying a unitary operation $U_1 \otimes U_2$ to ρ is the equivalent of applying orthogonal rotation matrices O_1 and O_2 to $\vec{r}^A$, $\vec{r}^B$, and T, such that:

$$\vec{r}^A_f = O_1\vec{r}^A \qquad \vec{r}^B_f = O_2\vec{r}^B \qquad T_f = O_1 T O_2^\dagger \tag{3.23}$$

Thus, given any ρ, we can find[2] some $U_1 \otimes U_2$ such that $\rho' = U_1 \otimes U_2 \rho U_1^\dagger \otimes U_2^\dagger$ has diagonal T [40], which, coupled with Lemma 16, gives us that we can assume without loss of generality that T is diagonal, as desired.

Given any two-qubit state ρ with diagonal T then, we derive a closed expression for $\mathrm{d}_{\max}(\rho)$.

Lemma 17. *Given two-qubit quantum state ρ with diagonal correlation matrix T, let $\lambda_i = T_{ii}$. Then*

$$\mathrm{d}_{\max}(\rho) = \frac{1}{\sqrt{2}}\sqrt{\lambda_0^2(1-n_0^2) + \lambda_1^2(1-n_1^2) + \lambda_2^2(1-n_2^2)}, \tag{3.24}$$

where if $\rho_B \neq I/2$, $\vec{n} = \vec{r}_B / \|\vec{r}_B\|_2$, and otherwise $n_i = 1$ for $\lambda_i = \min_k \lambda_k$ (with $\|\vec{n}\|_2 = 1$).

Proof. Assume for now that $\rho_B \neq I/2$. We manipulate Equation 3.21 to achieve the claimed form. Specifically, in order to characterize T^f in terms of T and U^B, apply Equation 3.23 for $U_1 = I$ and $U_2 = U^B$, and note that an explicit formula for O_2 in terms U^B is given by [26]:

$$\begin{aligned} O_2 &= I + \sin\theta A + (1-\cos\theta)A^2, \text{ where} \\ A &= \begin{pmatrix} 0 & -n_2 & n_1 \\ n_2 & 0 & -n_0 \\ -n_1 & n_0 & 0 \end{pmatrix} \end{aligned} \tag{3.25}$$

Here, $\vec{n} = (n_0, n_1, n_2)$ and θ are the axis and angle of rotation for O_2, respectively, with $\|\vec{n}\|_2 = 1$. Now, observe that demanding $[U^B, \rho_B] = 0$ is equivalent to U^B being a rotation about the Bloch vector of ρ_B. Thus, assuming $\vec{r}_B \neq \vec{0}$ (such that $\rho_B \neq I/2$), any cyclic unitary U^B is given by choosing $\vec{n} = \vec{r}_B / \|\vec{r}_B\|_2$.

In order to next choose a maximizing θ, note that setting $T^f = TO_2^T$, and using Equations 3.23 and 3.25 allows us to simplify Equation 3.21 to:

$$\mathrm{d}(\rho, U^B) = \frac{1}{2}\sqrt{\sum_{i=0}^{2} \lambda_i^2 (1-\cos\theta)(1-n_i^2)} \tag{3.26}$$

Since $\|\vec{n}\|_2 = 1$, we have $(1-n_i^2) \geq 0$, and so this equation is maximized when $\cos\theta = -1$, or $\theta = \pi$.

Finally, if $\rho_B = I/2$, any choice of U^B is cyclic, and hence we can choose any axis of rotation $\vec{n}$ we like. Clearly, choosing $n_i = 1$ for $\lambda_i = \min_k \lambda_k$ is the optimal choice, as claimed. □

[2] This holds in the two qubit case due to the existence of a surjective homomorphism from $SU(2)$ to $SO(3)$.

We now have a closed expression for $\mathrm{d}_{\max}(\rho)$. Let us finally show the main results of this section.

Theorem 18. *Let ρ be a two-qubit quantum state with diagonal correlation matrix T (and definitions as in Lemma 17). Then, $\mathrm{d}_{\max}(\rho) > 1/\sqrt{2} \Rightarrow M(\rho) > 1$, but $M(\rho) > 1 \nRightarrow \mathrm{d}_{\max}(\rho) > 1/\sqrt{2}$.*

Proof. Assume without loss of generality that $|\lambda_0| \geq |\lambda_1| \geq |\lambda_2|$. First, let us assume $\mathrm{d}_{\max}(\rho) > 1/\sqrt{2}$, and let $\vec{n} = (\sqrt{\epsilon_0}, \sqrt{\epsilon_1}, \sqrt{\epsilon_2})$, where $\epsilon_0 + \epsilon_1 + \epsilon_2 = 1$. Then, substitution into Equation 3.24 gives

$$(1-\epsilon_0)\lambda_0^2 + (1-\epsilon_1)\lambda_1^2 + (1-\epsilon_2)\lambda_2^2 > 1, \tag{3.27}$$

from which it is easy to see that $M(\rho) = \lambda_0^2 + \lambda_1^2 > 1$, since setting $\epsilon_2 = 1$ can only increase the left hand side of Equation 3.27.

Conversely, consider $\rho = |\psi\rangle\langle\psi|$, where $|\psi\rangle = a_0|00\rangle + a_1|11\rangle$, for $a_0, a_1 \in \mathbb{R}^+$, and $a_0^2 + a_1^2 = 1$. Then, we have $\vec{n} = (0,0,1)$, $\lambda_0 = 2a_1a_2$, $\lambda_1 = -2a_1a_2$, and $\lambda_2 = 1$. Thus, $M(\rho) > 1$, but $\mathrm{d}_{\max}(\rho) \leq 1/\sqrt{2}$ if $a_0 \leq 0.3827$ or $a_0 \geq 0.9239$. □

Hence, by Theorem 18, if we can reliably detect entanglement in ρ using $\mathrm{d}_{\max}(\rho)$, then we can also violate the CHSH inequality using ρ. The converse of this statement, however, is not necessarily true, implying the Fu shift is generally a weaker entanglement criterion (in the two-qubit case) than the CHSH inequality.

Given this relationship between the CHSH inequality and Fu shift, one can next ask, are there specific classes of states for which for which Equation 3.22 does hold? It turns out that this is indeed the case.

Theorem 19. *Given two-qubit quantum state ρ with diagonal correlation matrix T (and definitions as in Lemma 17), consider the following conditions:*

1. $\lambda_i = \min_k |\lambda_k|$, *and* $|n_i| = 1$, *where* $\vec{n} = \vec{r}_B / \|\vec{r}_B\|_2$, *and* $\vec{r}_B \neq (0,0,0)$.
2. $|\lambda_0| = |\lambda_1| = |\lambda_2|$.
3. $\rho_B = I/2$.

Then, $M(\rho) > 1 \Leftrightarrow \mathrm{d}_{\max}(\rho) > 1/\sqrt{2}$ holds if and only if one of these conditions holds.

Proof. We proceed case-by-case.

1. Suppose without loss of generality that λ_2 is the eigenvalue of T of smallest absolute value, and $\vec{n} = (0,0,1)$. Then, Equation 3.24 simplifies to:

$$\mathrm{d}_{\max}(\rho) = \frac{1}{\sqrt{2}}\sqrt{\lambda_0^2 + \lambda_1^2} = \frac{1}{\sqrt{2}}\sqrt{M(\rho)} \tag{3.28}$$

Therefore, achieving $M(\rho) > 1$ is equivalent to $\mathrm{d}_{\max}(\rho) > 1/\sqrt{2}$, as required.

2. Suppose $|\lambda_0| = |\lambda_1| = |\lambda_2|$. Then, since $\|\vec{n}\|_2 = 1$, Equation 3.24 simplifies to:

$$\mathrm{d}_{\max}(\rho) = \frac{1}{\sqrt{2}}\sqrt{\lambda_0^2(3 - n_1^2 - n_2^2 - n_3^2)} = \frac{1}{\sqrt{2}}\sqrt{\lambda_0^2 + \lambda_1^2} = \frac{1}{\sqrt{2}}\sqrt{M(\rho)}, \tag{3.29}$$

   and we arrive at the same conclusion as in Case 1.

3. Suppose $\rho_B = I/2$. Then by Theorem 17, it straightforwardly follows that we are reduced to to Case 1.

Finally, in order to show the demonstrated equivalence holds if and only if one of these conditions hold, assume without loss of generality that $|\lambda_0| \geq |\lambda_1| \geq |\lambda_2|$. Then, unless $|\lambda_0| = |\lambda_1| = |\lambda_2|$ (Case 2), the only way to guarantee equivalence is to have in Equation 3.24 $(1 - r_0^2) = 1$ and $(1 - r_1^2) = 1$, which implies that $\vec{r}_B = (0, 0, \pm 1)$. But such a choice of $\vec{r}_B$ can only correspond to a cyclic unitary operation if we have Case 1 or 3 above, as required. □

Thus, there exist classes of states for which violation of the CHSH inequality and reliable entanglement detection via $\mathrm{d}_{\max}(\rho)$ are equivalent. Specifically, note that the Werner state ρ we considered in the beginning of Section 3.4 falls into such a class, since we have $\rho_B = I/2$ for ρ. This explains the coincidence observed.

3.5 Bound Entanglement Detection

Let us now investigate the Fu shift for bound entangled states, in an attempt to discern whether it can be used to detect bound entanglement. Bound entangled states are mixed quantum states which are entangled, but from which no pure (state) entanglement can be distilled [36]. Discovered in the context of Peres' Positive Partial Transpose separability criterion [55, 38], it is known that any entangled state with a positive partial transpose is undistillable, or equivalently, bound entangled [36] (the converse of this statement, however, is an intriguing open problem [13]). Due to their undistillability, bound entangled states are not useful on their own for communication purposes, but it turns out that when used as a resource additional to free entanglement (i.e. distillable entanglement), they allow feats unachievable by free entanglement alone (called *activation of bound entanglement*) [39].

Here, we focus on three distinct constructions of bound entangled states of two qutrits. Throughout this section, we denote an unknown matrix $U^B \in \mathbb{C}^{3\times 3}$ (not necessarily unitary) by:

$$U^B = \begin{pmatrix} u_1 & u_2 & u_3 \\ u_4 & u_5 & u_6 \\ u_7 & u_8 & u_9 \end{pmatrix}. \tag{3.30}$$

3.5.1 P. Horodecki Construction

We begin by looking at the two-qutrit bound entangled state construction given by Horodecki [38]. Denote by $P_\psi = |\psi\rangle\langle\psi|$ the projector onto any quantum state $|\psi\rangle$, and the computational basis for qutrits as $\{|0\rangle, |1\rangle, |2\rangle\}$. Then, define, as per [38]:

$$\begin{aligned} Q &= I \otimes I - (\sum_{k=0}^{2} P_k \otimes P_k) - P_2 \otimes P_0 && (3.31) \\ \Psi &= \frac{1}{\sqrt{3}}(|0\rangle \otimes |0\rangle + |1\rangle \otimes |1\rangle + |2\rangle \otimes |2\rangle) && (3.32) \\ \rho_{\text{ent}} &= \frac{3}{8} P_\Psi + \frac{1}{8} Q && (3.33) \\ \Phi_a &= |2\rangle \otimes (\sqrt{\frac{1+a}{2}}|0\rangle + \sqrt{\frac{1-a}{2}}|2\rangle) && (3.34) \\ \rho_a &= \frac{8a}{8a+1}\rho_{\text{ent}} + \frac{1}{8a+1} P_{\Phi_a}, && (3.35) \end{aligned}$$

for $0 \leq a \leq 1$. Note that ρ_{ent} is entangled, as its partial transpose has a negative eigenvalue, and P_{Φ_a} is clearly separable, since Φ_a is a product state. The state of interest, ρ_a, is bound entangled for $0 < a < 1$ [38]. Observe that for $a = 0$, ρ_a reduces to a product state, and hence we know from Section 3.2 that $\mathrm{d}_{\max}(\rho_0) = 0$. Let us now determine $\mathrm{d}(\rho_a, U^B)$ in terms of a.

Theorem 20. *Let ρ_a be defined as above, restricted to domain $0 < a < 1$. Then $\mathrm{d}_{\max}(\rho_a) = \frac{2\sqrt{2}a}{8a+1}$, obtained, for example, using diagonal unitary matrix $U^B \in \mathbb{C}^{3\times3}$ with $U^B_{0,0} = -U^B_{1,1} = U^B_{2,2}$.*

Proof. Let U^B be an arbitrary complex 3×3 matrix, as defined in Equation 3.30. Setting up the commutator relation $[\rho_B, U^B]$ gives, for $\beta := \sqrt{1-a^2}$:

$$[\rho_B, U^B] = \frac{1}{2(8a+1)} \begin{pmatrix} \beta(u_7 - u_3) & (1-a)u_2 + \beta u_8 & \beta(u_9 - u_1) \\ u_4(a-1) - \beta u_6 & 0 & u_6(a-1) - \beta u_4 \\ \beta(u_1 - u_9) & (1-a)u_8 + \beta u_2 & \beta(u_3 - u_7) \end{pmatrix}$$

It is easy to see that since $0 < a < 1$, in order to have $[\rho_B, U^B] = 0$, we require U^B to be of the form:

$$U^B = \begin{pmatrix} u_1 & 0 & u_3 \\ 0 & u_5 & 0 \\ u_3 & 0 & u_1 \end{pmatrix}.$$

Plugging this U^B and ρ_a into Equation 3.3 and enforcing unitary constraints on the entries of U^B gives (where $\mathrm{Re}(x)$ denotes the real part of some $x \in \mathbb{C}$):

$$\mathrm{d}(\rho_a, U^B) = \frac{a}{8a+1}\sqrt{6 - 2\left(|u_1|^2 + 2\,\mathrm{Re}(u_1^* u_5)\right)} \quad (3.36)$$

To maximize $\mathrm{d}(\rho_a, U^B)$, we thus need to minimize $|u_1|^2 + 2\,\mathrm{Re}(u_1^* u_5)$. To do so, set $u_5 = -1$ and let $u_1 = re^{i\theta}$ for $0 \le r \le 1$, $\theta \in [0, 2\pi]$. Then,

$$|u_1|^2 + 2\,\mathrm{Re}(u_1^* u_5) = r^2 - 2r\cos(\theta), \tag{3.37}$$

which clearly achieves a minimum at $\theta = 0$ and $r = 1$, or equivalently for $u_1 = 1$ and $u_5 = -1$, giving the result $\mathrm{d}_{\max}(\rho_a) = (2\sqrt{2}a)/(8a+1)$, as claimed. It is easy to see that in fact any diagonal unitary matrix of the form

$$U^B = \begin{pmatrix} u_1 & 0 & 0 \\ 0 & -u_1 & 0 \\ 0 & 0 & u_1 \end{pmatrix},$$

gives the same optimum value, as required. □

Note that the limiting value for $\mathrm{d}_{\max}(\rho_a)$ as a approaches 1 is $\frac{2\sqrt{2}}{9}$, which is less than the bound of $\mathrm{d}_{\max}(\rho_{\mathrm{cc}}) \le \frac{1}{\sqrt{2}}$ for two-qubit classically correlated states ρ_{cc} (where consider that one could embed a two-qubit state in $\mathcal{H}^3 \otimes \mathcal{H}^3$). Thus, we cannot detect entanglement in bound entangled states of this specific Horodecki construction using $\mathrm{d}_{\max}(\rho_a)$.

3.5.2 Horodecki$^{\otimes 3}$ Construction

Next, we investigate $\mathrm{d}_{\max}(\rho)$ for a second bipartite one-parameter qutrit bound entangled class of states due to Paweł, Michał, and Ryszard Horodecki [39]. Define, for $2 \le \alpha \le 5$:

$$\begin{aligned} \sigma_+ &= \frac{1}{3}(|01\rangle\langle 01| + |12\rangle\langle 12| + |20\rangle\langle 20|) & (3.38) \\ \sigma_- &= \frac{1}{3}(|10\rangle\langle 10| + |21\rangle\langle 21| + |02\rangle\langle 02|) & (3.39) \\ \rho_\alpha &= \frac{2}{7}|\psi_+\rangle\langle\psi_+| + \frac{\alpha}{7}\sigma_+ + \frac{5-\alpha}{7}\sigma_-, & (3.40) \end{aligned}$$

where $|\psi_+\rangle = \frac{1}{\sqrt{3}}(|00\rangle + |11\rangle + |22\rangle)$. What makes σ_α interesting is that it is separable for $\alpha \in [2,3]$, bound entangled for $\alpha \in (3,4]$, and free entangled for $\alpha \in (4,5]$. Hence an interesting question is, how does $\mathrm{d}(\rho_\alpha, U^B)$ react to these "boundaries" for α?

Note first that $\mathrm{Tr}_A(\rho_\alpha) = I/3$, and so any choice of unitary U^B is cyclic. Letting U^B be some unknown matrix as in Equation 3.30, and substituting into Equation 3.3 and simplifying as much as possible using unitary constraints on U^B, we find:

$$\mathrm{d}(\rho_\alpha, U^B) = \frac{1}{21}\sqrt{9\alpha^2 - 45\alpha + 81 + f(\alpha)}, \tag{3.41}$$

where

$$f(\alpha) = (-3\alpha^2 + 15\alpha - 19)(|u_1|^2 + |u_5|^2 + |u_9|^2) - 8(\mathrm{Re}(u_1 u_5^*) + \mathrm{Re}(u_1 u_9^*) + \mathrm{Re}(u_5 u_9^*)). \tag{3.42}$$

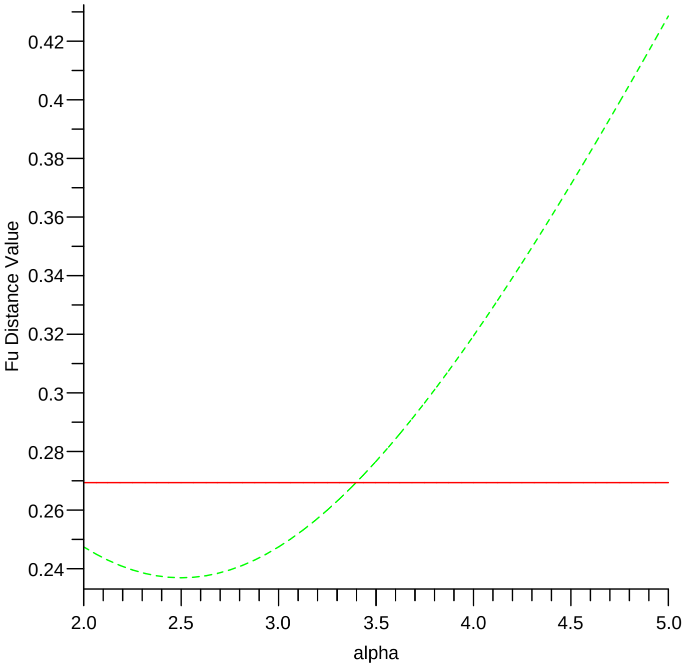

Figure 3.3: Fu shift values for ρ_α for two choices of U^B, as given by Equations 3.43 and 3.44, respectively. The horizontal solid line plots the former, and the curved dashed line plots the latter.

Unfortunately, this is a fairly difficult expression to optimize. By inspection, two seemingly "good" choices of U^B, however, along with corresponding Fu shifts are given below (with the two unitaries denoted U_1^B and U_2^B, respectively):

$$u_1 = -1, u_5 = 1, u_9 = 1 \quad \Rightarrow \quad \mathrm{d}(\rho_\alpha, U_1^B) = \frac{4\sqrt{2}}{21} \tag{3.43}$$

$$u_1 = 0, u_5 = 0, u_9 = 0 \quad \Rightarrow \quad \mathrm{d}(\rho_\alpha, U_2^B) = \frac{1}{21}\sqrt{9\alpha^2 - 45\alpha + 81} \tag{3.44}$$

Here, U_1^B is diagonal, and U_2^B has an all-zero diagonal, as given by Equations 3.43 and 3.44, respectively. Figure 3.3 plots the functions $\mathrm{d}(\rho_\alpha, U_1^B)$ and $\mathrm{d}(\rho_\alpha, U_2^B)$. Note first that depending on our choice of U^B, we could get a Fu shift which may or may not be dependent on the mixing parameter α. Hence, $\mathrm{d}(\rho_\alpha, U_1^B)$ yields a horizontal line, while $\mathrm{d}(\rho_\alpha, U_2^B)$ forms part of a parabola. Second, observe that $\mathrm{d}(\rho_\alpha, U_2^B)$ has its minimum at $\alpha = 2.5$, exactly halfway through the range for α in which ρ_α is separable. It is indeed strange that the Fu shift both decreases and increases within this separable range. Third, contrary to what one may have expected, for both plots, there is no jump or strange behavior in the graph as α increases (and ρ_α correspondingly goes from separable to bound entangled to free entangled). In fact, for $\mathrm{d}(\rho_\alpha, U_1^B)$, the strength of global effect produced is constant regardless of whether σ_α is separable, bound entangled, or free entangled, demonstrating just how drastically the choice of U^B can affect the global shift produced. Finally, for $\mathrm{d}(\rho_\alpha, U_2^B)$, since the ranges of values for separable, bound entangled, and free entangled states are disjoint, if we are promised ρ_α is a state of the given form, we can accurately distinguish between all three cases.

Let us make two further points before closing this section. Note that by Figure 3.3, there may not be a single choice for U^B which is optimal for all α — as seen in Figure 3.3, each choice of U^B has a certain range of α where it outperforms the other choice of U^B. Second, as in Section 3.5.1, note that the maximum value achieved for neither $\mathrm{d}(\rho_\alpha, U_1^B)$ nor $\mathrm{d}(\rho_\alpha, U_2^B)$ is greater than our bound of $\mathrm{d}_{\max}(\rho_{\mathrm{cc}}) \leq \frac{1}{\sqrt{2}}$ for classically-correlated two-qubit states ρ_{cc}. Thus, without prior knowledge that our state is in the form ρ_α, as far as our choices of unitaries here go, we cannot reliably detect bound entanglement for this construction of bound entangled states either using the Fu shift.

3.5.3 Unextendible Product Bases Construction

Finally, let us investigate $\mathrm{d}_{\max}(\rho)$ for a class of bound entangled states whose construction is based on *unextendible product bases (UPB)*. This construction applies to arbitrary dimensions, although we restrict ourselves to the case of two-qutrit states here for simplicity. A UPB is defined as follows:

Definition 10 (Unextendible Product Basis (UPB) [9])**.** Consider a bipartite quantum system $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2$, both subsystems of arbitrary dimension. Define an incomplete orthogonal *product basis* (PB) as a set S of pure orthogonal product states spanning proper subspace $\mathcal{H}_S$ of $\mathcal{H}$. Then an *unextendible product basis* (UPB) is a PB whose complementary subspace $\mathcal{H} - \mathcal{H}_\mathcal{S}$ contains no product state.

Note that this definition extends straightforwardly to the multipartite case [9]. Using a UPB, one can then systematically construct bound entangled states using the following theorem.

Theorem 21 (Bennett et al. [9])**.** *Given UPB $\{|\psi\rangle_i\}_{i=0}^{n-1}$ in a Hilbert space of total dimension D, the following state is bound entangled:*

$$\rho = \frac{1}{D-n}\left(I - \sum_{k=0}^{n-1} |\psi_k\rangle\langle\psi_k|\right) \tag{3.45}$$

Thus, we simply need to find an UPB in Hilbert space $\mathcal{H}^3 \otimes \mathcal{H}^3$, and apply Theorem 21 to obtain a bound entangled state ρ to test $\mathrm{d}_{\max}(\rho)$ on. An example of such an UPB (and corresponding bound entangled state ρ) is given by [9]:

$$\begin{aligned}
|\psi_0\rangle &= \frac{1}{\sqrt{2}}|0\rangle(|0\rangle - |1\rangle) \\
|\psi_1\rangle &= \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)|2\rangle \\
|\psi_2\rangle &= \frac{1}{\sqrt{2}}|2\rangle(|1\rangle - |2\rangle) \\
|\psi_3\rangle &= \frac{1}{\sqrt{2}}(|1\rangle - |2\rangle)|0\rangle \\
|\psi_4\rangle &= \frac{1}{3}(|0\rangle + |1\rangle + |2\rangle)(|0\rangle + |1\rangle + |2\rangle) \\
\rho &= \frac{1}{4}\left(I - \sum_{k=0}^{4}|\psi_k\rangle\langle\psi_k|\right).
\end{aligned}$$

Let us analyze $\mathrm{d}_{\max}(\rho)$. Letting U^B be an arbitrary 3×3 complex matrix as per Equation 3.30, and investigating the commutator relation $[\rho_B, U^B]$, we find that in order to be cyclic, U^B must be of the following form:

$$\begin{pmatrix} u_9 & u_8 & u_7 \\ u_8 & 5u_8 + u_7 + u_9 & u_8 \\ u_7 & u_8 & u_9 \end{pmatrix} \tag{3.46}$$

Substituting U^B into Equation 3.3 and simplifying as much as possible using unitary constraints on U^B, we have the following:

$$\mathrm{d}(\rho, U^B) = \frac{1}{2}\sqrt{-\frac{35}{8}|u_8|^2 + \frac{1}{2}|u_7|^2 - \frac{11}{2}\mathrm{Re}(u_7 u_8^*)} \tag{3.47}$$

This again proves difficult to optimize, but we can derive a (loose) upper bound on $\mathrm{d}(\rho, U^B)$ based on Equation 3.47 which shall suffice for our purposes, as follows. If we relax the unitary constraints on U^B and demand only that $|u_7| \leq 1$ and $|u_8| \leq 1$, then by inspection it is always optimal to choose $u_7 = -1$ in Equation 3.47, since $\frac{35}{8} < \frac{11}{2}$. The resulting expression, $\frac{1}{2}\sqrt{-\frac{35}{8}u_8^2 + \frac{11}{2}u_8 + \frac{1}{2}}$, is maximized for $u_8 = \frac{22}{35}$. Clearly, such choices for u_7 and u_8 give non-unitary U^B, but they also give an (unattainable) upper bound on $\mathrm{d}_{\max}(\rho)$ of $\mathrm{d}_{\max}(\rho) < 0.7464$. Thus, with our only known bound of $\mathrm{d}_{\max}(\rho_{\mathrm{cc}}) \leq \frac{2\sqrt{2}}{3} \approx 0.9428$ for classically correlated two-qutrit states ρ_{cc}, we cannot detect bound entanglement in this construction based on unextendible product bases either. We close by remarking that for sake of comparison, by inspection, a "fairly good" choice of U^B in this case is

$$U^B = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \tag{3.48}$$

with which we achieve $\mathrm{d}(\rho, U^B) = \frac{\sqrt{2}}{4} \approx 0.3536$.

3.6 Can We Avoid Classical Correlations?

We have seen that although $\mathrm{d}_{\max}(\rho)$ exhibits anomalies similar to those observed for non-locality measures (as per Section 3.3.1), $\mathrm{d}_{\max}(\rho)$ is itself not a non-locality measure, since there exist classically correlated states ρ_{cc} for which $\mathrm{d}_{\max}(\rho_{\mathrm{cc}}) > 0$. We now briefly ask, is there some variant $\mathrm{d}'(\rho)$ of $\mathrm{d}_{\max}(\rho)$, also based on locally invariant unitary operations, but for which $\mathrm{d}'(\rho_{\mathrm{cc}}) = 0$ for all ρ_{cc}?

A first attempt is to consider alternate metrics. A simple counterexample, however, shows that at least for any metric induced by a norm on the given Hilbert space, avoiding $\mathrm{d}_{\max}(\rho_{\mathrm{cc}}) > 0$ is impossible. Consider, for example, classically correlated state $\rho_{\mathrm{cc}} = \frac{1}{2}|00\rangle\langle 00| + \frac{1}{2}|11\rangle\langle 11|$, and $U^B = \sigma_x$, for σ_x the Pauli X operator. Then, we have $(I \otimes U^B)\rho_{\mathrm{cc}}(I \otimes U^{B\dagger}) = \frac{1}{2}|01\rangle\langle 01| + \frac{1}{2}|10\rangle\langle 10| \neq \rho_{\mathrm{cc}}$, and so $\mathrm{d}'(\rho_{\mathrm{cc}}) > 0$ for any $\mathrm{d}'(\rho)$ based on a norm-induced metric, since by definition, $\|\vec{x}\|_2 = 0$ if and only if $\vec{x} = \vec{0}$ for any norm and vector $\vec{x}$ [6].

A second attempt is to instead consider $\mathrm{d}_{\min}(\rho) := \min_{\text{cyclic } U^B} \mathrm{d}(\rho, U^B)$, where we must demand $U^B \neq I$, since otherwise any state ρ would clearly achieve $\mathrm{d}_{\min}(\rho) = 0$. Despite constituting a poor quantification scheme, $\mathrm{d}_{\min}(\rho)$ demonstrates the following interesting point. Consider two-qubit classically-correlated state $\rho_{\mathrm{cc}} = \frac{1}{2}|00\rangle\langle 00| + \frac{1}{2}|--\rangle\langle --|$, where $|-\rangle = \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)$. Using the results of Section 3.4.1, specifically Equation 3.26, we have $\vec{n} = (-\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}})$, $T_{11} = T_{33} = \frac{1}{2}$, $T_{22} = 0$, and so the only way to achieve $\mathrm{d}_{\min}(\rho) = 0$ is to set $\theta = 0$, which implies $U^B = I$. Thus, there exist classically correlated states which are incapable of *not* achieving a global shift under locally invariant unitary operations (unless $U^B = I$), and so $\mathrm{d}_{\min}(\rho)$ does not possess our desired property either.

3.7 A Brief Conclusion and Open Problems

We have investigated the global power of locally invariant unitary operations, mostly within the context of entanglement detection and non-locality. Specifically, we have seen closed formulas for $\mathrm{d}_{\max}(\rho)$ for the bipartite cases of (pseudo)pure quantum states, Werner states, and two-qubit states. The first of these reveals the existence of non-maximally entangled states capable of achieving a maximal global shift, both disallowing one from defining an entanglement measure based on $\mathrm{d}_{\max}(\rho)$, and revealing similarities to anomalies seen in non-locality measures. Since $\mathrm{d}_{\max}(\rho)$ is neither an entanglement measure, nor a non-locality measure, yet as demonstrated here possesses clear connections to the CHSH inequality in the two-qubit case, it would be interesting to have a better intuitive understanding of the correlations (both classical and quantum) that $\mathrm{d}_{\max}(\rho)$ is quantifying, and if and how the anomalies mentioned above are related to those seen in non-locality measures.

There are a number of questions which remain open. First, despite the fact that our formula for $\mathrm{d}_{\max}(\rho)$ for Werner states demonstrates diminishing distinguishability between classical and quantum correlations in Werner states as the dimension grows, it remains for a tight upper bound on $\mathrm{d}_{\max}(\rho)$ to be found for general classically correlated states of (subsystem) dimension $d \geq 3$ in order to conclusively state the efficacy of $\mathrm{d}_{\max}(\rho)$ as an entanglement detection criteria. This would also allow one to extend our analysis here for bound entangled states to higher dimensions. Second, although we have demonstrated that any entangled pseudopure or Werner state achieves $\mathrm{d}_{\max}(\rho) > 0$, it is still not known whether this holds for all entangled bipartite states. Third, it would of course be of interest to determine whether a closed formula for $\mathrm{d}_{\max}(\rho)$ can be derived for mixed states in general, the existence of which would not contradict known hardness results for the quantum separability problem [31]. Finally, as mentioned briefly in Section 3.1, the principle behind the Fu shift is implicitly applied in superdense coding, and we would be curious to know whether there exist any other applications in quantum computing and information.

Chapter 4

Conclusion

As we have included concluding comments and open problems at the close of each chapter, we shall keep our discussion here brief, so as not to duplicate previous sections. In this thesis, we have discussed both the computational complexity of the Quantum Separability problem, as well as one specific attempt at solving it via locally invariant unitary operations. Specifically, we have seen that when phrased as an instance of the Weak Membership problem over the set of separable quantum states, the Quantum Separability Problem can be shown to be strongly NP-hard, meaning it is still a computationally intractable problem (in the general case, unless $\mathbf{P} = \mathbf{NP}$) even when one allows "moderate" or inverse polynomial error (with respect to dimension). We observed that this implies immediate lower bounds on the Euclidean distance from the set of separable quantum states in which bound entangled states are known to exist, as well as strong NP-hardness of Weak Membership over the set of entanglement-breaking maps.

We next investigated the abilities of locally invariant unitary operations in detecting entanglement, as quantified by what we called the Fu shift. Closed formulas for this distance for the cases of (pseudo)pure quantum states, Werner states, and two-qubit states were derived, from which we drew a number of conclusions. The first of these was that the Fu distance cannot be used to define an entanglement measure, since there exist non-maximally entangled states capable of achieving shifts equally large as maximally entangled states. Despite being surprising in itself, this latter property is similar to anomalies seen for non-locality measures. We also showed that for certain classes of two-qubit states, there exists an equivalence between the Fu distance and the CHSH inequality, but that generally the former is a weaker entanglement detection criterion than the latter (at least for two-qubit states).

We close by remarking that both main topics discussed here highlight only the beginning of what has been and promises to continue being an interesting area of research for many years to come. With NP-hardness results in place, the challenge has been extended to researchers to find clever ways of tackling the quantum separability problem efficiently for special cases of interest. Further, anomalies as those seen for non-locality measures, and the similar behavior observed here for the Fu distance hopefully make it clear that our understanding of entanglement and non-locality is in many ways still in its early stages of infancy.

Appendix A

Lemma 22. *Given density operators ρ_1 and ρ_2 of quantum states of dimension N, with corresponding Bloch vectors $\vec{\alpha}$ and $\vec{\beta}$ given by equations $\rho_1 = \frac{I}{N} + \frac{1}{2}\vec{\alpha}\cdot\vec{\sigma}$ and $\rho_2 = \frac{I}{N} + \frac{1}{2}\vec{\beta}\cdot\vec{\sigma}$, respectively, we have $\sqrt{2}\,\|\rho_1 - \rho_2\|_{\mathrm{F}} = ||\vec{\alpha} - \vec{\beta}||_2$.*

Proof. Via straightforward manipulation, we have:

$$\begin{aligned}
\|\rho_1 - \rho_2\|_{\mathrm{F}} &= \left\| \left(\frac{I}{N} + \frac{1}{2}\sum_{i=1}^{N^2-1} \alpha_i \sigma_i \right) - \left(\frac{I}{N} + \frac{1}{2}\sum_{j=1}^{N^2-1} \beta_j \sigma_j \right) \right\|_{\mathrm{F}} \\
&= \frac{1}{2} \left\| \sum_{i=1}^{N^2-1} (\alpha_i - \beta_i)\sigma_i \right\|_{\mathrm{F}} \\
&= \frac{1}{2}\sqrt{\mathrm{Tr}\left((\sum_{i=1}^{N^2-1} (\alpha_i - \beta_i)\sigma_i)^\dagger (\sum_{j=1}^{N^2-1} (\alpha_j - \beta_j)\sigma_j) \right)} \\
&= \frac{1}{2}\sqrt{\sum_{i=1,j=1}^{N^2-1} (\alpha_i - \beta_i)(\alpha_j - \beta_j) 2\delta_{ij}} \\
&= \frac{1}{\sqrt{2}} ||\vec{\alpha} - \vec{\beta}||_2
\end{aligned}$$

□

Lemma 23. *In Chapter 2, Section 2.3, combining Theorem 2 and Lemma 5 gives $\|\hat{\mathbf{c}}\|_2 \in O(\sqrt{N})$.*

Proof. By definition, we have:

$$\|\hat{\mathbf{c}}\|_2 = \sqrt{\sum_{i=1}^{m} \left[\frac{1}{2}\mathrm{Tr}(C\sigma_i) \right]^2}, \tag{A.1}$$

where $m = M^2N^2 - 1$.

Recall now the definition of C from Lemma 4 of Chapter 2, where in our case, each $A_i \in \mathbb{R}^{N\times N}$ is all zeroes except for its upper left corner, which is set to submatrix $B_i \in \mathbb{R}^{l \times l}$ from Theorem 2. Each B_i in turn is all zeroes, except for some index (k, l) (and hence (l, k), by symmetry), $1 \leq k < l \leq n$, which is set to the (l, k)th entry of the adjacency matrix A_G of graph G.

We also require an explicit construction for the generators σ_i of $SU(D)$, where $D = m + 1$, as given for example in [45], where $\{\sigma_i\}_{i=1}^{D^2-1} = \{U_{pq}, V_{pq}, W_r\}$, such that for $1 \leq p < q \leq D$ and $1 \leq r \leq D - 1$, and $\{\mathbf{x}_i\}_{i=1}^{D}$ an orthonormal basis for Hilbert space $\mathcal{H}^D$:

$$U_{pq} = \mathbf{x}_p \mathbf{x}_q^\dagger + \mathbf{x}_q \mathbf{x}_p^\dagger \tag{A.2}$$

$$V_{pq} = -i\mathbf{x}_p \mathbf{x}_q^\dagger + i\mathbf{x}_q \mathbf{x}_p^\dagger \tag{A.3}$$

$$W_r = \sqrt{\frac{2}{r(r+1)}} \left[\left(\sum_{k=1}^{r} \mathbf{x}_k \mathbf{x}_k^\dagger \right) - r\mathbf{x}_{r+1} \mathbf{x}_{r+1}^\dagger \right]. \tag{A.4}$$

Due to the symmetry of C and the fact that $\mathrm{Tr}(C) = 0$, it is clear that only the generators of the form U_{pq} will contribute to the sum in Equation A.1. Further, for each edge in G, $\mathrm{Tr}(CU_{pq}) = 2$ for each U_{pq} whose non-zero indices match those of the entries in C corresponding to that edge. Since each edge contributes four (symmetrically placed) entries of 1 to C, we hence have $\|\hat{\mathbf{c}}\|_2 = \frac{1}{2}\sqrt{(2\hat{e})2^2} = \sqrt{2\hat{e}}$, where $\hat{e}$ denotes the number of edges in G. Since $\hat{e} \in O(n^2)$ (n the number of vertices in G), and $M \in \Theta(n^2)$, we have $\|\hat{\mathbf{c}}\|_2 \in O(\sqrt{M})$, as required. □

Institut für Theoretische Physik III
Prof. Dr. Dagmar Bruß

HEINRICH HEINE
UNIVERSITÄT
DÜSSELDORF

Universitätsstr. 1
Gebäude 25.32
D-40225 Düsseldorf
Tel.: (0211) 81-13992
Fax: (0211) 81-11337
E-Mail:
bruss@thphy.uni-duesseldorf.de

Theoretische Physik, Universitätsstr. 1, D-40225 Düsseldorf

July 11, 2008

Sevag Gharibian
Institute for Quantum Computing
University of Waterloo
200 University Ave. W.

N2L 3G1
Canada

To whom it may concern

I consent to the inclusion of our joint work on Local Cyclic Operations in the Masters thesis of Sevag Gharibian, the latter written in fulfillment of the requirements for a degree of Master of Mathematics at the University of Waterloo in Canada, as well as to the pertinent signing of a Theses Non-Exclusive License for Library and Archives Canada, which I have read and understood.

Sincerely yours

(Prof. Dr. D. Bruß) (Dr. H. Kampermann)

References

[1] A. Acín, T. Durt, N. Gisin, and J. I. Latorre. Quantum nonlocality in two three-level systems. *Phys. Rev. A*, 65(052325), 2002. 2, 26, 31

[2] T. Ando. Cones and norms in the tensor product of matrix spaces. *Linear Alg. Appl.*, 379:3, 2004. 21

[3] K. Audenaert, F. Verstraete, and B. De Moor. Variational characterizations of separability and entanglement of formation. *Phys. Rev. A*, 64:052304, 2001. 31

[4] P. Badziąg, P. Deuar, M. Horodecki, P. Horodecki, and R. Horodecki. Concurrence in arbitrary dimensions. *J. of Modern Optics*, 49(8):1289–1297, 2002. 31

[5] J. S. Bell. On the Einstein Podolsky Rosen paradox. *Physics*, 1:195, 1964. 1

[6] I. Bengtsson and K. Życzkowski. *Geometry of Quantum States: An Introduction to Quantum Entanglement.* Cambridge University Press, 2006. 2, 8, 21, 22, 23, 26, 44

[7] C. Bennett, G. Brassard, C. Crépeau, R. Josza, A. Peres, and W. K. Wootters. Teleporting an unknown quantum state via dual classical and Einstein-Podolsky-Rosen channels. *Phys. Rev. Lett.*, 70:1895—1899, 1993. 1

[8] C. Bennett and S. Wiesner. Communication via one- and two-particle operators on Einstein-Podolsky-Rosen states. *Phys. Rev. Lett.*, 69:2881–2884, 1992. 1, 25

[9] C. H. Bennett, D. P. DiVincenzo, T. Mor, P. W. Shor, J. A. Smolin, and B .M. Terhal. Unextendible product bases and bound entanglement. *Phys. Rev. Lett.*, 82(26), 1999. 42, 43

[10] C. H. Bennett, D. P. DiVincezo, J. Smolin, and W. K. Wootters. Mixed-state entanglement and quantum error correction. *Phys. Rev. A*, 54:3824, 1996. 1

[11] D. Bertsimas and S. Vempala. Solving convex programs by random walks. *J. of the ACM*, 51(4), 2004. 4, 23

[12] N. Brunner, N. Gisin, and V. Scarani. Entanglement and non-locality are different resources. *New Journal of Physics*, 7:88, 2005. 2, 26

[13] D. Bruß. Characterizing entanglement. *J. Math. Phys.*, 43:4237, 2002. 2, 21, 34, 38

[14] M. D. Choi. Completely positive linear maps on complex matrices. *Linear Alg. Appl.*, 10:285, 1975. 21

[15] M. Christandl and A. Winter. Squashed entanglement: An additive entanglement measure. *J. Math. Phys.*, 45:829, 2004. 1

[16] J. F. Clauser, M. A. Home, A. Shimony, and R. A. Holt. Proposed experiment to test local hidden-variable theories. *Phys. Rev. Lett.*, 23:880–884, 1969. 2, 25, 26, 34

[17] R. Cleve and H. Buhrman. Substituting quantum entanglement for communication. *Phys. Rev. A*, 56:1201–1204, 1997. 1

[18] T. H. Cormen, C. E. Leiserson, and R. L. Rivest. *Introduction to Algorithms.* The MIT Press, 1990. 17

[19] G. B. Dantzig. Discrete-variable extremum problems. *Operations Res.*, 5:266–277, 1957. 4

[20] M. J. Donald, M. Horodecki, and O. Rudolph. The uniqueness theorem for entanglement measures. *J. Math. Phys.*, 43:4252–4272, 2002. 1

[21] P. Eberhard. Background level and counter efficiencies required for a loophole free Einstein-Podolsky-Rosen experiment. *Phys. Rev. A*, 47:R747–R750, 1993. 2, 26

[22] A. Einstein, B. Podolsky, and N. Rosen. Can quantum-mechanical description of physical reality be considered complete? *Phys. Rev.*, 47:777, 1935. 1

[23] A. Ekert. Quantum cryptography based on Bell's theorem. *Phys. Rev. Lett.*, 67:661–663, 1991. 1

[24] U. Fano. Pairs of two-level systems. *Rev. Mod. Phys.*, 55:855–874, 1983. 34

[25] L. B. Fu. Nonlocal effect of bipartite system induced by local cyclic operation. *Europhys. Lett.*, 75:1–7, 2006. 2, 25, 26, 27, 28, 31, 34, 35

[26] J. Gallier. Geometric methods and applications for computer science and engineering. 2000. 36

[27] M. R. Garey and D. S.Johnson. *Computers and Intractability: A Guide to the theory of NP-completeness.* W. H. Freeman and Company, New York, 1979. 4

[28] N. Gisin. Hidden quantum nonlocality revealed by local filters. *Phys. Lett. A*, 210(3). 22

[29] M. Grötschel, L. Lovász, and A. Schrijver. *Geometric Algorithms and Combinatorial Optimization.* Springer-Verlag, Berlin, 1988. 4, 7, 8, 13, 17, 18, 19

[30] L. K. Grover. Quantum telecomputation. *quant-ph/9704012*, 1997. 1

[31] L. Gurvits. Classical deterministic complexity of Edmonds problem and quantum entanglement. In *Proceedings of the thirty-fifth ACM symposium on Theory of computing*, pages 10–19, New York, 2003. ACM Press. 1, 2, 3, 4, 9, 11, 45

[32] L. Gurvits and H. Barnum. Largest separable balls around the maximally mixed bipartite quantum state. *Phys. Rev. A*, 66(062311), 2002. 8

[33] S. Hill and W. K. Wootters. Entanglement of a pair of quantum bits. *Phys. Rev. Lett.*, 78:5022, 1997. 31

[34] A. S. Holevo. Coding theorems for quantum channels. *Russian Math. Surveys*, 53:1295–1331, 1999. 21

[35] M. Horodecki, P. Horodecki, and R. Horodecki. Separability of mixed states: necessary and sufficient conditions. *Phys. Lett. A*, 223:1, 1996. 1, 21

[36] M. Horodecki, P. Horodecki, and R. Horodecki. Mixed-state entanglement and distillation: Is there a “bound” entanglement in nature? *Phys. Rev. Lett.*, 80(24), 1998. 5, 20, 38

[37] M. Horodecki, P. W. Shor, and M. B. Ruskai. Entanglement breaking channels. *Rev. Math. Phys.*, 15:629, 2003. 5, 21, 22

[38] P. Horodecki. Separability criterion and inseparable mixed states with positive partial transposition. *Phys. Lett. A*, 232(5), 1997. 38, 39

[39] P. Horodecki, M. Horodecki, and R. Horodecki. Bound entanglement can be activated. *Phys. Rev. Lett.*, 82(5), 1999. 38, 40

[40] R. Horodecki and P. Horodecki. Perfect correlations in the Einstein-Podolsky-Rosen experiment and Bell’s inequalities. *Phys. Lett. A*, 210:227, 1996. 36

[41] R. Horodecki, P. Horodecki, and M. Horodecki. Violating Bell inequality by mixed spin-1/2 states: necessary and sufficient condition. *Phys. Lett. A.*, 200:340–344, 1995. 34

[42] R. Horodecki, P. Horodecki, M. Horodecki, and K. Horodecki. Quantum entanglement. arXiv:quant-ph/0702225v2, 2007. 2

[43] L. Ioannou. Computational complexity of the quantum separability problem. *Quantum Info. and Comp.*, 7(4):335, 2007. 2, 4, 9, 12, 14

[44] A. Jamiołkowski. Linear transformations which preserve trace and positive semi-definiteness of operators. *Rep. Math. Phys.*, 3:275, 1972. 21

[45] G. Kimura. The bloch vector for n-level systems. *Phys. Lett. A*, 314(5), 2003. 9, 12, 49

[46] K. Kraus. General state changes in quantum theory. *Ann. Phys.*, 64:311, 1971. 21, 23

[47] J. Leinaas, J. Myrheim, and E. Ovrum. Geometrical aspects of entanglement. *Phys. Rev. A*, 74(012313), 2006. 22

[48] Y.K. Liu. The complexity of the consistency and N-representability problems for quantum states. `http://arxiv.org/abs/0712.3041`, 2007. PhD Thesis. 2, 4, 7, 8, 9, 14, 16, 17, 18, 19, 23

[49] A. A. Méthot. Private communication, 2008. 26

[50] A. A. Méthot and V. Scarani. An anomaly of non-locality. *Quantum Info. and Comp.*, 7:157–170, 2007. 2, 26

[51] D. S. Mitrinovic, J. Pecaric, and V. Volenec. *Recent Advances in Geometric Inequalities.* Springer, 1989. 29

[52] T. S. Motzkin and E. G. Straus. Maxima for graphs and a new proof of a theorem of Túran. *Canadian J. Math.*, 17:533–540, 1965. 10

[53] M. A. Nielsen and I. L. Chuang. *Quantum Computation and Quantum Information.* Cambridge University Press, 2000. 2, 28

[54] M. A. Nielsen and J. Kempe. Separable states are more disordered globally than locally. *Phys. Rev. Lett.*, 86:5184, 2001. 1

[55] A. Peres. Separability criterion for density matrices. *Phys. Rev. Lett.*, 77(8), 1996. 1, 21, 38

[56] P. Rungta, V. Bŭzek, C. M. Caves, M. Hillery, and G. J. Milburn. Universal state inversion and concurrence in arbitrary dimensions. *Phys. Rev. A*, 64:042315, 2001. 31

[57] E. Schrödinger. Discussion of probability relations between separated systems. *Proc. Cambridge Phil. Soc.*, 31:555–563, 1935. 1

[58] P. Shor. Algorithms for quantum computation: discrete logarithms and factoring. In *Proc. of 35th Ann. Symp. on Found. of Comp. Sci.*, pages 124–134, 1994. 1

[59] A. Uhlmann. Fidelity and concurrence of conjugated states. *Phys. Rev. A*, 62:03237, 2000. 31

[60] V. Vedral, M. B. Plenio, M. A. Rippin, and P. L. Knight. Quantifying entanglement. *Phys. Rev. Lett.*, 78:2275, 1997. 1

[61] G. Vidal and R. F. Werner. Computable measure of entanglement. *Phys. Rev. A*, 65:032314, 2002. 1

[62] E. B. Vinberg. *Linear Representations of Groups.* Birkhäuser Basel, 1989. 8

[63] R. F. Werner. Quantum states with Einstein-Podolsky-Rosen correlations admitting a hidden-variable model. *Phys. Rev. A*, 40(8):4277–4281, 1989. 25, 26, 32

[64] E. T. Whittaker and G. N. Watson. *A Course of Modern Analysis.* Cambridge University Press, 1958. 7

[65] W. K. Wootters. Entanglement of formation of an arbitrary state of two qubits. *Phys. Rev. Lett.*, 80:2245, 1998. 31

[66] W. K. Wootters. Entanglement of formation and concurrence. *Quantum Info. and Comp.*, 1(1):27–44, 2001. 31

[67] D. B. Yudin and A. S. Nemirovskii. Informational complexity and efficient methods for the solution of convex extremal problems (in Russian). *Ekonomica i Matematicheskie Metody*, 12:357–369, 1976. 4