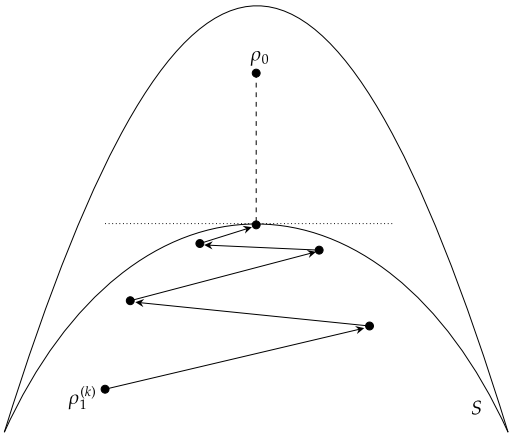
APPLICATION OF CHOSEN OPTIMIZATION ALGORITHMS FOR RECOGNITION OF NONCLASSICAL EFFECTS

PALASH PANDYA



SUPERVISOR:
dr. hab. Marcin Wieśniak

Institute of Theoretical Physics and Astrophysics
Faculty of Mathematics, Physics and Informatics
University of Gdańsk

STRESZCZENIE

Praca pod tytułem „Application of chosen optimization algorithms for recognition of nonclassical effects" składa się z sześciu rozdziałów, a także spisów tabel i figur, bibliografii oraz dodatku matematycznego. W pierwszym rozdziale sformułowany został problem, którego będzie dotyczyć rozprawa: problem splątania stanów wielu podukładów kwantowych, zdefiniowany jest problem k-separowalności. Następnie następuje bardzo syntetyczny przegląd wybranych kryteriów i miar splątania. Ma on na celu podkreślenie trudności, z jakimi się spotykamy próbując je zastosować lub policzyć. Część z tych trudności jest szczegółowo przedyskutowana w podrozdziale 1.5. Rozdział drugi opisuje wyniki z pracy "Hilbert-Schmidt distance and entanglement witnessing" [39]. Rozdział ten rozpoczyna formalizacja problemów słabej separacji i optymalizacji, których wariantem jest detekcja splątania w stanie. Następnie wprowadzona jest miara Hilberta-Schmidta dla macierzy kwadratowych dowolnego wymiaru, będąca bezpośrednim uogólnieniem odległości kartezjańskiej w przestrzeniach wektorowych. Jest to jedyna miara niezmiennicza względem operacji unitarnych i jednocześnie nie wymagająca diagonalizacji, co czyni ją bardzo efektywną w obliczaniu. W kolejnej części rozdziału zaprezentowany został algorytm Gilberta, który w zbiorze wypukłym znajduje przybliżenie najbliższego punktu do danego. Jeżeli interesujący nas punkt leży wewnątrz zbioru, algorytm wskazuje ten punkt, w przeciwnym razie, jednym z wyników będzie przybliżona odległość punktu od zbioru. W podrozdziale 2.4 algorytm jest adaptowany do analizy stanów kwantowych. W ostatniej części drugiego rozdziału w szczególności przedyskutowana została generacja stanów czystych, przy pomocy których algorytm optymalizuje zwracany stan. Są one generowane zgodnie z miarą Haara, by algorytm był równie skuteczny dla wszystkich stanów. Wskazana jest też jednoznaczność najbliższego stanu separowalnego, a także dane wyjściowe z algorytmu. Rozdział 3 prezentuje wyniki algorytmu dla wybranych przykładów. Stany maksymalnie splątane dwóch kuditów (kwantowych układów d-poziomowych) pokazują konieczność używania liczb zespolonych w optymalizacji. Analiza stanów GHZ omówiona w podrozdziale 3.2 prowadzi do analitycznej formy najbliższego stanu separowalnego, co nie udaje się

ze stanami W omówionymi w następnej części pracy, lecz pokazujemy możliwą analityczną postać. W kolejnym podrozdziale algorytm zostaje zastosowany do problemu biseparowalności, gdzie są zaprezentowane nowe własności geometryczne dla zbioru stanów separowalnych oraz biseparowalnych. Rozdział 4 opiera się na pracy "Hilbert-Schmidt distance and entanglement witnessing" [39] i dotyczy użycia algorytmu Gilberta do konstrukcji świadectw splątania. Pojęcie świadectwa (świadka splątania) zostało opisane w pierwszych dwóch podrozdziałach. Podrozdział 4.3 opisuje związek pomiędzy świadectwami a najbliższymi stanami separowalnymi, a kolejny przedstawia próbę ich dalszej optymalizacji. Podrozdziały 4.5 i 4.6 opisują przykłady takich świadectw, ze szczególnym uwzględnieniem stanów ze splątaniem związanym z nierozszszerzalnych baz produktowych. Piąty rozdział dotyczy manuskryptu "An elegant proof of self-testing for multipartite Bell inequalities" [40] i dyskutuje samotestowanie się szerokiej klasy nierówności Bella. W pierwszym podrozdziale przedstawione jest znaczenie samotestowania schematów Bellowskich dla kryptografii. Następnie praca opisuje dowód samotestowania wszystkich nierówności Bella z dwoma lokalnymi, projektywnymi obserwablami dla korelacji dwucząstkowych. Szczególne znaczenie mają tutaj nierówności Uffinka będące nieliniowym kryterium obecności splątania N-cząstkowego.

ABSTRACT

The work entitled "Application of chosen optimization algorithms for recognition of nonclassical effects" consists of six chapters, as well as lists of tables and figures, a bibliography and a mathematical appendix. The first chapter formulates the problem that will be central to the thesis: the problem of detecting entanglement in multiparty systems and the problem of k-separability. This is followed by a very synthetic review of the selected entanglement criteria and measures. Its purpose is to highlight the difficulties we encounter when trying to apply or measure them in higher dimensional Hilbert spaces. Some of these difficulties are discussed in detail in subsection 1.5. The second chapter describes the results of our work "Hilbert-Schmidt distance and entanglement witnessing" [39]. This chapter begins with the formalization of weak separation and optimization problems, an equivalent problem of which is entanglement-in-state detection. Then the Hilbert-Schmidt measure is introduced for square matrices of any dimension, which is a direct generalization of Cartesian distance in vector spaces. It is the only measure invariant in relation to unitary operations and, at the same time, does not require diagonalization, which makes it very effective in calculations. The next part of the chapter presents Gilbert's algorithm, which in a convex set finds the approximation of the closest point in the set to a given point. If the point of interest lies inside the set, the algorithm points to that point, otherwise one of the results will be the approximate distance of the point from the set. In subsection 2.4, the algorithm is adapted to the quantum state analysis. In the last part of the second chapter, the generation of pure states, with the help of which the algorithm optimizes the returned state, is discussed in particular. They are generated according to the Haar measure to ensure that the algorithm covers the whole state space effectively. The uniqueness of the closest separable state is also discussed and proved. Then the output data from the algorithm is described as well. Chapter 3 presents the results of the algorithm for selected examples. The maximally entangled states of two qu$d$its (quantum d-level systems) show the necessity to use complex numbers in optimization. The analysis of GHZ states discussed in subsection 3.2 leads to the analytical form of the closest separable state. The analysis fails with the W

states discussed in the next part of the work, but we still provide a possible analytical form. In the next section, the algorithm is applied to the biseparability problem, and novel geometrical insights for the set of separable and biseparable states are presented. Additional examples are described in section 3.5. Chapter 4 builds on [58] and deals with the application of Gilbert's algorithm to construct Entanglement Witnesses. The concept of Witnessing (Entanglement Witnesses) entanglement is described in the first two sections. Section 4.3 describes the relationship between the Entanglement Witnesses and the closest separable states, and in the next one we attempt to further optimize the Entanglement Witnesses. Sections 4.5 and 4.6 describe examples of such evidence, with particular emphasis on Bound Entangled states associated with Unextendible Product Bases. The fifth chapter deals with the manuscript "An elegant proof of self-testing for multipartite Bell inequalities" [40] and discusses self-testing of the broad class of Bell inequalities. The first section outlines the importance of device independence and self-testing in Bell scenarios for cryptography. The paper then describes the proof of self-testing for all Bell inequalities with two local, projective observables for N-particle correlations. The Uffink inequalities, which are a nonlinear criterion for the presence of N-particle entanglement, are of particular importance here.

PUBLICATIONS

The results presented in this thesis are based on the following works:

[1] Palash Pandya, Omer Sakarya, and Marcin Wie śniak. "Hilbert-Schmidt distance and entanglement witnessing." In: *Phys. Rev. A* 102 (1 July 2020), p. 012409. DOI: 10.1103/PhysRevA.102.012409. URL: https://link.aps.org/doi/10.1103/PhysRevA.102.012409.

[2] Ekta Panwar, Palash Pandya, and Marcin Wieśniak. "An elegant proof of self-testing for multipartite Bell inequalities." In: (2022). arXiv: 2202.06908 [quant-ph].

[3] Marcin Wieśniak et al. "Distance between Bound Entangled States from Unextendible Product Bases and Separable States." In: *Quantum Reports* 2.1 (2020), pp. 49–56. ISSN: 2624-960X. DOI: 10.3390/quantum2010004. URL: https://www.mdpi.com/2624-960X/2/1/4.

ACKNOWLEDGMENTS

I would first like to acknowledge my supervisor, dr hab. Marcin Wieśniak, prof. UG. His kind, patient guidance and support have been invaluable for this work as well as for me. I would also like to thank my dear friends and colleagues, Bianka and Omer, for making this a home away from home, I look up to you both; to Anubhav, Mahasweta, Ekta, Ray, Tanmoy for numerous discussions, and constant support. I would also like to thank Małgorzata Szczekocka, one of the kindest and most helpful person I've ever known.

I would like to express my gratitude to Daria, who made this work possible with her love and support. And lastly, Avni, you are the source of all my joy.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

1

PRELIMINARIES

One of the most important problem in the field of Quantum Information that still remains without a solution that can be computed efficiently is the *Separability problem*, that is, deciding if a given state is separable or entangled. The classification of any given quantum state residing in an arbitrary Hilbert space as *Entangled* or *Separable* has been proven to be a *NP-HARD* problem. Essentially, finding an algorithm to establish such a classification that has a complexity polynomial in the dimension of the Hilbert space is impossible. While several separability criteria exist, most of them are only necessary and sufficient in the bipartite case, and even then at best in the 2×2 or 2×3 Hilbert spaces. In this thesis our aim is to discuss a simple yet effective way to explore the quantum state space of arbitrary dimensions. A way to detect and quantify the entanglement of a given state based on the minimum Hilbert-Schmidt Distance of the state to the convex set of separable states, and in the process provide state-specific Entanglement Witnesses that are close to optimal. All of this is accomplished by employing the Gilbert's algorithm for minimizing a quadratic form over a convex set [22], which is a widely used algorithm in the fields of Optimal Control, collision detection and classification problems that employ Principal Component analysis or Support Vector Machines.

Let us begin with some definitions and review of concepts that shall help us along the way, starting with separability in the bipartite and multipartite scenarios.

1.1 SEPARABILITY IN THE BIPARTITE SCENARIO

A quantum state represented by a d-dimensional density matrix [1], in the Hilbert space $\mathcal{H}_d$, is called *separable* if and only if it can be written as a convex combination of pure product states. For example, if the state ρ_{AB} is *pure*[2] then

1 A valid density matrix is unit trace and positive semi-definite.
2 $\mathrm{Tr}(\rho^2) = 1$

$\rho_{AB} = \left|\phi_i^{AB}\right\rangle\!\left\langle\phi_i^{AB}\right|$ is called a product state, and if it is a *mixed* state[3], then it is separable if and only if it can be written as,

$$\rho_{AB} = \sum_{i=1} p_i \left|\phi_i^{AB}\right\rangle\!\left\langle\phi_i^{AB}\right|, \tag{1.1}$$

where $\sum_{i=1} p_i = 1$, and $\left|\phi_i^{AB}\right\rangle = \left|\phi_i^{A}\right\rangle \left|\phi_i^{B}\right\rangle$ are pure product states. Both the definition and detection of separability are the simplest in the bipartite case, and as we shall see, the notion of partial separability comes into play for number of subsystems $N > 2$, for which several results [18, 19] establish a sort of hierarchy.

1.2 NOTIONS OF SEPARABILITY IN MULTIPARTITE SCENARIO

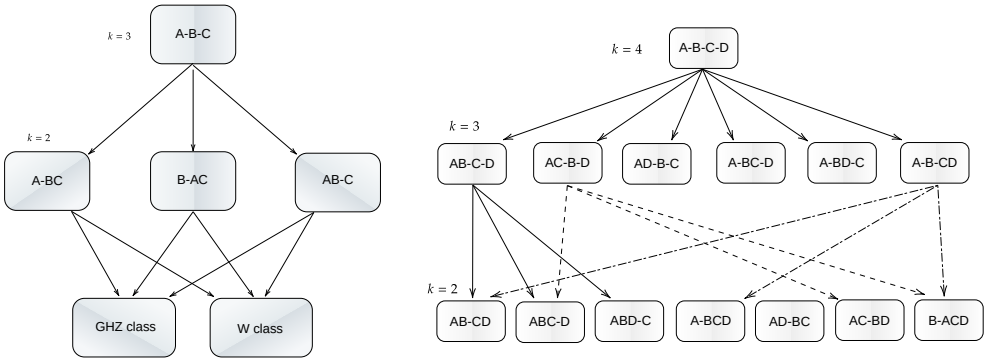
In the multiparty scenario ($N > 2$), a given state can be separable in a few different ways, for example, the state can be written as a tensor product of all the individual subsystems, or two of the subsystems can be entangled while the rest are separable, and so on. To better understand the possibilities that arise, we need to define the concept of *k-separability*. Let us say that the set of subsystems is labeled as $L = \{1, 2, 3, \ldots, N\}$ and $I_1, I_2 \ldots I_k$ are subsets of L such that $I_1 \cup I_2 \cup \ldots I_k = L$ and $I_1 \cap I_2 \cap \ldots I_k = \emptyset$. Such a partition of the subsystems is termed as a *k-partite split*. The state $\rho_{12\ldots N}$ is then called k-separable if it can be written as the mixture,

$$\rho_{12\ldots N} = \sum_i p_i \rho_i^{I_1} \otimes \rho_i^{I_2} \otimes \cdots \rho_i^{I_K}. \tag{1.2}$$

One can also combine separability over different partitions of the set of subsystems and all of them essentially form a part of the levels provided by k-separability. When $k = N$, then the state is called fully separable, and can be written as the tensor product of all the individual subsystems. If $k = 2$, then it is called *biseparability*, which means the state is separable across a particular bipartition of the set L into I_1 and I_2. If we take $N = 3$ as an example then we have three possible bipartitions, $\{I_1 = \{1\}, I_2 = \{2, 3\}\}$, $\{I_1 = \{2\}, I_2 = \{1, 3\}\}$ and $\{I_1 = \{3\}, I_2 = \{1, 2\}\}$. In each case if the state is also separable within I_2 then it becomes fully separable. But as is the case for the 3-qubit bound entan-

3 $\frac{1}{d} \leq \mathrm{Tr}(\rho^2) < 1$, where d is the dimension of ρ

gled states in [9], the fact that they are separable within all the three bipartitions does not guarantee that the state is fully separable.



(a) $N = 3$ (b) $N = 4$

Figure 1.1: Illustrations of the k-separability hierarchy where l-partite splits contain k-partite splits when $l < k$. The direction of the arrow denotes the relation 'contained in'. (a) Shows this hierarchy for $N = 3$ where the full separability guarantees separability on any of the 2-partite splits. (b) In the case of $N = 4$, we have a lot more partitions and interconnections. Note that not all the connections are shown.

In [19] three qubit separability was discussed and was generalized to N-qubits in [18], where a hierarchy is proposed. Considering k- and l-separability for some N qubit state with $l < k$, a l-partite split contains a k-partite split if by joining some of the parties in the k-partite split we obtain the l-partite split. Taking the 3-qubit example again, if $k = 3$ (full separability), then joining any two parties shall give us biseparability, $l = 2$. This is not a one-to-one relation and there are several interconnections between the splits related by permutations of the subsystems. The number of partitions under permutation increases dramatically, with N. Figure 1.1a shows this hierarchy and interconnections for three qubits and Figure 1.1b for four qubits.

It is also important to note, if a state is separable across some k-partite split, then it is also separable across all the l-partite splits that contain it ($l < k$). On the other hand, as we observed for 3-qubit Bound entangled states, it is not necessary that if a state is separable across l-partite splits, then it will also be separable across k-partite splits contained in that l-partite split. In other words, separability with respect to all the l-partite splits containing a particular k-partite split provide a necessary but not a sufficient condition for k-separability with respect to the k-partite split. This implies that if we determine a state to be

separable over some k-partite split, then we do not need to check separability across any l-partite splits that contain it, thus saving effort in classification.

1.3 SEPARABILITY CRITERIA

In the bipartite $d \times d$ case there are several effective ways to check if a state is separable or not. This proves to be the simplest for pure bipartite states. One of the reasons is that it is always possible to write the *Schmidt Decomposition* [36] of a pure state, using which the state can always be written in terms of a product basis $\{|i_A\rangle |i_B\rangle\}$,

$$|\psi_{AB}\rangle = \sum_i \lambda_i |i_A\rangle |i_B\rangle \ , \tag{1.3}$$

where $0 \leq \lambda_i \leq 1$ and $\sum_i \lambda_i^2 = 1$ are the *Schmidt coefficients*. The number of nonzero λ_is is the *Schmidt Rank* (r_s) of the state and can be used to deduce if the state is a product state ($r_s = 1$) or an entangled state ($r_s > 1$). In Equation 1.1, we have already seen how a separable mixed state can be written as a convex combination of pure states that are product vectors themselves. Consequently, a separable pure state has pure reduced states, while an entangled one has mixed reduced density operators for each subsystem. While the number of terms in such a convex combination is not fixed, it can be bounded by using the Caratheodory theorem [52], to be less or equal to the dimension of the total Hilbert space, squared, $\dim(\mathcal{H}_{AB})^2$. We will denote the convex set of separable states by $\mathcal{S}$, with or without subscripts like $\mathcal{S}_{AB}$, according to context, to denote the subsystems it concerns. Only if a state does not belong to the set of separable states, it is said to be *entangled*.

We will now briefly discuss some separability criteria for bipartite states that are discussed in the context of multipartite states as well.

- `Partial Transposition`

  This is one of the most widely used indicators of separability and constitutes a necessary and sufficient condition for the cases of 2×2 and 2×3 dimensions. The criterion can be stated as follows: if ρ_{AB} is a separable state then the new matrix $\rho_{AB}^{T_B}$ obtained by applying the transposition map

on the subsystem B, is also a positive semi-definite matrix. The relation between the matrix elements is given by,

$$(\rho_{AB}^{T_B})_{ij,kl} = (\rho_{AB})_{ij,lk}\,, \tag{1.4}$$

where the indices $\{i, j\}$ and $\{k, l\}$ correspond to the subsystems A and B respectively. We can similarly define the partial transpose over the subsystem A. In [46], it was shown that the partial transpose map can be interpreted as a time reversal in the partially transposed subsystem.

When a density matrix has a positive partial transpose we shall say it is *PPT*. On the other hand, if the partially transposed density matrix has at least one negative eigenvalue then it is called *NPT* (Negative Partial Transpose). While NPT states are definitely entangled, in higher dimensions it is possible to have entangled states that are still PPT. Such states are called *PPT entangled states* (PPTES), or *Bound entangled states*, because their entanglement cannot be used to create maximally entangled state with infinitely many copies of the state. There are several examples of such states in the literature [9, 16].

While the PPT criterion provides a complete characterization of separability in qubit-qubit and qubit-qutrit systems, in higher dimensions one can only conclude if a state is entangled or bound entangled. As it provides a simple condition to check for a given state computationally, and is unable to provide a full picture in higher dimensions, it has to be combined with one or more of the separability criteria available.

- `Positive but not Completely Positive maps`

  The PPT criterion can be written in terms of the linear map $\mathbb{1} \otimes T_B$ or $T_A \otimes \mathbb{1}$, where T_X is the transposition map on the subsystem X. The condition that a state is PPT and thus separable is equivalent to requiring the linear map $(\mathbb{1} \otimes T_B)\rho_{AB}$ to be positive. While the transposition map is positive (i. e., non-negative spectrum), it is not Completely Positive (CP, i. e., strictly positive spectrum) as the map $\mathbb{1} \otimes T_B$ can be shown to be not positive (whereby comes the power of the PPT criterion). Such maps are called *Positive not Completely Positive* maps or PnCP maps. We can state the separability criterion using PnCP maps as follows: a state ρ_{AB} is separable if and only if $(\mathbb{1} \otimes \Lambda_B)\rho_{AB}) \geq 0$ for all PnCP maps Λ_B.

Thus the separability problem is equivalent to the characterization of the set of all PnCP maps. The characterization of the set of all PnCP maps is again a hard problem in linear algebra as well as computationally, and remains unsolved.

- `Reduction Criterion`

  The Reduction Criterion is a special case of the PnCP maps, defined as $\Lambda^{\text{red}}(\rho) = \mathbb{1}\,\text{Tr}(\rho) - \rho$. The separability condition for the state ρ_{AB} is defined as $(\mathbb{1}_A \otimes \Lambda_B^{\text{red}})(\rho_{AB}) \geq 0$ or equivalently $\rho_A \otimes \mathbb{1}_B - \rho_{AB} \geq 0$.

  In the 2×2 case this condition is equivalent to the PPT criterion, and in higher dimensions forms a weaker condition than the PPT criterion. This is an example of a decomposable map. A positive map is called decomposable if it can be written in the form:

$$\Lambda_{dec}^{P} = \Lambda_1^{CP} + \Lambda_2^{CP} \circ T, \tag{1.5}$$

  where T is the transposition map. If this is not the case then the map is called *non-decomposable*. A decomposable map only detects entanglement if the transposition map is able to detect it [24]. Therefore, to detect PPT entanglement one needs to employ non-decomposable maps.

- `Range Criterion`

  In [25], a criterion was formulated to detect the entanglement of a class of PPT states, called the *Range criterion*. If ρ_{AB} is separable then there exist the sets of product vectors $\{|\psi_A\rangle_i \otimes |\psi_B\rangle_i\}$ and $\{|\psi_A\rangle_i \otimes |\psi_B^*\rangle_i\}$ that span the range of ρ_{AB} and $\rho_{AB}^{T_B}$ respectively, where ψ^* indicates the complex conjugate.

  An interesting application of the Range criterion is the formulation of the *Unextendible Product Basis*, the complement space of which does not contain any product states, thus leading to Bound Entangled states (see Section 4.6.1 for more information). Some formulations to quantify entanglement and ascertain separability of given states that use the Range criterion are discussed later in the chapter.

- `Matrix Realignment criterion or Computable Cross Norm`

This is an operational criterion that can be written as a map R that reshuffles the matrix entries [13, 44, 45],

$$(R(\rho_{AB}))_{ij,kl} = (\rho_{AB})_{il,jk} \tag{1.6}$$

such that the *Trace norm*[4], $\|R(\rho_{AB})\|_1 \leq 1$.

The general form of such a condition is comprised of *contraction conditions* defined as follows: if for all pure product states $|\psi_A\rangle |\psi_B\rangle$,

$$\|\Lambda_{con}(|\psi_A\rangle |\psi_B\rangle \langle\psi_A| \langle\psi_B|)\|_1 \leq 1. \tag{1.7}$$

then for any separable state ρ_{AB}, one has $\|\Lambda_{con}(\rho_{AB})\|_1 \leq 1$.

The realignment criterion satisfies such a contraction condition over the set of pure product states and is useful to detect certain classes of PPT entangled states but is limited in it's application.

- `Entanglement Witnesses`

  The origin of Entanglement Witnesses (EW) lies in the geometry of convex sets. An entanglement witness is essentially a hyperplane separating a given entangled state from the set of separable states. Entanglement Witnesses take the form of hermitian operators that have at least one negative eigenvalue while having a non-negative expectation value over the set of separable states. If W is an EW for a state ρ and $\mathrm{Tr}(W\sigma) \geq 0$ for all $\sigma \in \mathcal{S}$ (the set of separable states), then ρ is entangled only if $\mathrm{Tr}(W\rho) < 0$. In higher dimensional Hilbert spaces, it is even possible to detect PPT entanglement using EWs, which makes them quite versatile in any given dimension. See Figure 1.2 for an illustration. A more in depth discussion about Entanglement Witnesses and their connection with other entanglement measures and separability criterion are discussed in Chapter 4.

  Due to the fact that the set of separable states does not form a polytope, the characterization of the shape and boundary of the set in arbitrary dimensions is extremely hard. Entanglement Witnesses that are *optimal* in the sense that the hyperplane representing the witness is also a supporting hyperplane of the convex set of separable states provide a way to characterize the set $\mathcal{S}$.

4 Trace norm of X is given by $\|X\|_1 = \mathrm{Tr}\left(\sqrt{X^\dagger X}\right)$

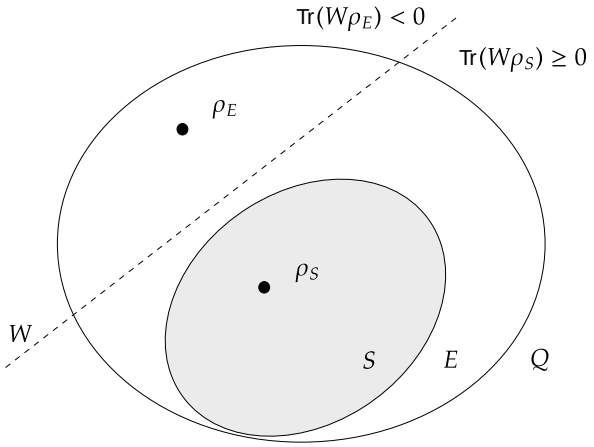


Figure 1.2: Illustration of the Entanglement Witness W as a separating hyperplane. W witnesses the entanglement of an entangled state ρ_E as well as the states in it's neighborhood.

Again, the problem of finding an optimal Entanglement Witness is an optimization problem, that is not solvable efficiently (polynomial complexity in the dimension of the problem).

1.4 ENTANGLEMENT MEASURES

The complementary problem to that of detecting separability is that of detecting and quantifying Entanglement, and this field of research has indeed received a lot of attention for a long time and in the following section we will discuss some of the measures of Entanglement.

1.4.1 *Desirable Properties for an Entanglement Measure*

An entanglement Measure is a function of the state that outputs a real value as a measure of entanglement in the given state. The desired properties that an entanglement measure should have for it to be useful are listed below. We denote a general Entanglement measure as $E(\rho)$.

- If the state ρ is separable, $E(\rho) = 0$, otherwise $E(\rho) > 0$.

- $E(\rho)$ should be invariant under local unitary transformations,

$$E(\rho_{12\cdots N}) = E(U_1 \otimes U_2 \cdots \otimes U_N \rho U_1^\dagger \otimes U_2^\dagger \cdots \otimes U_N^\dagger)\,. \tag{1.8}$$

- Local operations and Classical communication between subsystems does not increase $E(\rho)$.

$$E(\Lambda_{LOCC}(\rho)) \leq E(\rho), \quad \text{for any LOCC map } \Lambda_{LOCC} \tag{1.9}$$

- $E(\rho)$ is convex, i. e., on mixing two or more states the entanglement decreases.

$$E(\sum_i p_i \varrho_i) \leq \sum_i p_i E(\varrho_i) \tag{1.10}$$

- (optional) A particularly demanding condition for an entanglement measure is additivity, so that if two parties share between them 2 different states ρ_1 and ρ_2 then

$$E(\rho_1 \otimes \rho_2) = E(\rho_1) + E(\rho_2) \tag{1.11}$$

  or if there are n copies of the same state, then we have a slightly relaxed condition,

$$E(\rho^{\otimes n}) = nE(\rho). \tag{1.12}$$

1.4.2 *Convex Roof Extension*

Once an entanglement measure has been defined for pure states with all the properties mentioned above, then the next step is to extend the definition to mixed states. A method of accomplishing this is called the *convex roof extension* of the entanglement measure. If an entanglement measure $E(|\psi\rangle)$ has been defined, then its definition is extended to mixed states as,

$$E(\rho) = \inf_{p_i, |\psi_i\rangle} \sum_i p_i E(|\psi_i\rangle) \tag{1.13}$$

where the infimum is over all possible pure state decompositions of the state ρ, so that $E(\rho) \leq \sum_i p_i |\psi_i\rangle$. The resulting quantity by definition follows the

desired property of the entanglement measures. It is quite apparent that such a decomposition is extremely hard to find except in some special cases, where the properties of the state might ease the complexity.

1.4.3 *Examples*

Let us look at some examples of entanglement measures for the bipartite and multipartite cases.

- `Entanglement of Formation`

  Entanglement of formation $E_F(\rho)$, a bipartite measure, is defined as the convex roof of *von Neumann entropy*[5], denoted $S(\rho)$,

  $$E_F(\rho) = \inf_{p_i, \left|\psi_i^{AB}\right\rangle} \sum_i p_i S\left(\mathrm{Tr}_B\left(\left|\psi_i^{AB}\right\rangle\left\langle\psi_i^{AB}\right|\right)\right), \tag{1.14}$$

  where Tr_B denotes partial trace over the subsystem B.

  Entanglement of Formation of a state ρ can be interpreted as the minimal number of singlet states ($\frac{1}{\sqrt{2}}(|00\rangle - |11\rangle)$) that are required to build a copy of ρ. As we already saw, the optimization over all pure state decompositions here presents a difficult problem.

- `Concurrence`

  Concurrence is a bipartite measure defined for pure states as

  $$C(\left|\psi^{AB}\right\rangle) = \sqrt{2 - 2\,\mathrm{Tr}\left[\left(\mathrm{Tr}_B\left(|\psi^{AB}\rangle\langle\psi^{AB}|\right)\right)^2\right]}. \tag{1.15}$$

  When extended to mixed states by convex roof extension it is possible to compute analytically:

  $$C(\rho) = \max\{0, \lambda_1 - \sum_{i=2}^{4} \lambda_i\}, \tag{1.16}$$

  where λ_i are the eigenvalues, in descending order, of the matrix,

  $$X = \sqrt{\sqrt{\rho}(\sigma_y \otimes \sigma_y)\rho^*(\sigma_y \otimes \sigma_y)\sqrt{\rho}} \tag{1.17}$$

5 von Neumann entropy of a state ρ is $S(\rho) = \mathrm{Tr}(\rho \log(\rho))$.

Concurrence and Entanglement of formation are connected by the following relation,

$$E_F(\rho) = H_b\left(0.5 + \frac{\sqrt{1 - C^2(\rho)}}{2}\right) \tag{1.18}$$

where $H_b(x) = -x\log(x) - (1-x)\log(1-x)$ is the binary entropy function.

- `Negativity`

  The next measure, Negativity of a state ρ [54], quantifies the violation of the PPT criterion by the state ρ, and is defined as the sum of the negative eigenvalues of the partial transposition,

  $$N(\rho) = \frac{1}{2}(\left\|\rho^{T_B}\right\|_1 - 1) \tag{1.19}$$

  The advantages of Negativity are that it is convex and very easy to compute. It can even be made additive by defining the logarithmic negativity, $E_N(\rho) = \log_2 \left\|\rho^{T_B}\right\|_1$. Although, logarithmic negativity loses the convexity of Negativity, it can be used to give an upper bound on the Entanglement of Distillation.

- `Distillable Entanglement`

  Distillable Entanglement, $E_D(\rho)$, is defined as the number of singlets that can be distilled out of a large number of copies of the state ρ using only Local Operations and Classical Communication (LOCC). For pure states, it can be written in terms of the von Neumann entropy of the reduced state $\rho_A = \mathrm{Tr}_B(\left|\psi^{AB}\right\rangle)$, and is bounded on the above by logarithmic Negativity.

  $$E_D\left(\left|\psi^{AB}\right\rangle\right) = S(\rho_A) = -Tr(\rho_A \log_2(\rho_A)) \leq E_N\left(\left|\psi^{AB}\right\rangle\right). \tag{1.20}$$

- `Geometric Measure of Entanglement`

Geometric measure of entanglement E_{GM} quantifies the distance of the given state from the set of separable states in a way. For a pure state and the extension to mixed state are defined as,

$$E_{GM}(|\psi\rangle) = 1 - \max_{|\phi\rangle \in \mathcal{S}} |\langle\phi|\psi\rangle|^2, \quad E_{GM}(\rho) = \min_{p_i, |\psi\rangle_i} \sum_i p_i E_{GM}(|\psi\rangle_i)\,. \tag{1.21}$$

In the bipartite case and multipartite case, if the distance is minimized from the set of fully separable states then we have E_{GM} as discussed. However, if in the multipartite case, the distance is minimized over all biseparable pure states, it is called the *Generalized Geometric Measure* E_{GGM},

$$E_{GGM}(|\psi\rangle) = 1 - \max_{|\phi\rangle \in \mathcal{S}_{bi}} |\langle\phi|\psi\rangle|^2, \quad E_{GGM}(\rho) = \min_{p_i, |\phi\rangle_i} \sum_i p_i E_{GGM}(\phi_i)\,, \tag{1.22}$$

where $\mathcal{S}_{bi}$ is the set of biseparable pure states. From the definition of these measures for pure states, a geometric interpretation becomes clear that is the minimization over the pure product states is finding the state $|\phi\rangle_i$ that minimize the sine of the angle between the state $|\psi\rangle$.

- `Other Distance measures`

  There are some more examples of distance measures that quantify the entanglement based on the distance of the given state from the set of separable states. In general, the distance measures are defined with a minimization over the set of separable states. Even though a separable state in some Hilbert space is easy to parameterize, such a minimization has exponential complexity in the dimension of the Hilbert space. Notably, we have the *Relative Entropy of Entanglement*,

$$E_R(\rho) = \min_{\sigma \in \mathcal{S}} S(\rho \| \sigma)\,, \tag{1.23}$$

  where $S(\rho\|\sigma) = \mathrm{Tr}(\rho \log \rho) - \mathrm{Tr}(\rho \log(\sigma))$ is the relative entropy of ρ with respect to σ, and the minimum is taken over the set of separable states [52].

Another notable example is the *Robustness of Entanglement*[55], defined for a state ρ as the minimum value of t such that

$$\rho^+ = \frac{\rho + t\rho^-}{1+t}, \tag{1.24}$$

is a separable state. It is the minimum weight of a separable state, ρ^-, which when mixed with ρ, destroys all quantum correlation in ρ. If the definition is restated to take ρ^- as any arbitrary state in the Hilbert space (separable or entangled), then the minimal weight t is called *Generalized Robustness of Entanglement*[48]. On the other hand, if one looks at how much white noise, $\rho^- = \mathbb{1}$, can be mixed before all the correlations are lost, then the weight is called *Random Robustness*.

Another example is the *Hilbert-Schmidt distance* which is central to this thesis. While not considered an entanglement measure, is nevertheless a good entanglement quantifier as we shall see in the coming chapters [38]. This distance between two states is based on the Hilbert-Schmidt Norm, or equivalently the Frobenius norm of a state

$$\|\rho\|_2 = \sqrt{\mathrm{Tr}(\rho^\dagger \rho)},$$

so that the distance between states ρ and σ is then,

$$D_{\mathrm{HS}}(\rho, \sigma) = \|\rho - \sigma\|_2 = \sqrt{\mathrm{Tr}[(\rho - \sigma)^2]}. \tag{1.25}$$

When the Hilbert-Schmidt distance of a given state ρ is minimized over all $\sigma \in \mathcal{S}$, it is possible to quantify the entanglement of the state,

$$D_{\mathrm{HS}_{\min}}(\rho) = \min_{\sigma \in \mathcal{S}} D_{\mathrm{HS}}(\rho, \sigma). \tag{1.26}$$

Intuition behind this is that the minimum distance $D_{\mathrm{HS}_{\min}}(\rho) = 0$ only if ρ is separable. For further discussion on Hilbert-Schmidt distance see Section 2.2.

1.5 ANALYSING ALGORITHMS FOR SEPARABILITY

Apart form Separability criteria and Entanglement measures that are functions of the given state, there are then a few examples of constructive algorithms that take as input the state on which we need to decide separability, and output, in addition to the classification as separable or entangled, a quantifier associated with the state. While we discuss a few examples of such algorithms, for example, the method of constructing the *Best Separable Approximation*, PPT symmetric extensions, and dual problem of the PPT symmetric extensions, it is important to understand that most studied cases are that of the bipartite case, specifically two qubit systems, while stating that the measure or criteria or algorithm can be generalised to higher dimensions easily. Although the generalisation might be easy, the runtime and complexity always scale exponentially in the dimension of the system. The reason for this becomes clear when we consider the naive approach of searching through the parameterized space of separable density matrices to minimize a function(/measure/criteria).

1.5.1 *Analysing the naive approach*

Let us say that the goal is to find if a state is separable or not, then one can try to find a convex decomposition of the state with the number of terms in the decomposition bounded by $\dim\mathcal{H}^2$ from Caratheodory's theorem. If such a decomposition exists, the answer is 'Separable'. Another solution is to minimize the Relative Entropy of Entanglement, Equation 1.23 over the set of separable states [52] by introducing a parameterization of the mixed two-qubit states,

$$\rho = \sum_{i=1}^{16} p_i^2 \rho_1^i \otimes \rho_2^i \tag{1.27}$$

where $\rho_k^i = \left|\psi_k^i\right\rangle\!\left\langle\psi_k^i\right|$ and $p_i = \sin\phi_{i-1}\prod_{j=1}^{15}\cos\phi_j$ with $\phi_0 = \pi/2$. The pure states are defined to be,

$$\begin{aligned}\left|\psi_1^i\right\rangle &= \cos\alpha_i\left|0\right\rangle + \sin\alpha_i e^{\imath\eta_i}\left|1\right\rangle \quad \text{and,}\\ \left|\psi_2^i\right\rangle &= \cos\beta_i\left|0\right\rangle + \sin\beta_i e^{\imath\mu_i}\left|1\right\rangle\ .\end{aligned} \tag{1.28}$$

N	2	3	4	5	6
$d = 2$	79	575	4351	33791	266239
$d = 3$	809	20411	538001	14407955	387951929

Table 1.1: The number of parameters for a mixed state following the parameterization from Equation 1.27 and Equation 1.28 grows exponentially with the dimension of the subsystem d and the number of such subsystems N.

This makes the space of parameters $\{\alpha_i, \beta_i, \phi_i, \eta_i, \mu_i\}$ periodic as they can only attain values in the interval $[0, 2\pi]$.

In just the two qubit case described here, there are 16 terms in the decomposition, 4 parameters do describe each pure product state in the decomposition and 15 parameters from the coefficients of the decomposition, for a total of 79 parameters.

Number of parameters is 15 and not 16 because $\sum_{i=1}^{16} p_i = 1$ fixes one coefficient.

Calculating the minimum relative entropy over the set of separable states becomes a search over this parameter space and different algorithms can be applied. One can search randomly, but even in the two qubit case, searching over 79 parameters is not an easy task. A better approach, if the problem allows it, is to calculate the gradient of the function that we are trying to minimize in terms of the parameters.

There are many well defined and tested implementations and variations of the *Gradient Descent method*, where fundamentally all of them calculate the gradient and move in the direction that decreases the function value. Although efficient and fast the basic gradient descent method can get stuck in a local minimum very easily, but there are implementations that deal with this in different manners, one of which is using a bunch of random starting points and then choosing the minimum answer found from all the runs.

However, even a very efficient implementation of gradient descent method will have difficulty in successfully finding the global minimum when the parameter space increases exponentially. The parameterization scheme discussed above gives the total number of parameters for an N-qu$d$it, $d^{2N} \times d^N + (d^{2N} - 1)$. See Table 1.1 for the trend. Quite clearly even for 3-qubits, the dimension of the parameter space is too high to apply naive minimization algorithms and expect them to find the global minimum.

*A **qudit** is a d-dimensional system.*

A partial solution to the problem is designing algorithms for checking one-way criteria. For example, the algorithm stops if the state is Entangled, but

cannot confirm if the sate is separable and vice versa. A straightforward example of such a one way criteria is the PPT criterion in higher dimensions, a state being PPT does not imply separability. Another example is the PPT symmetric extension [17], which we shall discuss below with its dual problem [26].

1.5.2 *Improving on the naive approach*

The algorithm presented in [26] for bipartite states, instead of searching the whole space like in the naive approach, constructs a countable set of product vectors that is dense in the set of separable states, which already reduces the set of feasible solutions to a countable number. The algorithm follows the ensuing steps:

1. Constructs a countable set of product vectors, C, that is dense in the set of separable states.
2. The set C is then divided into tuples of a fixed length $l = (\dim \mathcal{H}_A \times \dim \mathcal{H}_B)^2$.
3. (`$i^{th}$-iteration`) Check if the i^{th} l-tuple is linearly independent, if not move to the next l-tuple.
4. Check if the given state is in the convex hull of the product vectors from l-tuple.
5. If the state is in the convex hull of the vectors, then the algorithm terminates, with the answer 'Separable'. But if the state is not inside the convex hull, then move on to the next tuple.

The algorithm works using a slight redefinition of 'Separability': A state is separable if and only if there exists a set of pure product vectors such that the state lies in the convex hull of the elements of this set. By not using the probabilities that form the coefficients in the pure state decomposition, the dimension of the parameter space is reduced, but it still has exponential complexity in the dimension of the total Hilbert space.

The reduction in the dimension is $d^{2N} - 1$.

The algorithm provides a one-way condition by only terminating if the state is separable. If the state is separable then it is guaranteed to stop in a finite number of steps. As such it was proposed as the dual of the algorithm in the

next part, which provides a one-way condition that terminates in a finite number of steps only if the state is entangled. By running the two in parallel over a given state, we are assured of a classifying answer in a finite amount of time and number of steps.

1.5.3 *Semi-definite Programs and PPT symmetric extension*

A popular optimization method in the recent time is *Semi-Definite Programming* (SDP), which is a special case of convex optimization paradigm. A Semi-definite program optimizes a linear function under linear matrix inequality constraints [17]. The usual form of SDPs is,

$$\begin{aligned} &\text{minimize} && c^T\mathbf{x} \\ &\text{subject to} && M(\mathbf{x}) \geq 0, \end{aligned} \tag{1.29}$$

where c is a given constant vector, $\mathbf{x}$ is the optimization variable, $M(\mathbf{x}) = M_0 + \sum_i x_i M_i$ is a linear combination of Hermitian matrices M_i. The constraint in Equation 1.29 requires positive semi-definiteness of $M(\mathbf{x})$. Due to their convex nature, it has been possible to develop efficient numerical methods for solving SDPs. Consequently, any problem that can be expressed as an SDP or a series of SDPs, called an SDP hierarchy, such that on each level the constraints increase in tightness, then an efficient solution is possible. At least in the lower dimensions 2×2 and 2×3 cases, where the equivalence of the set of PPT states and the set of separable states again makes it simpler to solve the optimization problem using an SDP efficiently. An SDP hierarchy is said to be complete if it can be shown to definitely succeed in achieving it's goal at some finite level of the hierarchy.

When trying to reduce a problem like determining if the state is separable or not to an SDP in the higher dimensions, due to the constraints being positive semi-definite, one can at most demand that the partial transpose is positive, which as a consequence only allows us to distinguish between the set of PPT states from the set of entangled states. This is the basis of the algorithm for checking *PPT symmetric extensions* using SDP. It provides an infinite hierarchy of SDPs that are complete, and will always decide if the state is entangled in a finite number of steps.

A *symmetric extension* of a state $\rho = \sum_i p_i |\psi_i\rangle\langle\psi_i| \otimes |\phi_i\rangle\langle\phi_i|$ on the Hilbert space $\mathcal{H}_A \otimes \mathcal{H}_B$ is defined as the state,

$$\tilde{\rho} = \sum_i p_i |\psi_i\rangle\langle\psi_i| \otimes |\phi_i\rangle\langle\phi_i| \otimes |\psi_i\rangle\langle\psi_i| \,, \tag{1.30}$$

which is in turn defined on the Hilbert space $\mathcal{H}_A \otimes \mathcal{H}_B \otimes \mathcal{H}_A$, such that $\mathrm{Tr}_C(\tilde{\rho}) = \rho$ (C is the copy of A), and $\tilde{\rho}$ is symmetric under exchange of the copies of $\mathcal{H}_A$ is called the symmetric extension of ρ.

This definition can be generalized to arbitrary number of copies of $\mathcal{H}_A$ or $\mathcal{H}_B$. The number of such copies defines the level in the SDP hierarchy. If at any level, there are k copies of the subsystem A and l copies of the subsystem B, then the density matrix $\tilde{\rho}$ is classified as 'entangled' if it is NPT on any partition of the total $k + l$ parties into two groups. Otherwise, the program moves on to the next level of the hierarchy.

At each level it is enough to check the PPT criterion for partitions of the total Hilbert space not related by permutations of the copies of the subsystems. Therefore, at each level, there are a total of $\lceil (k+1)(l+1)/2 \rceil$ PPT checks.

The algorithm is again exponential in the dimension of the total Hilbert space and stops only for the entangled states. As suggested in the discussion of the previous algorithm, the two combined together give a complete solution to separability, that unfortunately, quickly becomes intractable in higher dimensional Hilbert spaces.

1.5.4 *Best Separable Approximation*

First presented in [31], construction of the *Best Separable Approximation* (BSA) provides a necessary separability condition based on the Range Criterion. We discuss this here and not earlier, because the construction of BSA is algorithmic as we shall see now.

Firstly, BSA is defined as follows. For any density matrix ρ, and a set of vectors V in the range of ρ, $R(\rho)$, there exists a separable (not normalized) matrix,

$$\rho_s^* = \sum_i \Lambda_i P_i \,, \quad \text{such that} \quad \Lambda_i \geq 0, \quad P_i = |\psi_i\rangle\langle\psi_i| \in V, \tag{1.31}$$

and $\Delta\rho = \rho - \rho_s^* \geq 0$ then the matrix ρ_s^* is called the Best separable approximation of ρ if $\mathrm{Tr}(\Delta\rho)$ is minimal or equivalently $\mathrm{Tr}(\rho_s^*) \leq 1$ is maximal.

Using this definition, a density matrix ρ is separable if and only if there exists a set of product vectors $V \subset R(\rho)$ such that $\mathrm{Tr}(\rho_s^*) \leq 1$ and thus $\mathrm{Tr}(\Delta\rho) = 0$. For the construction, a concept of *maximality* of Λ_is is used. A Λ_i is called *maximal* with respect to ρ and some projection $P_i = |\psi_i\rangle\langle\psi_i|$ if the matrix $\rho - \Lambda_i P_i \geq 0$ and $\rho - (\Lambda_i + \epsilon)P_i < 0$ for any $\epsilon > 0$. Therefore, if the projection $P_i \notin R(\rho)$, then corresponding maximal $\Lambda_i = 0$ and P_i has no contribution in the BSA.

Similarly, the maximality of a pair of Λ_a and Λ_b is also defined with respect to two projectors P_a and P_b in the three cases where one, both or none of the two projectors are in $R(\rho)$.

- If $P_a, P_b \notin R(\rho)$, then $\Lambda_a = \Lambda_b = 0$.
- If $P_a \in R(\rho)$ and $P_b \notin R(\rho)$, then $\Lambda_b = 0$ and Λ_a is maximal. Similarly when $P_a \notin R(\rho)$ and $P_b \in R(\rho)$, then then $\Lambda_a = 0$ and Λ_b is maximal..
- If $P_a, P_b \in R(\rho)$, then Λ_a and Λ_b are maximal if
  - Λ_a is maximal with respect to $\rho - \Lambda_b P_b$ and P_a.
  - Λ_b is maximal with respect to $\rho - \Lambda_a P_a$ and P_b.
  - $\Lambda_a + \Lambda_b$ is maximal under the constraint $\rho - \Lambda_a P_a - \Lambda_b P_b \geq 0$.

The construction of BSA begins with a randomly chosen set product vectors V. Then the matrix,

$$\rho_s^*(V) = \sum_{P_i \in V} \Lambda_i P_i, \tag{1.32}$$

is constructed where all the Λ_is are maximal individually and pairwise with respect to ρ and projectors P_is. If any projector $P_i \notin R(\rho)$ then the corresponding $\Lambda_i = 0$.

In the case of 2 qubits it was shown that the construction of BSA provides an entanglement measure, and it is unique. In higher dimensions, the construction is complicated by the presence of PPT entangled states, but the BSA still remains unique [29]. It is difficult to calculate in higher dimensions, an example of which is in [43] where analytical results were given for the symmetric classes of states that are invariant under permutations of the subsystems. [43].

1.6 SUMMARY

In the chapter, we have discussed the many and varied ways and methods to check separability and quantify entanglement for a given state. While going over these necessary concepts that are required to get the complete context, we also discussed how computation of Entanglement measures for a given state and the checking separability criteria are easy and give a complete picture in the Hilbert spaces of the lower dimensions (2×2 and 2×3). Whereas, they are either too hard to compute or fail to give an answer in the higher dimensions. Even in the bipartite case, the computation of measures for mixed states is a very hard problem.

The *Separability problem* has been shown to be NP-hard, and while there are several algorithms that tackle this problem, none do so efficiently. It is an open question if it is at all possible. A review of such algorithms can be found in [27]. It is also important to note that NP-hardness of the problem does not exclude the possibility of finding efficient solutions of the separability problem in special cases, or classes of quantum states.

In the following chapters we formulate the Gilbert's algorithm to find the closest separable state from a given state in arbitrary dimensional Hilbert space. While this algorithm again scales exponentially with the dimension of the Hilbert space, our algorithm provides a two-way condition as we shall see.

In Chapter 2, we formally define the convex optimization problems that are equivalent to the Separability problem. We will discuss the geometry of the quantum state space under the Hilbert-Schmidt norm and how these optimization problems apply to this state space. The Gilbert's algorithm is described and then we discuss the modifications that simplify the computation based on our work "Hilbert-Schmidt distance and entanglement witnessing" [39].

In Chapter 3, we present the results of applying the algorithm to some well known classes of states. Detailed analysis of the output of the algorithm is done whilst comparing it to the previously known analytical knowledge in some cases. Where there was a lack of previous knowledge we present new findings, trends and analytical insights. Some of these results were already presented in [39], while others, crucially results concerning biseparability and some accompanying geometrical insights are novel results presented in this thesis.

In Chapter 4 we discuss the application of the results from Gilbert's algorithm to form close to optimal Entanglement Witnesses by adding one more step, an

optimization procedure, to the algorithm. Here we present a few results from "Distance between Bound Entangled States from Unextendible Product Bases and Separable States" [58]. Also, new results are presented for some specific classes of states.

Then in Chapter 5 we will look at another way to certify non-local correlations, namely the paradigm of *Self-Testing*. It is a self contained chapter and the relevant notions are described in place. The results presented are a part of our work, "An elegant proof of self-testing for multipartite Bell inequalities" [40], where we have provided a very general framework for Self-testing with minimal assumptions and wide ranging applicability to all two-setting and two-outcome Bell scenarios.

2

GILBERT'S ALGORITHM AND HILBERT-SCHMIDT DISTANCE

In the last chapter we saw the difficulty we face in tackling the problem of separability. We now discuss our approach to decide separability, using the widely used Gilbert's algorithm for minimizing a quadratic function over a convex set [22]. To start we will first define a few optimization problems that are equivalent in this case to solving the Separability problem. The fact that the set of separable states is convex, is central to the formulation of an algorithm to compute any measure or criterion on the set, because the whole set can be characterised using convex combinations of the pure product states which are the extreme points or vertices.

2.1 SEPARATION, OPTIMIZATION AND MINIMUM DISTANCE

First we shall formally define the *Weak Separation* and the *Weak Optimization* problems.

- `Weak Separation` (`WSEP`): Given a point $\mathbf{x} \in \mathbb{R}^N$, and a convex set $\mathcal{S} \subset \mathbb{R}^N$, either find $\mathbf{s} \in \mathcal{S}$ such that $\|\mathbf{x} - \mathbf{s}\| < \delta$ for some $\delta > 0$ or find a vector $\mathbf{c} \in \mathbb{R}^N$ such that $\mathbf{x} \cdot \mathbf{c} > \max_{\mathbf{s} \in \mathcal{S}} \mathbf{c} \cdot \mathbf{s}$.

  If $\delta = 0$, it becomes the *strong* separation problem.

- `Weak Optimization` (`WOPT`): Given a point $\mathbf{x} \in \mathbb{R}^N$, and a convex set $\mathcal{S} \subset \mathbb{R}^N$, find a point $\mathbf{r} \in \mathcal{S}$ such that $\mathbf{x} \cdot \mathbf{r} > \max_{\mathbf{s} \in \mathcal{S}} \mathbf{x} \cdot \mathbf{s} - \delta$.

  Again, if $\delta = 0$, it becomes the *strong* optimization problem.

δ defines the radius of the neighbourhood of the feasible point. Therefore, $\delta = 0$ implies infinite precision in finding the feasible point.

`WSEP` tries to ascertain if the point $\mathbf{x}$ is in the set $\mathcal{S}$ and if not, then it tries to find the separating hyperplane. `WOPT`, on the other hand, tries to find the point that maximizes a linear function. Here δ defines the radius of spherical neighborhood of the optimum point, i. e., the algorithm accepts a point as the *nearly optimum* solution when the point is in the neighborhood defined by δ. It

is clear that the *Membership problem* which tries to find out if a given point is in the convex set can be solved by solving WSEP.

The description of the convex set plays significant role in deciding how well any algorithm that is designed to use the convexity of the set performs over it. For instance, if the set is described as a convex hull of vertices $\mathcal{S} = \text{conv}\{\mathbf{s}_1, \mathbf{s}_2 \dots \mathbf{s}_n\}$, then optimization of any function over the set (i. e., solving WOPT) is just a matter of calculating the function value over the set of vertices. While solving WSEP is significantly more complex, and would require a linear program to solve. Previously we saw in Section 1.5.2, this is the step that is the source of the computational complexity of the separability criterion.

Moreover, solving WSEP is easy when the set is described as a convex polytope, which is the intersection of a finite number of half-spaces specified by the set of inequalities, $\mathcal{S} = \{\mathbf{s} \mid \mathbf{a} \cdot \mathbf{s} \leq b_i\}$. The inequalities defining the convex polytope are all satisfied when the point is inside the set, whereas if any one inequality fails, the point is outside the set and the failed inequality provides the separating hyperplane. With this description of the set, we require a linear program to solve WOPT. The complexity of the linear programs in both cases is exponential in the dimension of the set $\mathcal{S}$.

There are methods to convert WSEP to WOPT and vice versa, depending on which representation of the convex set suits the problem at hand, although this might cause an exponential increase in the size of the problem. In [28], such a reduction is discussed at length. To avoid the aforementioned increase in problem size, we can define the *weak minimum distance* problem [12].

- Weak Minimum Distance (WDIST): Given a point $\mathbf{x} \in \mathbb{R}^n$, find $\mathbf{s} \in \mathcal{S}$ such that $\|\mathbf{x} - \mathbf{s}\| \leq \text{dist}_{\text{min}}(\mathbf{x}, \mathcal{S}) + \delta$ for some $\delta > 0$.

  Here $\text{dist}_{\text{min}}(\mathbf{x}, \mathcal{S})$ is the minimum distance of the point $\mathbf{x}$ from the set $\mathcal{S}$ and if $\delta = 0$, it is the *strong* Minimum Distance problem.

It is clear that if $\mathbf{x}$ is in $\mathcal{S}$, then $\|\mathbf{x} - \mathbf{s}\| \leq \delta$, otherwise, a call to WOPT gives the point $\mathbf{c}$ that separates $\mathbf{x}$ from $\mathcal{S}$.

We will see that the WDIST problem is precisely what the Gilbert's iterative algorithm solves. To adapt and apply the Gilbert algorithm to Separability, we must first discuss the Hilbert-Schmidt norm and the geometry it imposes over the set of quantum states. The quadratic function that is to be minimized over the convex set becomes the Hilbert-Schmidt distance between two density matrices.

2.2 HILBERT-SCHMIDT NORM AND THE GEOMETRIC PICTURE

In finite dimensional Hilbert spaces, the Hilbert-Schmidt norm of a density matrix ρ is equivalent to the Frobenius Norm. The norm is independent of the orthonormal basis used, and therefore has the following equivalent definitions.

$$\|\rho\|_{HS} = \sqrt{\mathrm{Tr}(\rho^\dagger \rho)} \tag{2.1}$$

$$= \sqrt{\sum_i \sum_j |a_{ij}|^2} \qquad \text{(sum of squares of matrix elements)} \tag{2.2}$$

$$= \sqrt{\sum_i \lambda_i^2} \qquad \text{(sum of squares of eigenvalues)} \tag{2.3}$$

The norm then induces the Hilbert-Schmidt distance over the set of quantum states, $D_{\mathrm{HS}}(\rho_1, \rho_2) = \sqrt{\mathrm{Tr}(\rho_1 - \rho_2)^2}$. When we consider the Hilbert space $\mathbb{C}^2$, any density matrix ρ can be written in the Pauli basis,

$$\rho = \frac{\mathbb{1}}{2} + \mathbf{n} \cdot \boldsymbol{\sigma} \tag{2.4}$$

where $\mathbb{1}$ is the 2×2 identity matrix, $\mathbf{n}$ a vector in $\mathbb{R}^3$ and $\boldsymbol{\sigma} = \{\frac{1}{\sqrt{2}}\sigma_x, \frac{1}{\sqrt{2}}\sigma_y, \frac{1}{\sqrt{2}}\sigma_z\}$ is a vector of the normalized Pauli matrices. See Equation a.1 for the definition of Pauli matrices. The vector $\mathbf{n}$ is called the *statevector* of ρ.

Normalized so that the Hilbert-Schmidt Norm, $\sqrt{\mathrm{Tr}(A^\dagger A)}$ equals 1.

The conditions ρ needs to satisfy to be a valid density matrix also constrain $\mathbf{n}$. The hermiticity of ρ ensures $\mathbf{n} \in \mathbb{R}^3$, positivity ($\mathrm{Tr}(\rho^2) \leq 1$) implies the length of the vector $\|\mathbf{n}\| \leq \frac{1}{\sqrt{2}}$. By demanding $\mathrm{Tr}(\rho^2) = 1$, we get the distinction between pure states, $\|\mathbf{n}\| = 1/\sqrt{2}$ and mixed states $\|\mathbf{n}\| < 1/\sqrt{2}$.

Therefore, under the Hilbert-Schmidt norm, the set of all states takes on the geometry of a ball, $\mathbf{B}^3$, called the *Bloch ball*. The pure states lie on the surface (and form the *Bloch sphere*) and the mixed states lie in the interior. At the center of the ball lies the maximally mixed state $\frac{\mathbb{1}}{2}$, and all the pure states lie at an equal Hilbert-Schmidt distance of $1/\sqrt{2}$ from the center (Figure 2.1).

This picture can be generalized for N d-dimensional systems,

$$\rho = \frac{\mathbb{1}}{D} + \mathbf{n} \cdot \boldsymbol{\Lambda}, \tag{2.5}$$

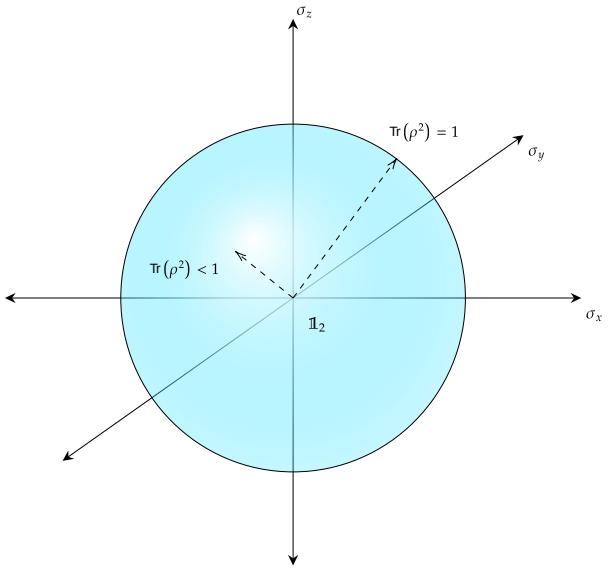


Figure 2.1: Bloch ball representation of a qubit. Every vector inside and on the surface corresponds to a quantum state. Pure states lie on the surface and mixed states lie in the interior.

where $D = d^N$ is the dimension of the the density matrix and now the statevector $\mathbf{n} \in \mathbb{R}^{D^2-1}$ so that the set of all states take on the geometry of a hyperball, $\mathbf{B}^{D^2-1}$ of radius $\sqrt{\frac{(D-1)}{D}}$.

For all pure states $\|\mathbf{n}\| = \sqrt{\frac{(D-1)}{D}}$.

In the general case, the vector $\mathbf{\Lambda}$ is comprised of the (normalized) generalized Gell-Mann Matrices. The hyperball is sometimes called generalized Bloch ball. See Section a.1.2 for a way to calculate the generalized Gell-Mann matrices for arbitrary dimensions.

Considering the the Hilbert-Schmidt distance of two density matrices,

$$\rho = \mathbb{1}_D + \mathbf{n}_\rho \cdot \mathbf{\Lambda} \tag{2.6}$$

$$\sigma = \mathbb{1}_D + \mathbf{n}_\sigma \cdot \mathbf{\Lambda}\,, \tag{2.7}$$

we find $D_{\mathrm{HS}}(\rho, \sigma) = D_E(\mathbf{n}_\rho, \mathbf{n}_\sigma)$, where D_E is the Euclidean distance between the two vectors. Moreover, in the case of $\mathbb{C}^2$, every point in the Bloch ball corresponds to a valid density matrix, while for $D > 2$, this is not true and the set of all quantum states is a subset of the generalized Bloch ball [62].

If we now turn to look at the set of separable states, $\mathcal{S}$, we find it is neither a ball or a convex polytope. The infinite extreme points of the set, i. e., pure product vectors all lie on the hyper-surface of the hyperball and all the convex

combinations reside in the interior. The hyper-surface of $\mathcal{S}$ is then curved where it coincides with the surface of the hyperball and has faces lying in the interior. Thus, it is neither possible to define the set as a convex hull of vertices nor as an intersection of a finite number of half-spaces, wherein lies the difficulty in formulating algorithms for `WSEP`, `WOPT` and `WDIST` problems.

*For a D dimensional convex set, the faces with dimension $D-1$ are called **facets**. The set $\mathcal{S}$ has no facets.*

Armed with this picture we take on the task of applying the Gilbert's algorithm to the Separability problem.

2.3 GILBERT'S ALGORITHM

Proposed originally in 1966, the Gilbert's algorithm is widely used in its many variations in the fields of Optimal control, Classification using Principal Component Analysis and Support Vector Machines, Collision detection and related areas. Its popularity as a real-time algorithm is due to its low computational requirements, effectiveness and a guarantee of convergence.

The original algorithm is an iterative method that attempts to minimize a quadratic function over a given convex set. Gilbert in [22] used the minimum norm problem to prove the convergence of the algorithm and then showed that any quadratic form to be minimized can be reduced to the minimum norm problem. Let's define the minimum norm problem.

- `Weak Minimum Norm (WMNORM)` Given a convex set $\mathcal{K} \in \mathbb{R}^n$, find $\mathbf{z}_{\min} \in \mathcal{K}$ such that $\|\mathbf{z}_{\min}\| = \min_{\mathbf{z}\in\mathcal{K}} \|\mathbf{z}\| + \delta$ for some $\delta > 0$.

  With $\delta = 0$, we have the *strong* Minimum Norm problem.

The algorithm generates a sequence of points $\{\mathbf{z}_k\}$, employing a *contact function* $s(-\mathbf{z}_k)$ for the set $\mathcal{K}$. The contact function, $s(-\mathbf{z}_k)$, outputs the point on the boundary of $\mathcal{K}$ such that the supporting hyperplane at the point has normal $-\hat{\mathbf{z}}_k$. This is to ensure that next feasible point is towards the origin and such a point will have a lower norm. The contact function in itself comprises an optimization problem (`WOPT`) on $\mathcal{K}$ that requires a linear program. The algorithm can be stated as follows:

`Input`: Convex set $\mathcal{K} \in \mathbb{R}^n$;
`Output`: Sequence $\{z_k\}$ with decreasing norm;

---

1. Choose an arbitrary $\mathbf{z}_0 \in \mathcal{K}$.

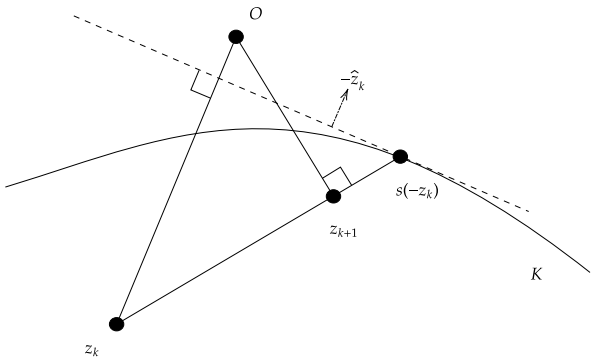


Figure 2.2: An iteration of Gilbert's algorithm with current best point $\mathbf{z}_k$, for which we evaluate $s(-\mathbf{z}_k)$ to obtain the point on the boundary of $\mathcal{K}$ such that if a hyperplane is tangent at this point, it has the normal $-\hat{\mathbf{z}}_k$. We obtain $\mathbf{z}_{k+1}$ by projecting $-\mathbf{z}_k$ onto $s(-\mathbf{z}_k) - \mathbf{z}_k$, which definitely has the minimum norm on the line joining $s(-\mathbf{z}_k)$ and $\mathbf{z}_k$.

2. (`$k^{\text{th}}$-iteration`) With $\mathbf{z}_k$ obtain point $s(-\mathbf{z}_k)$.

3. If $\mathbf{z}_k = s(-\mathbf{z}_k)$, we have found the optimum, else continue.

4. Update $\mathbf{z}_{k+1} \leftarrow \mathbf{z}_k + \epsilon(s(-\mathbf{z}_k) - \mathbf{z}_k)$, where

$$\epsilon = \min\left\{\frac{\mathbf{z}_k \cdot (\mathbf{z}_k - s(-\mathbf{z}_k))}{\|\mathbf{z}_k - s(-\mathbf{z}_k)\|^2}, 1\right\}. \tag{2.8}$$

   gives the point with minimum norm on the line $s(-\mathbf{z}_k) - \mathbf{z}_k$.

5. Check *HALT* condition. If *FALSE*, go to step 2. If *TRUE*, exit.

Figure 2.2 illustrates the k^{th}-iteration where from $\mathbf{z}_k$ we obtain $\mathbf{z}_{k+1}$ by projecting the vector $-\mathbf{z}_k$ onto the vector $s(-\mathbf{z}_k) - \mathbf{z}_k$. The intuition behind the min function in Equation 2.8 is that if $s(-\mathbf{z}_k)$ is close enough to the vector $\mathbf{z}_k$ then the projection of $-\mathbf{z}_k$ on $s(-\mathbf{z}_k) - \mathbf{z}_k$ might lie outside the set $\mathcal{K}$, in which case we update $\mathbf{z}_{k+1} = s(-\mathbf{z}_k)$. The HALT condition can be a limit on the number of iterations, a time constraint or required precision. In the case where the origin is in $\mathcal{K}$, the optimum point $\mathbf{z}_{\min}$ is obtained as the origin, otherwise Gilbert proved that the optimum point will lie on the boundary of the set $\mathcal{K}$. In other words, the algorithm finds the point closest to the origin.

It is easy to see that, in this case, if we wanted to minimize the distance of a given point $\mathbf{z}$ from the set $\mathcal{K}$, we could translate the set by $-\mathbf{z}$ and the WDIST problem is reduced to the minimum norm problem.

WDIST reduces to WMNORM.

2.4 ADAPTING THE GILBERT'S ALGORITHM

In this section we will discuss the application of the Gilbert's algorithm to minimize the Hilbert-Schmidt distance over the convex set of separable states. The main hurdle in using the Gilbert's algorithm as is, turns out to be the optimization step: calculating $s(-\mathbf{z}_k)$, which requires a linear program exponential in the dimension of the Hilbert space.

Let us lay down the parameters of the WDIST (equivalently, minimum norm) problem that we would like to solve. Our goal is to find the *Closest Separable State* (CSS) to a given state ρ_0, and in the process, classify ρ_0 as separable or entangled. We will do this by minimizing the squared Hilbert-Schmidt distance $D^2_{\text{HS}}(\rho_0, \sigma)$ for all $\sigma \in \mathcal{S}$. We can now define the problem of finding the Closest Separable State with respect to the Hilbert-Schmidt distance formally:

- `Weak Closest Separable state (WCSS)`: Given a state ρ_0, find $\rho_{\text{CSS}} \in \mathcal{S}$ such that $D^2_{\text{HS}}(\rho_0, \rho_{\text{CSS}}) = \min_{\sigma \in \mathcal{S}} D^2_{\text{HS}}(\rho_0, \sigma) + \delta$ for some $\delta > 0$.

  The corresponding strong optimization problem is defined with $\delta = 0$.

$\delta = 0$ finds the exact CSS.

We continue to define the optimization problem with non-zero δ to emphasize the point that we cannot achieve infinite precision ($\delta = 0$) without adding optimization procedures, which causes the computational complexity to significantly increase. For this reason, the simplified Gilbert's algorithm gives us a close approximation of the CSS, and as the number of iterations increases, the approximation moves closer still to the CSS.

We can now define the quantity minimum Hilbert-Schmidt distance from the set of separable states for a given state ρ_0,

$$D^2_{\text{HS}_{\text{min}}}(\rho_0) = \min_{\sigma \in \mathcal{S}} D^2_{\text{HS}}(\rho_0, \sigma) = D^2_{\text{HS}}(\rho_0, \rho_{\text{CSS}}), \tag{2.9}$$

where ρ_{CSS} is the closest separable state to ρ_0. For all $\rho_0 \in \mathcal{Q}$ (the set of quantum states), the minimum distance $D^2_{\text{HS}_{\text{min}}}(\rho_0) \geq 0$ where the equality holds only if $\rho_0 \in \mathcal{S}$. We can also see that when $\rho_0 \notin \mathcal{S}$, ρ_{CSS} lies on the boundary of $\mathcal{S}$.

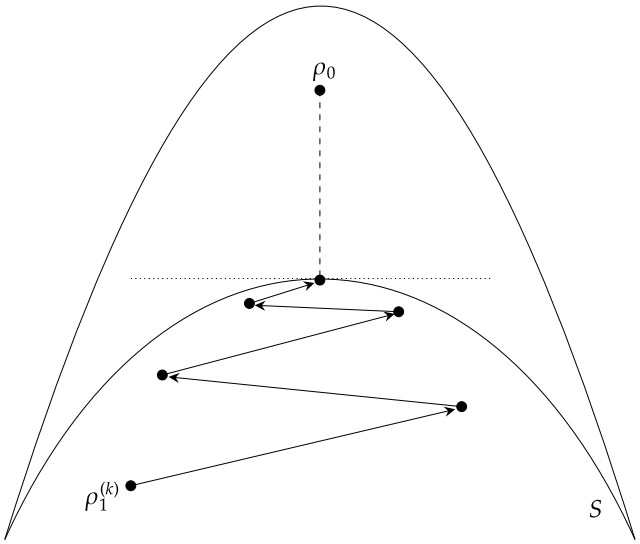


Figure 2.3: An illustration of the iterative process by which the Gilbert's algorithm converges to the CSS.

We accomplish this task by using the *Simplified* Gilbert's Algorithm. The simplification here is omitting the optimization step and replacing it with uniform random sampling from the set of pure product states that form the vertices of the set $\mathcal{S}$, combined with a preselection condition. The preseletion condition ensures that the randomly chosen state will definitely reduce the distance, or it will be rejected.

In each iteration of the algorithm, we have the reference state ρ_0, the current approximation of the CSS, ρ_1, and the random pure separable state generated in this iteration ρ_2. The preselection condition, for ρ_2 to be viable, can be stated as $\mathrm{Tr}[(\rho_0 - \rho_1)(\rho_2 - \rho_1)] > 0$. The intuition behind this is that the angle made by the line $(\rho_0 - \rho_1)$ with $(\rho_2 - \rho_1)$ has to be acute for the projection of $(\rho_0 - \rho_1)$ on to $(\rho_2 - \rho_1)$ to give us a reduction in the distance from ρ_0. This simple formulation of the preselection condition involves just one matrix multiplication and a constant number of addition and subtraction operations, making it very computationally undemanding. The preselection condition is the reason why the convergence of simplified Gilbert's algorithm towards the optimum is guaranteed.

The point of projection lies on the line joining ρ_1 and ρ_2 and can be expressed as the convex combination $\rho_1(p) \coloneqq p\rho_1 + (1-p)\rho_2$ where $p \in [0,1]$. The optimal value of p that minimizes the distance from ρ_0 can be found by minimiz-

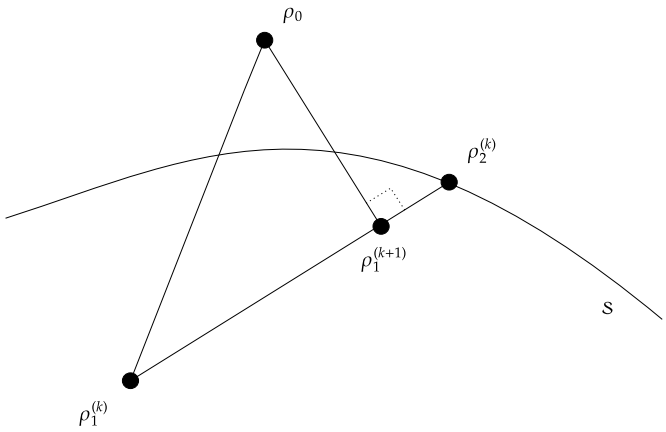


Figure 2.4: An iteration of simplified Gilbert' algorithm where the next feasible point $\rho_1^{(k+1)}$ is the projection of $\rho_0 - \rho_1^{(k)}$ on $\rho_2^k - \rho_1^{(k)}$.

ing $D^2_{\mathrm{HS}}(\rho_0, \rho_1(p))$ with respect to p. As it turns out $D^2_{\mathrm{HS}}(\rho_0, \rho_1(p))$ is just a quadratic function in p,

$$D^2_{\mathrm{HS}}(\rho_0, \rho_1(p)) = \mathrm{Tr}(\rho_0 - \rho_2)^2 + 2\,\mathrm{Tr}[(\rho_0 - \rho_2)(\rho_2 - \rho_1)]p + \mathrm{Tr}(\rho_1 - \rho_2)^2 p^2\,, \tag{2.10}$$

the vertex of which lies at

The vertex of $ax^2 + bx + c$ lies at $x = -\frac{b}{2a}$.

$$p_{\min} = -\frac{\mathrm{Tr}[(\rho_0 - \rho_2)(\rho_2 - \rho_1)]}{\mathrm{Tr}(\rho_1 - \rho_2)^2}. \tag{2.11}$$

so that the new optimum is given by $\rho_1(p_{\min})$. Now that we have the relevant information let us look at the steps of the algorithm. The k^{th}-iteration is also illustrated in the Figure 2.4.

`Input`: ρ_0, ρ_1;
`Output`: Approximate ρ_{CSS}, lists $\{\rho_1^{(k)}\}$, $\{D^2_{\mathrm{HS}}(\rho_0, \rho_1)_k\}$ (optional);

---

1. (`$k^{\text{th}}$-iteration`) Choose a random pure separable state $\rho_2^{(k)}$.
2. Check preselection condition for $\rho_2^{(k)}$, if FALSE, increment c_{rej} and go to step 1 else increment c_s.
3. Calculate $p_{\min}$.

4. Update $\rho_1^{(k+1)} \leftarrow p_{\min}\rho_1^{(k)} + (1 - p_{\min})\rho_2^{(k)}$. Add $\rho_1^{(k+1)}$, and $D^2_{\text{HS}}(\rho_0, \rho_1^{(k+1)})$ to the lists.

5. Check HALT, if TRUE exit, else go to step 1.

Here c_s counts the number of successful iterations, c_{rej} counts the number of rejected states (that fail the preselection) at each iteration. The HALT condition can be checks on one or a combination of any of the following:

- Number of corrections, c_s,
- Number of rejections per iteration, c_{rej},
- Tolerance: $D^2_{\text{HS}}(\rho_0, \rho_1)_{k+1} - D^2_{\text{HS}}(\rho_0, \rho_1)_k$, or
- Time elapsed.

A check on number of corrections bounds the number of successful iterations performed by the algorithm, while a check on number of rejections per iteration terminates the algorithm if it has generated a set number of random pure separable states and all of them failed the preselection condition. This condition is helpful in terminating the algorithm when the algorithm has a very small feasible space compared to the search space, which happens because the feasible space shrinks on every iteration, while the search space remains the same. Geometrically, the feasible space is the half space defined by the hyperplane with the normal $\rho_0 - \rho_1^{(k)}$, so that all the states that lie on the same side of the hyperplane as ρ_0 will pass the preselection condition.

Random states generated in the feasible space reduce the distance.

2.5 OBSERVATIONS REGARDING THE SIMPLIFIED GILBERT'S ALGORITHM

In this section we shall discuss some of the important aspects of the simplified Gilbert's Algorithm. We will look at how to generate the random pure separable states uniformly distributed over the set $\mathcal{S}_{\text{pure}}$, we shall prove that for any state there can only be one unique closest separable state due to the geometry imposed by the Hilbert-Schmidt norm and then we shall discuss the output of the algorithm and its interpretation.

2.5.1 *Generating Random Pure Separable states*

An important part of the algorithm is the generation of random pure separable states. It is sufficient to only pick uniformly from the set of pure separable states, $\mathcal{S}_{\text{pure}}$, because a convex combination of them provides us with the rest of the set $\mathcal{S}$. We generate random states that are uniformly distributed in $\mathcal{S}_{\text{pure}}$ using the following method. For an N-qudit system, we construct the tensor product of N individual d-dimensional states that are pure. To make a d-dimensional state, we take a list of $2d$ random real numbers drawn from a Normal distribution with a fixed deviation and the mean equal to 0. Then, the d consecutive pairs in the list are combined to form d complex numbers, and the list is normalized. Another way is to draw $2d$ random real numbers uniformly distributed over the interval $[0, 1]$ and again with consecutive pairs (a_i, b_i), build a complex variable $e^{2\pi i a_i}\sqrt{-2\ln b_i}$. We can then construct the density matrix ρ_2 from the pure state [62].

$\mathcal{S}_{\text{pure}} \subset \mathcal{S}$

2.5.2 *Uniqueness of the Closest Separable State*

We can also demonstrate by a simple argument, that the CSS found by the algorithm is unique. Suppose there are two states, $\rho_A,\ \rho_B \in \mathcal{S}$ that are at the same distance from our reference state ρ_0. Then by the convexity property all combinations of the two $p\rho_A + (1-p)\rho_B$ for $p \in [0, 1]$ also lie in $\mathcal{S}$. Then the distance of ρ_0 from the convex combination is in the form similar to Equation 2.10 and the minimum distance occurs at the vertex $p_{\min}$, as in Equation 2.11, of this quadratic function in p. At which point using the fact that $D^2_{\text{HS}}(\rho_0, \rho_A) = D^2_{\text{HS}}(\rho_0, \rho_B)$, we can show $p_{\min} = 1 - p_{\min}$. Therefore, $p_{\min} = 0.5$ and the minimum distance is attained for the equal mixture of the two states. The same reasoning can be applied to any number of states equidistant from the reference state, and one will always find the minimum distance occurring at the equal mixture of the states equidistant from ρ_0.

Every state has a unique CSS.

2.5.3 *Output of the Simplified Gilbert's Algorithm*

The output of the algorithm is not only the minimum distance reached at HALT, but also the series of states $\rho_1^{(k)}$ and the corresponding distance from ρ_0. Both are useful in providing insights about the algorithm as well as the relation between

a reference state and it's CSS. The series $\{D^2_{\text{HS}}(\rho_0, \rho_1^{(k)})\}$ has a decaying trend that is a very good indicator of how entangled a state is. When the algorithm is run for a separable state, the decay is more drastic than compared to an entangled state and the distance quickly tends towards zero. Within the limit of the defined precision for the problem we can define such states as practically separable. On the other hand, for entangled states the distance is without a doubt always non-zero. This final minimum distance obtained becomes a good quantifier of the entanglement in the state. It is higher for states that posses more entanglement, as can be verified in the cases where other entanglement measures can be calculated, for example, the bipartite maximally entangled states. The minimum distance then is lower when considering mixed entangled states, an example of which is progressively adding more white noise. The maximally mixed state or the normalized Identity operator is generally referred to as white noise and thus adding white noise means making a mixture of the state with the identity operator. In such a mixture the minimum distance decreases as the visibility (weight) of white noise increases in the mixture, eventually creating a separable state. This critical visibility for which all entanglement is destroyed is directly related to the entanglement measure *Random robustness* of the state.

2.6 SUMMARY

The algorithm is not limited to a particular convex set, this makes it incredibly versatile.

Before we move on, however, it is crucial to mention this fact about the versatility of the Gilbert's algorithm. The algorithm is designed to work on convex sets and as such it does not matter which convex set. Therefore, just as easily, we could substitute in all the previous analysis about running the algorithm, the fully separable set $\mathcal{S}$, with the convex set of k-separable states. The sole requirement that changes is the generation of random pure separable states, where instead of fully separable pure states one would generate random pure k-separable states. Therefore, running parallel instances of the algorithm searching in the sets of varying separability properties gives us the *Closest k-Separable States* for different values of k. This is a very valuable tool to reveal a detailed picture about the entanglement and separability properties for different classes of states. For example, in the case of three qubits, we can simultaneously look for the Closest (fully) separable state and the Closest *Biseparable* states in any bipartition $A - BC$, $B - AC$ and/or $AB - C$. In this case a normal desktop com-

puter, nowadays, is perfectly capable of running four parallel instances of the algorithm to obtain the CSS and the closest biseparable states simultaneously.

In the following chapter, we discuss examples of the application of the algorithm on some well known classes. We will compare the obtained results to the known analytical results. We will also make some interesting observations, discuss some important trends and use the output of the algorithm to gain analytical insights.

3

APPLICATION OF GILBERT'S ALGORITHM

In this chapter, we shall consider various well known classes of states and present the results that we obtain from the Gilbert's algorithm. We will look at states of varying dimensions and entanglement properties. As was discussed at the end of the previous chapter, the algorithm does not care in which convex set it runs, and it is guaranteed to converge so long as it is supplied with the random pure separable states from the correct set. We will discuss this in context of multiple qubits and provide new results with regards to closest k-separable states. We shall also change our perspective and look at what happens in the state-vector picture and the relation between the state-vectors of the reference states and their closest (k-)separable states.

3.1 BIPARTITE MAXIMALLY ENTANGLED STATES

The bipartite maximally entangled states are defined as

$$|\psi_d\rangle = \frac{1}{\sqrt{d}} \sum_{i=0}^{d} |i,i\rangle \,. \tag{3.1}$$

Due to the fact that we know the analytical form of the closest separable states in this case, these states serve as the testbed for our implementation of the simplified Gilbert's algorithm. The closest separable states for bipartite maximally entangled states are known to be the d-dimensional Werner states,

$$\rho_{\text{Wer}} = p\,|\psi_d\rangle\langle\psi_d| + (1-p)\frac{\mathbb{1}}{d}, \quad 0 \leq p \leq 1. \tag{3.2}$$

with $p = (d+1)^{-1}$. This gives us the analytical minimum distance

$$D^2_{\text{HS}_{\text{min}}}(|\psi_d\rangle\langle\psi_d|) = \frac{d-1}{d+1}. \tag{3.3}$$

Running the algorithm with $\rho_0 = |\psi_d\rangle\langle\psi_d|$, we get an approximation of the CSS and the sequence of distances from each iteration. The decay of the dis-

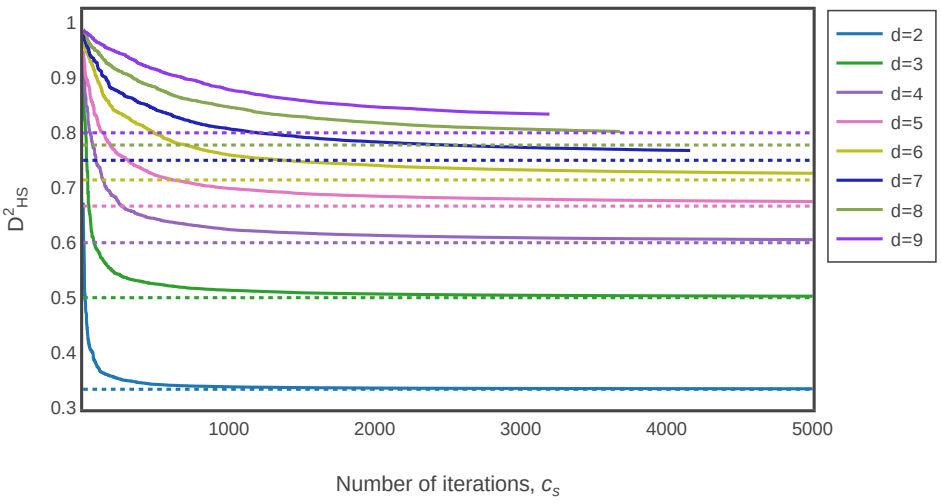


Figure 3.1: The plot shows the decay and convergence of $D^2_{\mathrm{HS}_{\mathrm{min}}}$ for 2 qudit maximally entangled states to the analytical minimum distance in Equation 3.3 (shown here in dotted lines) for values of d ranging from 2 to 9. It is apparent from the figure that increase in the dimension slows down the convergence.

tance with each iteration for dimensions $d = 2, 3, 4 \ldots, 9$ is plotted in Figure 3.1. The increasing dimension of the subsystems makes the search space grow exponentially, as we saw earlier, and this leads to a slower convergence to the analytical minimum. We can visualize this another way by using the number of rejected states per iteration. Such a plot is presented in the Figure 3.2 where the algorithm was run with HALT set as c_s >2,000, with input state dimensions $d = 2, 3, 4, 5$. To make the relation between the number of corrections/iterations with the number of rejections clear, the plot shows cumulative number of rejections instead of rejections per iteration with the axes being log scaled. The reason we accumulate the number of rejections for plotting is that in the last 100 iterations the number of rejections per iteration fluctuates wildly, for example, for $d = 5$ in the last 100 iterations the number was between 1,342 to 523,458, which leads to an obscure plot.

While generating random pure separable states, it is necessary to generate complex matrix entries. If we generate strictly real matrices then the outcome can be misleading. For example, in the case 2 qubit maximally entangled state,

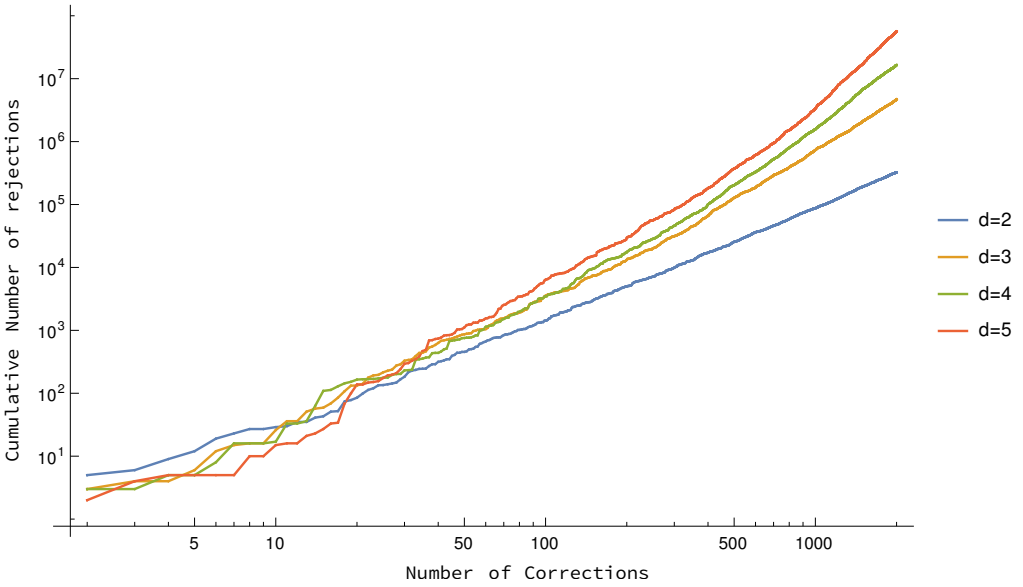


Figure 3.2: Cumulative rejection count for qudit maximally entangled states

by generating only real or only complex matrices we obtain two different closest separable states:

$$\lim_{c_s\to\infty}\rho_1^{\mathbb{R}} = \frac{1}{8}\begin{pmatrix} 3 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 3 \end{pmatrix} \qquad \lim_{c_s\to\infty}\rho_1^{\mathbb{C}} = \frac{1}{6}\begin{pmatrix} 2 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 2 \end{pmatrix} \tag{3.4}$$

The distance of these states from $\rho_0 = |\psi_2\rangle\langle\psi_2|$ is 0.625 for $\rho_1^{\mathbb{R}}$ and 0.5 for $\rho_1^{\mathbb{C}}$. Clearly, only generating the real matrices at random results in states that are not evenly distributed on the set $\mathcal{S}_{\text{pure}}$.

3.1.1 *Predicting $D^2_{HS_{\min}}$ with Linear Model Fitting*

In Figure 3.1, the distance $D^2_{\text{HS}}(\rho_0, \rho_1^{(k)})$ decreases drastically as it progresses from the initial guess but then the decay slows down as it reaches the analytical limit. The shape of this decay is similar to an exponential curve that has it's

asymptote at the analytical minimum distance. For instance we can employ curve fitting methods and try to fit the following function to our data,

$$D^2_{\text{HS}}(\rho_0, \rho_1^{(k)}) = \mathrm{e}^{y_k^{1/b}} + D^2_a, \tag{3.5}$$

where y_k is the fit variable, b and D^2_a are the fit parameters, so that D^2_a is the asymptote for this exponential function. This would constitute a non-linear model, instead we could rewrite the above into a linear form and employ linear regression with,

$$y_k = \left| \ln[D^2_{\text{HS}}(\rho_0, \rho_1^{(k)}) - D^2_a] \right|^b. \tag{3.6}$$

True only in the case of simple linear regression.

To judge the goodness of the fit we look at the *sample correlation coefficient*, in this case, the square or R^2 is the *coefficient of determination*, which is calculated using the formula,

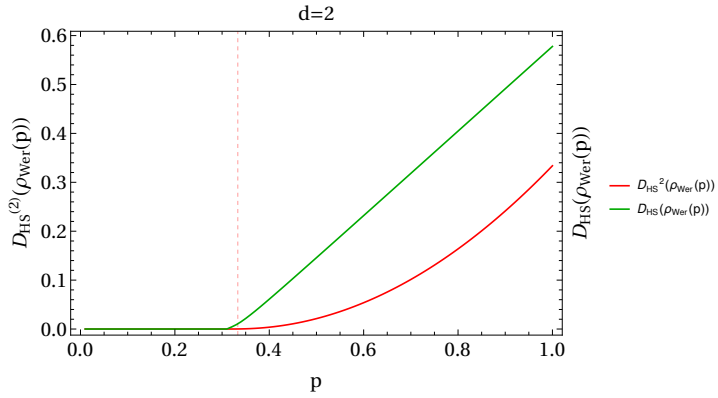
$$R(k, y_k) = \frac{\langle k\, y_k \rangle - \langle k \rangle \langle y_k \rangle}{\sqrt{\left(\langle k^2 \rangle - \langle k \rangle^2\right)\left(\langle y_k^2 \rangle - \langle y_k \rangle^2\right)}} \in [-1, 1], \tag{3.7}$$

and $R^2 \in [0, 1]$. Here $k = 1, 2 \ldots$ is a positive integer sequence denoting the iteration numbers. We maximize the sample correlation coefficient by varying the fit parameters b and D^2_a. Let's take the example of the bipartite maximally entangled states discussed above. On maximizing the said correlation coeffiecient, a value close to 1 indicates a very good fit. The results of our linear model fitting are in Table 3.1.

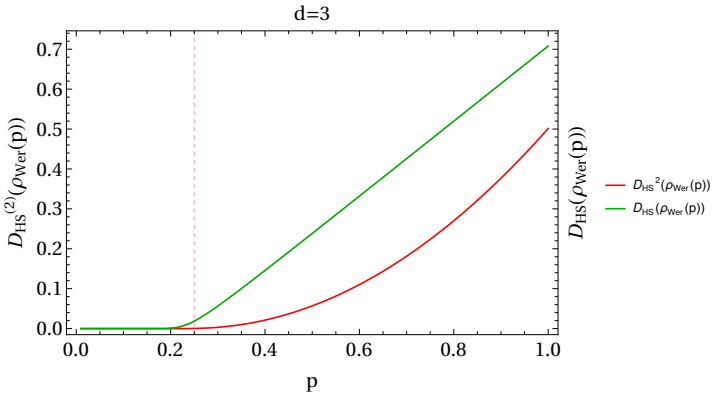
We obtain very good fits, which is not only indicated by R^2 being close to 1, but also the fact that the asymptotic minimum distance lies closer to the analytical minimum distance for these states, and is less than the minimum distance obtained from the algorithm after 2,000 iterations. Therefore, indicating that the algorithm successfully reaches very close to the closest separable state, and the approximation only gets better with increasing iterations.

3.1.2 *A case study of bipartite Werner states*

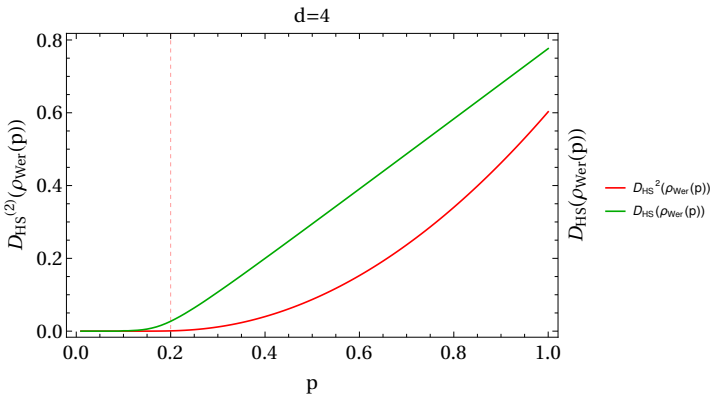
We defined the bipartite Werner states in Equation 3.2, where $p \in [0, 1]$ defines all the states on the line joining $|\psi_d\rangle\langle\psi_d|$ to the d-dimensional identity $\mathbb{1}_d$. Given



(a) $d = 2$



(b) $d = 3$



(c) $d = 4$

Figure 3.3: The plots shows the minimum Hilbert-Schmidt distance of bipartite Werner states as a function of p when the algorithm is run for 10^4 iterations. (a) d=2, (b) d=3 and (c) d=4. In each case the distance D_{HS} starts increasing at the critical visibility $p = (d+1)^{-1}$ (red dashed-line). Plotted in green is D_{HS}, and D_{HS}^2 is in red (solid curve).

d	R^2	b	D_a^2	D_{HS}^2	$D_{\text{HS}_{\text{min}}}^2(\|\psi_d\rangle\langle\psi_d\|)$
2	0.9998	5.58	0.33369	0.33535	0.33333
3	0.9994	4.84	0.50133	0.50573	0.5
4	0.9996	4.61	0.60049	0.61190	0.6
5	0.9997	4.48	0.66510	0.68334	0.66666

Table 3.1: The results of linear regression for d-dimensional bipartite maximally entangled states is shown. The second column, R^2 measures the goodness of fit (closer to 1 is better). Third and fourth columns show the fit parameters of which D_a^2 is the predicted asymptotic minimum distance. D_{HS}^2 is the minimum distance obtained after 2000 iterations. Finally, the last column shows the analytical minimum distance. We can observe the ordering $D_{\text{HS}}^2 \leq D_a^2 \leq D_{\text{HS}_{\text{min}}}^2(|\psi_d\rangle\langle\psi_d|)$.

the fact that the closest separable state lies on this line, it motivates us to see how the minimum distance, $D_{\text{HS}_{\text{min}}}^2(\rho_{\text{Wer}}(p))$ changes with p. We do this by running the algorithm for 100 values of p in the interval $[0, 1]$ for $d = 2, 3, 4$ with the HALT condition to be the combination $c_s > 10000$ or $D_{\text{HS}}^2(\rho_0, \rho_1^{(k)}) < 10^{-6}$. The results are shown in the Figure 3.3 and are in agreement with what we already know about the separability of these Werner states. Note that the plots show both D_{HS}^2 (red) and D_{HS} (green), because in this scenario the behaviour of D_{HS} provides us with more relevant information. For $d = 2, 3, 4$ the distance D_{HS} starts increasing at around the same value of $p = (d+1)^{-1}$, which we know is the critical visibility.

Therefore, we can estimate and give a lower bound for the critical visibility of a given state with respect to addition of noise using the Gilbert's algorithm and as critical visibility is directly related to Random robustness of the state, it can also be estimated.

3.2 N-QUBIT GHZ STATES

Next we consider another well known class of states, the Greenberger-Horne-Zeilinger (GHZ) states of $N(> 2)$ qubits. The N-qubit GHZ state is defined as,

$$|\text{GHZ}_N\rangle = \frac{1}{\sqrt{2}}\left(|0\rangle^{\otimes N} + |1\rangle^{\otimes N}\right) \tag{3.8}$$

and $\rho^N_{\text{GHZ}} = |\text{GHZ}_N\rangle\langle\text{GHZ}_N|$. This is another case where the analytical form of the closest separable states have been known in the literature [42] to be the convex combination, $\rho_{\text{CSS}} = t_N\varrho_N + (1-t_N)\Delta_N$ where ϱ_N and Δ_N are $2^N \times 2^N$ matrices defined as

$$\varrho_N = \frac{1}{2}\begin{pmatrix} 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & \dots & 0 & 1 \end{pmatrix}, \quad \text{and} \quad \Delta_N = \frac{1}{2^N}\begin{pmatrix} 1 & 0 & \dots & 0 & 1 \\ 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & 0 \\ 1 & 0 & \dots & 0 & 1 \end{pmatrix}, \tag{3.9}$$

and $t_N = (2^N-2)^2(4+4^N-2^{N+1})^{-1}$. This gives us the analytical minimum distance,

$$D^2_{\text{HS}_{\min}}(\rho^N_{\text{GHZ}}) = \frac{1}{2}\left(1 - \frac{1}{4^{N-1}-2^{N-1}+1}\right). \tag{3.10}$$

If we compare this to the case of bipartite maximally entangled states, the closest separable state is not on the line joining the state and the identity. It is interesting to note that the convex combination $p\varrho_N + (1-p)\Delta_N$ equals the state $p_S\rho^N_{\text{GHZ}} + (1-p_S)\mathbb{1}_{2^N}$ when $p = p_S = (2^{N-1}+1)^{-1}$.

Usually states of the form $p_S\rho^N_{\text{GHZ}} + (1-p_S)\mathbb{1}_{2^N}$ are called *generalized* Werner states as they are the counterparts of bipartite Werner states. Again it is a previously known result that N-qubit generalized Werner states $\rho_{\text{gWer}}(p)$ are fully separable for $0 \le p \le (2^{N-1}+1)^{-1}$. Therefore, both ρ_{CSS} and $\rho_{\text{gWer}}(p_S)$ lie on the boundary of the set $\mathcal{S}$.

The algorithm, when run for $N = 3, 4, 5$, gives us very good results and converges to the closest separable state discussed above. To give an example of how close the approximation is to the analytical ρ_{CSS}, we write both the ρ_{CSS} and the approximate closest separable state $\tilde{\rho}_{\text{CSS}}$ for $N = 3$,

$$\rho_{\text{CSS}} = \begin{pmatrix} 0.3846 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 0.0384 \\ \cdot & 0.0384 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & 0.0384 & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & 0.0384 & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & 0.0384 & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & 0.0384 & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 0.0384 & \cdot \\ 0.0384 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 0.3846 \end{pmatrix} \tag{3.11}$$

$$\tilde{\rho}_{\text{CSS}} = \begin{pmatrix} 0.3814 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 0.0368 \\ \cdot & 0.0393 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & 0.0393 & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & 0.0392 & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & 0.0393 & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & 0.0394 & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 0.0393 & \cdot \\ 0.0368 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 0.3823 \end{pmatrix} \tag{3.12}$$

That the two matrices are close is apparent when we look at the matrix entries, and it is reaffirmed by the fact that the Hilbert-Schmidt distance between the two is 0.0000118. Additionally, their distances from the GHZ state are $D^2_{\text{HS}}(\rho^3_{\text{GHZ}}, \rho_{\text{CSS}}) = 0.46153846$ and $D^2_{\text{HS}}(\rho^3_{\text{GHZ}}, \tilde{\rho}_{\text{CSS}}) = 0.46616934$.

Similarly, in the case of 4 qubit GHZ states we found the distance to be 0.49699147 compared to the analytical minimum distance of 0.49122807 after 5,000 iterations. The obtained closest separable state from the algorithm has the same structure as the analytical. This makes for an important observation that the *CSS and the reference state both have same eigenvectors* but different eigenvalues, and thus ρ_{CSS} is diagonal in the eigenbasis of the reference state.

3.3 N-QUBIT W STATES

We discuss now the class of N qubit W states,

$$|W_N\rangle = \frac{1}{N}(|100\cdots0\rangle + |010\cdots0\rangle + |001\cdots0\rangle + |0\cdots001\rangle). \tag{3.13}$$

While we do not have analytical knowledge about the closest separable states for this class, nevertheless, there is an interesting pattern that emerges in the closest separable states found by the algorithm.

Let's first consider the 3 qubit W state. After running the algorithm for 10,000 iterations, it yielded the following $\tilde{\rho}_{\text{CSS}}$ at a distance of 0.43216016 from the state $|W_3\rangle\langle W_3|$,

$$\tilde{\rho}_{\text{CSS}} = \begin{pmatrix} 0.219786 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & 0.180468 & 0.11966 & \cdot & 0.119807 & \cdot & \cdot & \cdot \\ \cdot & 0.11966 & 0.179624 & \cdot & 0.119951 & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & 0.068536 & \cdot & 0.063988 & 0.064144 & \cdot \\ \cdot & 0.119807 & 0.119951 & \cdot & 0.180505 & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & 0.063988 & \cdot & 0.068833 & 0.064126 & \cdot \\ \cdot & \cdot & \cdot & 0.064144 & \cdot & 0.064126 & 0.068945 & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 0.033303 \end{pmatrix} \tag{3.14}$$

Notice, there are non-zero matrix entries at the positions corresponding to two excitations, for instance $|011\rangle\langle011|$, $|011\rangle\langle101|$, $|011\rangle\langle110|$, etc. This suggests that the CSS for W state is a convex combination of N-qubit symmetric states, that are invariant under the exchange of subsystems, because the W state possesses

the same symmetry under exchange of subsystems. Indeed we find that the ρ_{CSS} in Equation 3.14 is at a distance of 10^{-6} from the following convex combination of density matrices which possess the same symmetry,

$$\begin{aligned}\rho'_{\text{CSS}} = 0.219786\,|000\rangle\langle 000| + 0.359418\,|W\rangle\langle W| + 0.192258\,|W_x\rangle\langle W_x| \\ + 0.181179\,\pi_W + 0.014056\,\pi_{W_x} + 0.033303\,|111\rangle\langle 111|\,, \quad (3.15)\end{aligned}$$

where

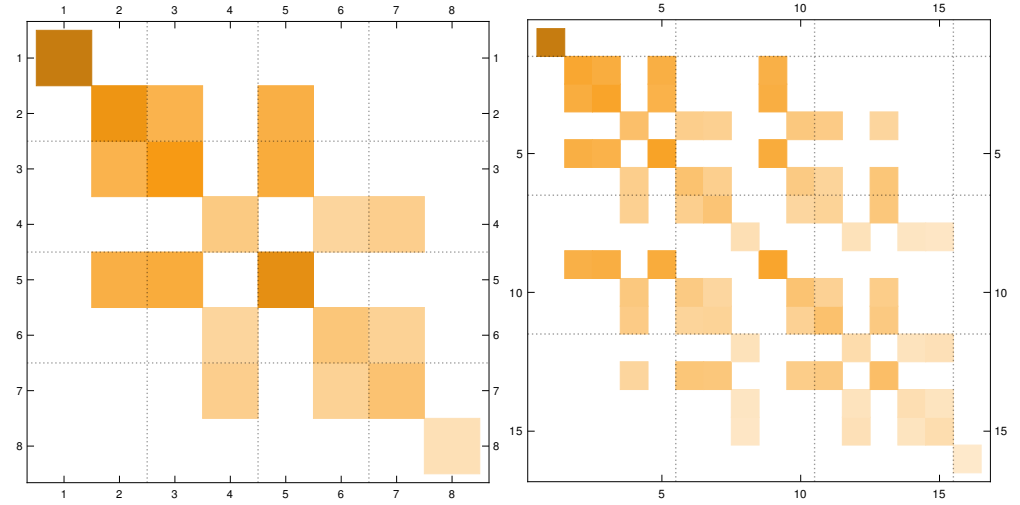
$\|W\rangle\langle W\|$	is the W state,
$\|W_x\rangle\langle W_x\|$	is $\sigma_x^{\otimes 3}\,\|W\rangle\langle W|\,(\sigma_x^{\otimes 3})^\dagger$
π_W	has projectors from the diagonal of $\|W\rangle\langle W\|$, and
π_{W_x}	has projectors from the diagonal of $\|W_x\rangle\langle W_x\|$.

We also ran the algorithm for $N = 4, 5, 6$ and obtained $D^2_{\text{HS}_{\text{min}}} = 0.473063$, 0.501705, 0.544239 respectively and with number of corrections $c_s = 9700$, 2600, 900. We observe a similar pattern in the closest separable states where in the case of $N = 3$ we had matrix entries corresponding to 1 and 2 excitations, for $N = 4$ we have 1,2 and 3 excitations and so on. Therefore they can be written as convex combinations of N-qubit symmetric states, like in the case of $N = 3$ W state. As the matrices are too large to fit here, we try to show this fact using a matrix plot that assigns colors to elements depending on their value. See Figure 3.4 where the matrix plot (a) represents the above CSS for $N = 3$ for reference.

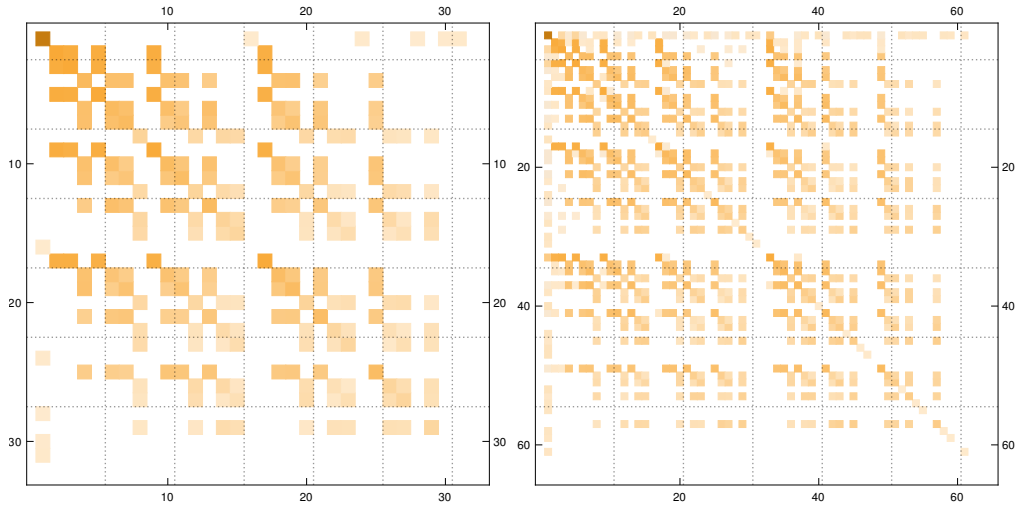
Another point of note is that in both $N = 3$ and $N = 4$ the GHZ states have a higher minimum distance compared to the W states for the same N, indicating higher entanglement in GHZ states, which we know to be true. Thus once again the method coincides with previously known facts.

3.4 CLOSEST BISEPARABLE STATES

Till now we have only discussed closest *fully separable* states for the discussed classes. A much less studied aspect is biseparability in $N \geq 3$ systems. In this section we will present our findings and some insights about the geometry of the set of biseparable states, combined with what we learned in the previous sections about the fully separable states. Let's take a moment to fix notation to avoid tedious repetition.



(a) $N = 3$ (b) $N = 4$



(c) $N = 5$ (d) $N = 6$

Figure 3.4: Matrix plots of the closest separable states for N-qubit W states. Darker the color, higher the value. (a) N=3, the corresponding matrix is shown in Equation 3.14. (b) N=4, (c) N=5 and (b) N=6

- CBiS_X : Closest Biseparable State on the bipartition denoted by the subscript as X-to-Rest, where X is a subset of the systems labeled in roman alphabet A, B, C and so on.
- CBiS : Overall Closest Biseparable state (without subscript) among all bipartitions.

In the case of biseparability, there are two different approaches that need to be considered. Taking the example of $N = 3$ case, if the systems are labeled A,B and C, there are three possible bipartitions: $A - BC$, $B - AC$ and $C - AB$. The first approach then is to only run the algorithm for the set of states that are separable across one of the bipartitions, say $A - BC$, to obtain the closest biseparable state from this bipartition, CBiS_A. Repeating this for the other two, we get as many closest biseparable states as there are possible bipartitions, here CBiS_B and CBiS_C. While the set of separable states across a particular bipartition is convex, all linear combinations of them are also biseparable and form a convex set. Therefore, as we proved the uniqueness of the CSS earlier, we find that a convex combinations of these states must be closer still to the reference state by simple geometry. If with respect to the reference state, all such closest biseparable states were symmetrically placed, i. e., at the same distance then the equal mixture of all them should, in principle, give us the overall Closest Biseparable state, CBiS. Otherwise, one could hypothesize that a certain convex combination of these, not necessarily equal, would result in the CBiS.

The second approach tries to directly find the overall closest biseparable state. This is done by randomly picking one of the bipartitions at each iteration, say $A - BC$. Then the random pure biseparable state generated is separable across $A - BC$. In the next iteration, the algorithm might pick $C - AB$ for generating the random pure state. In this way, the algorithm makes a convex combination of all the different bipartitions, resulting in the CBiS.

We would also expect the outcomes from the first and the second approaches to be the same, or at least very close together, indicating that they would converge given enough iterations. Intriguingly, this is not the case as we will see in the examples that follow. Even when the individual closest biseparable states are symmetrically placed with respect to the reference state, *the second approach finds an overall closest biseparable state that is closer compared to the overall closest biseparable state found from the first approach*. Although, a look at the eigenvalues

of both the states reveals that they are both indeed on the closest face of the convex set of biseparable states.

Before we look at the examples, we will consider the method for constructing the closest PPT state first mentioned in the work [53] by Verstraete et. al. We shall call this Verstraete's method for clarity.

3.4.1 *Verstraete's method for calculating the closest PPT state*

We mention this here because this is the only known analytical method to construct the Closest PPT state, which of course in some cases will coincide with closest biseparable state, because the set of biseparable states is a subset of the set of PPT states. Their argument is based on the fact that the Hilbert-Schmidt distance to the set of PPT states does not change under partial transposition of any one of the systems, or a subset of systems forming a bipartition.

For a given reference state ρ_0, ρ_0^{PT} denoting partial transposition, their algorithm to find the closest PPT state is as follows:

1. Calculate the eigenvalue decomposition of the state $\rho_0^{\mathrm{PT}} = UDU^\dagger$, such that d_is are the diagonal entries of D.

2. Define a new diagonal matrix E which has diagonal entries,

$$e_i = \max\{0, d_i + c\} \quad \text{if} \quad d_i > 0 \quad \text{else} \quad e_i = 0. \tag{3.16}$$

   where c is the sum of the negative eigenvalues divided by the number of strictly positive eigenvalues. Effectively, we are redistributing the negative eigenvalues to the positive ones, to make them smaller and sum up to 1.

3. Construct the closest PPT state as $\rho_{\mathrm{PPT}} = (UEU^\dagger)^{\mathrm{PT}}$.

It is important to note that it is possible that the resulting state ρ_{PPT} is not positive semi-definite, and hence not a valid density matrix. On the contrary, if it a valid density matrix, it is guaranteed to be the Closest PPT state (in the particular bipartition).

3.4.2 *GHZ states*

Lets go back to the N-qubit GHZ states and run the algorithm using the two approaches discussed above. We will discuss the case of $N = 3$ and $N = 4$ qubit GHZ states.

3.4.2.1 *First Approach*

Consider first the $N = 3$ GHZ state, with possible bipartitions $A - BC$, $B - AC$ and $C - AB$. The algorithm run for the three bipartitions for 1,000 iterations results in the following closest biseparable states.

In the bipartition $A - BC$,

$$\mathrm{CBiS}_A = \begin{pmatrix} 0.32147 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 0.15903 \\ \cdot & 0.0070622 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & 0.0069418 & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & 0.16515 & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & 0.16118 & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & 0.0072362 & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 0.0073022 & \cdot \\ 0.15903 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 0.32364 \end{pmatrix} \tag{3.17}$$

In the bipartition $B - AC$,

$$\mathrm{CBiS}_B = \begin{pmatrix} 0.31942 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 0.15943 \\ \cdot & 0.16218 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & 0.0073232 & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & 0.0071405 & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & 0.0075479 & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & 0.0072055 & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 0.16496 & \cdot \\ 0.15943 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 0.32420 \end{pmatrix} \tag{3.18}$$

In the bipartition $AB - C$,

$$\mathrm{CBiS}_C = \begin{pmatrix} 0.32435 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 0.15864 \\ \cdot & 0.0079352 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & 0.16226 & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & 0.0073087 & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & 0.0077868 & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & 0.16302 & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 0.0070293 & \cdot \\ 0.15864 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 0.32030 \end{pmatrix} \tag{3.19}$$

Their distances form the GHZ state are as follows, $D^2_{\mathrm{HS}}(\rho_{\mathrm{GHZ}}, \mathrm{CBiS}_A) = 0.34895$, $D^2_{\mathrm{HS}}(\rho_{\mathrm{GHZ}}, \mathrm{CBiS}_b) = 0.34921$ and $D^2_{\mathrm{HS}}(\rho_{\mathrm{GHZ}}, \mathrm{CBiS}_C) = 0.34932$. It seems they are very close to being equidistant from GHZ state within computational deviation. Another thing to notice is that all three of the above can be transformed

into each other by applying the swap operator across the bipartition. For instance, $(U_{\text{swp}} \otimes \mathbb{1})\text{CBiS}_A(U_{\text{swp}} \otimes \mathbb{1})^\dagger$ gives a state very close numerically to CBiS_B. These two observations strongly suggest that the exact closest biseparable states corresponding to these three lie symmetrically around the GHZ state.

If we take a convex combination of the three as

$$\sigma = p_1\text{CBiS}_A + p_2\text{CBiS}_B + (1 - p_1 - p_2)\text{CBiS}_C$$

and minimize the Hilbert Schmidt distance, $D^2_{\text{HS}}(\rho_{\text{GHZ}}, \sigma)$ for $0 \leq p_1, p_2 \leq 1$, we find that the minimum distance $\min_{p_1,p_2} D^2_{\text{HS}}(\rho_{\text{GHZ}}, \sigma) = 0.31678$ occurs at $p_1 = 0.3358$ and $p_2 = 0.3337$, which as we expected is very nearly an equal mixture of the three and is closer to the GHZ state. We shall call it CBiS_1,

$$\text{CBiS}_1 = \begin{pmatrix} 0.32174 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 0.15903 \\ \cdot & 0.059126 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & 0.058391 & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & 0.060257 & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & 0.059219 & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & 0.058701 & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 0.059835 & \cdot \\ 0.15903 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 0.32272 \end{pmatrix} \tag{3.20}$$

An interesting thing to check would be to calculate the closest PPT state to GHZ on each of these bipartitions. The resultant states are listed here,

$$\sigma^A_{\text{PT}} = \begin{pmatrix} 0.333333 & \cdot & \cdot & \cdot & \cdot & 0.166667 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & 0.166667 & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & 0.166667 & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0.166667 & \cdot & \cdot & \cdot & \cdot & 0.333333 \end{pmatrix} \tag{3.21}$$

$$\sigma^B_{\text{PT}} = \begin{pmatrix} 0.333333 & \cdot & \cdot & \cdot & \cdot & 0.166667 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & 0.166667 & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & 0.166667 & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0.166667 & \cdot & \cdot & \cdot & \cdot & 0.333333 \end{pmatrix} \tag{3.22}$$

$$\sigma^C_{\text{PT}} = \begin{pmatrix} 0.333333 & \cdot & \cdot & \cdot & \cdot & 0.166667 \\ \cdot & 0.166667 & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & 0.166667 & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0.166667 & \cdot & \cdot & \cdot & \cdot & 0.333333 \end{pmatrix} \tag{3.23}$$

Also, because they are all at an equal distance $1/3$ from the GHZ state, by symmetry their equal mixture will have the least distance among all the convex combinations. The equal mixture σ_{GHZ} is,

$$\sigma_{\text{GHZ}} = \begin{pmatrix} 0.3333 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 0.1666 \\ \cdot & 0.05555 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & 0.05555 & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & 0.05555 & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & 0.05555 & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & 0.05555 & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 0.05555 & \cdot \\ 0.1666 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 0.3333 \end{pmatrix} \tag{3.24}$$

and it is at a distance of $0.(296)$ from the GHZ state. Comparing the matrix entries of σ_{GHZ} and CBiS_1, we see they are quite close. The same pattern is seen in the corresponding states in each bipartition. Where the matrix entries are zero in σ_{PT}^X, the entries are about 0.007 in the CBiS_X, and the same can be seen for the other entries as well.

More importantly one can look at the eigenvectors and eigenvalues of the corresponding matrices. For example, the eigenvectors and corresponding eigenvalues of CBiS_A and σ_{PT}^A are,

$$ED(\text{CBiS}_A) = \begin{pmatrix} 0.481595 & \{-0.704699, 0., 0., 0., 0., 0., 0., -0.709506\} \\ 0.163526 & \{-0.709506, 0., 0., 0., 0., 0., 0., 0.704699\} \\ 0.165151 & \{0., 0., 0., 1., 0., 0., 0., 0.\} \\ 0.161185 & \{0., 0., 0., 0., 1., 0., 0., 0.\} \\ 0.00723626 & \{0., 0., 0., 0., 0., 1., 0., 0.\} \\ 0.00730223 & \{0., 0., 0., 0., 0., 0., -1., 0.\} \\ 0.00706222 & \{0., 1., 0., 0., 0., 0., 0., 0.\} \\ 0.00694187 & \{0., 0., 1., 0., 0., 0., 0., 0.\} \end{pmatrix} \tag{3.25}$$

$$ED(\sigma^{A}_{\mathrm{PT}}) = \begin{pmatrix} 0.5 & \{-0.707107, 0., 0., 0., 0., 0., -0.707107\} \\ 0.166667 & \{-0.707107, 0., 0., 0., 0., 0., 0., 0.707107\} \\ 0.166667 & \{0., 0., 0., 1., 0., 0., 0., 0.\} \\ 0.166667 & \{0., 0., 0., 0., 1., 0., 0., 0.\} \\ 0. & \{0., 0., 0., 0., 0., 1., 0., 0.\} \\ 0. & \{0., 0., 0., 0., 0., 0., -1., 0.\} \\ 0. & \{0., 1., 0., 0., 0., 0., 0., 0.\} \\ 0. & \{0., 0., 1., 0., 0., 0., 0., 0.\} \end{pmatrix} \tag{3.26}$$

where $ED(\rho)$ denotes the eigendecomposition of ρ. We have arranged the eigenvectors in the same order to bring out the similarity in both cases. It becomes apparent then the matrices CBiS_X approach σ^{X}_{PT} for $X = A, B, C$ as the algorithm progresses.

Therefore, it is safe to conclude that using the first approach we find closest biseparable states in each bipartition that are also the closest PPT states in that bipartition with respect to our reference state. And that the equal mixture of these is closer still.

3.4.2.2 *Second Approach*

Now in the second approach we don't calculate the closest biseparable states in each bipartition but randomly choose a bipartition to generate the random pure biseparable state in each iteration. After 5000 iterations we obtain the following state,

$$\mathrm{CBiS}_2 = \begin{pmatrix} 0.28574 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 0.20612 \\ \cdot & 0.07125 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & 0.07166 & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & 0.07007 & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & 0.07132 & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & 0.07098 & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 0.07321 & \cdot \\ 0.20612 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 0.28573 \end{pmatrix} \tag{3.27}$$

which as it turns out is remarkably close to the three qubit generalized Werner state with the critical visibility for biseparability.

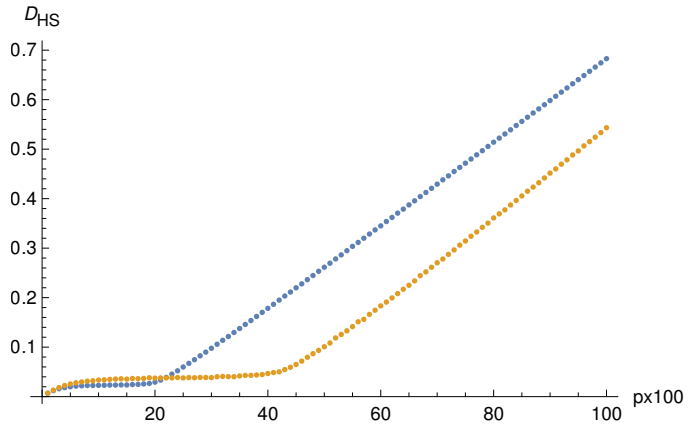


Figure 3.5: Plot showing the minimum Hilbert distance of the 3-qubit generalized Werner state for 100 values of p from the set of fully separable states (blue) and from the set of bisebarable states (yellow).

In the N-qubit scenario, the states $\rho_{\text{GWer}}(p) := p\rho_{\text{GHZ}}^{N} + (1-p)\mathbb{1}_{2^N}$ are,

$$\begin{aligned} \text{fully separable} \quad & 0 \leq p \leq \left(1 - \tfrac{1}{1+2^{1-N}}\right), \\ \text{biseparable} \quad & \left(1 - \tfrac{1}{1+2^{1-N}}\right) < p \leq \left(1 - \tfrac{1}{2(1-2^{-N})}\right), \\ \text{genuinely entangled} \quad & \left(1 - \tfrac{1}{2(1-2^{-N})}\right) < p \leq 1. \end{aligned} \tag{3.28}$$

For $N = 3$ the states ρ_{GWer} are biseparable for $\frac{1}{5} < p \leq \frac{3}{7}$, and the generalized Werner state with $p = 3/7$ is,

$$\rho_{\text{GWer}}(3/7) = \begin{pmatrix} 0.28571 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 0.21428 \\ \cdot & 0.07142 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & 0.07142 & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & 0.07142 & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & 0.07142 & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & 0.07142 & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 0.07142 & \cdot \\ 0.21428 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 0.28571 \end{pmatrix} \tag{3.29}$$

The distances of the states CBiS_2 and $\rho_{\text{GWer}}(3/7)$ from GHZ state are 0.2951 and 0.2857 respectively. Comparing the matrix entries as well as the Eigenvalue decompositions of both $\rho_{\text{GWer}}(3/7)$ and CBiS_2 leads us to conclude that the algorithm converges to the generalized Werner state at the boundary of biseparability for the mixture of GHZ with noise.

In particular, if one looks at the minimum Hilbert-Schmidt distance to the set of fully separable and biseparable states for the 3 qubit generalized Werner

states for different values of visibility p, we find that it agrees with the Equation 3.28. This is shown in the Figure 3.5. In this case full separability is up to 0.2 and biseparability is up to $3/7 = 0.428571$.

In both approaches it seems we have a different closest biseparable state overall, although as we verified both of them are on the boundary of the set of biseparable states. The most intriguing part is that the three states in question, CBiS_1, CBiS_2 and $\rho_{\text{GWer}}(3/7)$, all lie on the same line, written as the convex combination, $\varsigma(x) := x\,\rho_{\text{GWer}}(3/7) + (1-x)\varrho_3$, where $0 \leq x \leq 1$ and ϱ_3 is defined in Equation 3.9. The values of x corresponding to each of the three are,

$$\begin{aligned} x &= 7/9 = 0.7778 \quad & \varsigma(x) &= \text{CBiS}_1, \\ x &= 0.9836 & \varsigma(x) &= \text{CBiS}_2, \\ x &= 1 & \varsigma(x) &= \rho_{\text{GWer}}(3/7). \end{aligned}$$

3.4.2.3 *Summarizing the findings*

We can summarize the comprehensive findings that we have gathered above for 3 qubit GHZ states in the Figure 3.6. The figure combines the findings from the fully separable and biseparable runs of the algorithm, by illustrating the relative positions of all the relevant states in a 3-dimensional section of the hyperball B^{63} in which all the 3 qubit states reside. In the figure it is easier to see how the boundaries of the set of separable states and the set of biseparable states are relative to the line joining the states GHZ and $\mathbb{1}_8$, on which all the generalized Werner states lie.

The fact that the closest biseparable state to GHZ is the Werner state, implies that the boundary of the set of biseparable states is perpendicular to the line GHZ to $\mathbb{1}_8$, which then implies that the normal of the hyperplane that defines the face of the convex set of biseparable states is $\rho_{\text{GHZ}} - \mathbb{1}_8$, with all the biseparable states either lying on the plane or on the same side as $\mathbb{1}_8$.

When we do the same analysis for $N = 4$ GHZ states, we find the same patterns. Every step in the above analysis follows as for $N = 3$. The number of bipartitions is now higher, however, they provide closest separable states that are equivalent under swap operations and converge to the corresponding closest PPT states calculated by Verstraete's method. We find again that the line joining the 4-qubit GHZ to the $\mathbb{1}_{16}$ is perpendicular to the hyperplane defining the face of the set of biseparable states, thus, the Werner state with $p = 7/15$

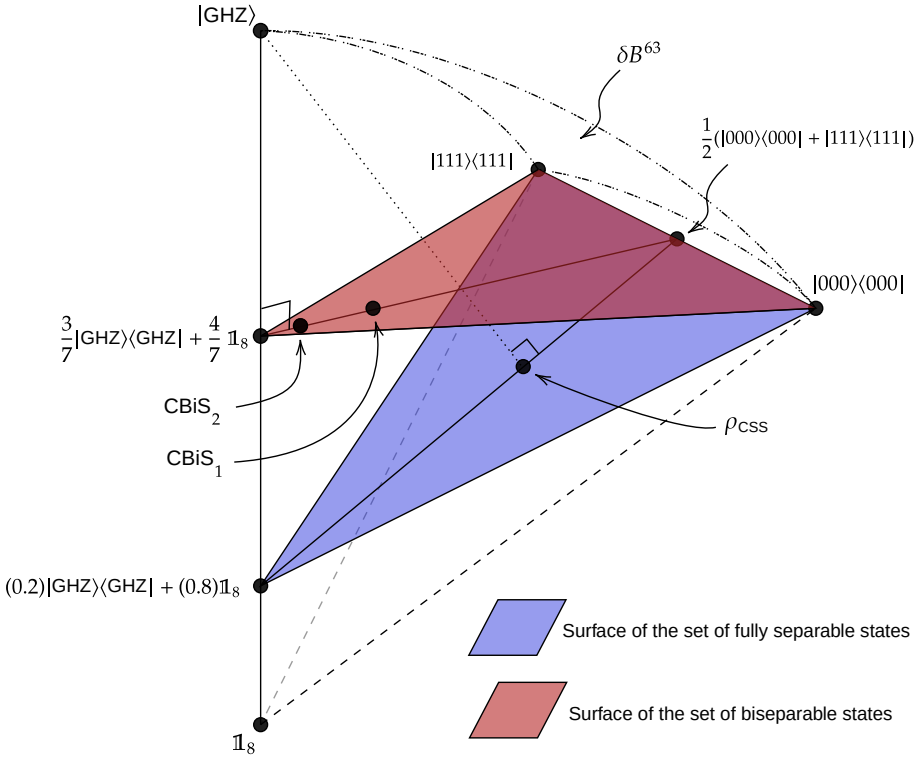


Figure 3.6: A section of the hyperball B^{63} where the 3 qubit GHZ resides, with its closest fully separable state ρ_{CSS}, closest biseparable state from the second approach of the algorithm, $CBiS_2$ and actual closest separable state, the generalized werner state, ρ_{GWer}, with $p = 3/7$. The maximally mixed state is in the center of the hyperball. ρ_{CSS} lies on the line joining ϱ_3 and werner state with visibility $p = 0.2$. The five states, $|000\rangle\langle 000|$,$|111\rangle\langle 111|$,$\frac{1}{2}(|000\rangle\langle 000| + |111\rangle\langle 111|)$, ρ_{CSS} and $\rho_{GWer}(0.2)$ lie on the hyperplane that constitutes the face of the convex set of fully separable states (blue). The six points $|000\rangle\langle 000|$,$|111\rangle\langle 111|$,$\frac{1}{2}(|000\rangle\langle 000| + |111\rangle\langle 111|)$, $\rho_{GWer}(3/7)$, $CBiS_1$ and $CBiS_2$ lie on the hyperplane defining the face of the convex set of biseparable states. This face is perpendicular to the line from GHZ to $\mathbb{1}_8$. The curves joining the pure states GHZ, $|000\rangle\langle 000|$ and $|111\rangle\langle 111|$ denote the surface of the hyperball, δB^{63}.

is the closest biseparable state to 4-qubit GHZ. Similar to the case of bipartite maximally entangled states, this leads to the following proposition:

- `Proposition:` The closest biseparable state to an N-qubit GHZ state is the generalized Werner state with visibility $p = (1 - 0.5(1 - 2^{-N})^{-1})$,

$$\rho_{\text{GWer}} = \frac{1 - 2^{1-N}}{2(1 - 2^{-N})} |\text{GHZ}_N\rangle\langle\text{GHZ}_N| + \frac{1}{2(1 - 2^{-N})} \mathbb{1}_{2^N} \tag{3.30}$$

  and its squared Hilbert-Schmidt distance from the GHZ state is $\frac{2^{N-2}}{2^N - 1}$.

Another interesting insight, which was tested and found to be correct, is that the closest fully separable states of the the generalized Werner states all lie on the line $x\rho_{\text{GWer}}(0.2) + (1 - x)\rho_{\text{CSS}}$.

3.4.3 *W state*

The same analysis was done for 3-qubit W state using the two approaches to find the closest biseparable state in each bipartition as well as the overall closest biseparable state.

When using the first approach, the closest biseparable states in each bipartition converge to the corresponding closest PPT state calculated using Verstraete's method.

Although, again, as we discovered in the GHZ case, the convex mixtures of the closest separable states in each bipartition are not enough to characterize the face of the set of biseparable states. The algorithm yet again finds a better candidate for the overall closest separable state. The two matrices are given below. Closest convex mixture from the first approach,

$$\text{CBiS}_1 = \begin{pmatrix} 0.188562 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & 0.239052 & 0.2 & \cdot & 0.2 & \cdot & \cdot & \cdot \\ \cdot & 0.2 & 0.239052 & \cdot & 0.2 & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & 0.031427 & \cdot & 0.0471405 & 0.0471405 & \cdot \\ \cdot & 0.2 & 0.2 & \cdot & 0.239052 & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & 0.0471405 & \cdot & 0.031427 & 0.0471405 & \cdot \\ \cdot & \cdot & \cdot & 0.0471405 & \cdot & 0.0471405 & 0.031427 & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \end{pmatrix} \tag{3.31}$$

and the closest biseparable state from the algorithm,

$$\mathrm{CBiS}_2 = \begin{pmatrix} 0.0428141 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & 0.301005 & 0.169739 & \cdot & 0.169198 & \cdot & \cdot & \cdot \\ \cdot & 0.169739 & 0.302724 & \cdot & 0.168604 & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & 0.0174193 & \cdot & 0.00675952 & 0.00726677 & \cdot \\ \cdot & 0.169198 & 0.168604 & \cdot & 0.301507 & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & 0.00675952 & \cdot & 0.0167853 & 0.00697841 & \cdot \\ \cdot & \cdot & \cdot & 0.00726677 & \cdot & 0.00697841 & 0.0174937 & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \end{pmatrix} \tag{3.32}$$

The distances from the W state for CBiS_1 and CBiS_2 are 0.26202 and 0.16769. Both the states have the same structure, and we again see the contribution from vectors with 2 excitations. Another interesting observation that can be made is that the set of eigenvectors for both CBiS_1 and CBiS_2 are the same as the eigenvectors of the state $\frac{1}{2}(|W\rangle\langle W| + \sigma_x^{\otimes 3} |W\rangle\langle W| (\sigma_x^{\otimes 3})^\dagger)$.

3.5 SPECIAL CLASSES OF STATES

In this section, we provide the results from running the algorithm for a few special classes of states like the generalized GHZ states and convex combination of 3-qubit GHZ and W states.

3.5.1 *Generalized GHZ states*

The N-qubit generalized GHZ states are defined as,

$$|GGHZ\rangle_N = \cos\theta\,|00\ldots0\rangle + \sin\theta\,|11\ldots1\rangle\,, \tag{3.33}$$

where $\theta \in [0, 2\pi]$. We will only consider $N = 3$ in this section for brevity, but as has been the trend with N-qubit GHZ states, we fully expect observations to scale to $N > 3$. The first observation to make is that varying θ traces a smooth path on the hypersurface δB^{63}. This path goes through the following pure states for particular values of θ:

$\theta = 0$ $\quad |000\rangle$

$\theta = \frac{\pi}{4}$ $\quad \frac{1}{\sqrt{2}}(|000\rangle + |111\rangle) = |GHZ\rangle$

$\theta = \frac{\pi}{2}$ $\quad |111\rangle$

$\theta = \frac{3\pi}{4}$ $\quad \frac{1}{\sqrt{2}}(|000\rangle - |111\rangle) =: |GHZ^-\rangle$

$\theta = \pi$ $\quad -|000\rangle$

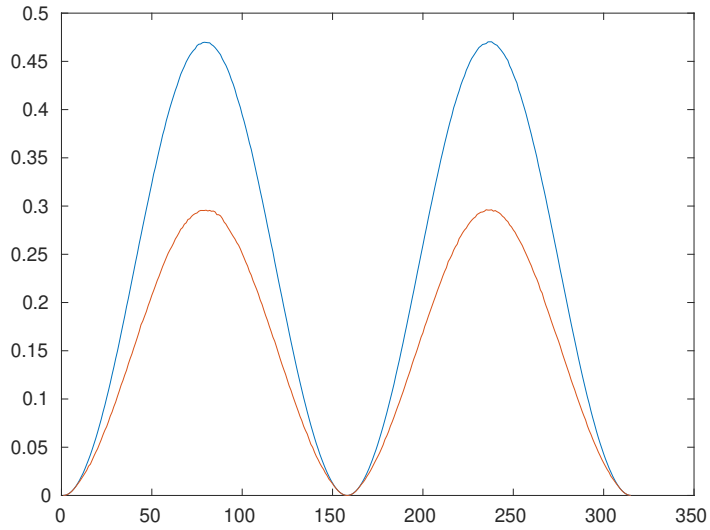


Figure 3.7: Minimum Hilbert-Schmidt distance of the generalized GHZ states as a function of θ. In blue is the Hilbert-Schmidt distance from the set of separable states and Hilbert-Schmidt distance from the set of biseparable states is in red. The plot starts at $|000\rangle$ and then the first pair of peaks is for $|GHZ\rangle$ and the second for $|GHZ\rangle^{-}$.

From $\theta \in [\pi, 2\pi]$ the states have additional global phase that does not affect the results in any manner. Therefore it is sufficient to take $\theta \in [0, \pi]$ for our consideration.

We ran the algorithm for this interval of θ, with the increments of 0.01, for a total of 314 states. The algorithm ran on both separable and biseparable sets. The resulting minimum distances from the set of separable and biseparable states is plotted in the Figure 3.7. The first pair peak corresponds to $|GHZ\rangle$ where we have already seen the analysis of the closest separable and biseparable states. The second pair of peaks correspond to the state $|GHZ^{-}\rangle$, for which the findings are exactly the same as for GHZ state but with appropriately placed minus signs. The closest separable state ρ^{-}_{CSS} is now a convex combination, $\rho^{-}_{\mathrm{CSS}} = t_N \varrho_N + (1 - t_N)\Delta^{-}_N$ with the same $t_N = (2^N - 2)^2(4 + 4^N - 2^{N+1})^{-1}$. The difference is in the matrix Δ^{-}_N which is equal to Δ_N in Equation 3.9, except on the anti-diagonal corners where it has minus signs.

Similarly the closest biseparable state is the generalized Werner state with visibility 3/7. Due to this symmetry in properties of the two states, we know a portion of the boundaries of the sets of separable and biseparable states. To

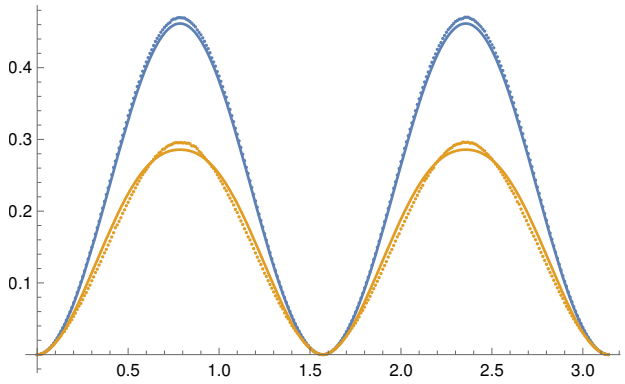


Figure 3.8: To check correspondence of closest separable and biseparable states found using geometry and the algorithm, we use convex combinations of the the vertices of separable face and biseparable face and minimize the Hilbert-Schmidt distance from the state $|GGHZ\rangle$. While the closest separable states all lie on the separable face that we know, the closest biseparable states do not. While transitioning from entangled to separable pure state, the path that θ traces derives its minimum Hilbert-Schmidt distance from an adjacent plane. Solid lines denote distance minimized from convex combinations. Blue denotes minimum distance from separable states and yellow from the set of biseparable states.

check if the closest separable and closest biseparable states for varying values of θ lie on these faces, we could minimize the distance of the $|GGHZ\rangle$ states to the convex combinations of the endpoints of the red and blue planes in Figure 3.6 for $\theta \in [0, \pi/2]$ and from the endpoints of the corresponding planes with respect to $|GHZ\rangle^{-}$, for $\theta \in (\pi/2, \pi]$.

While the closest separable states lie throughout the range of θ on the faces of the set of separable states that we indicate in blue in Figure 3.6, the closest biseparable states that the algorithm finds are closer to the reference states for a small range of θ while it transitions from $|GHZ\rangle$ to the separable pure states and back. See Figure 3.8. The conclusion we can draw from this is that the perpendicular dropped from the path traced by θ does not lie on the particular face of the biseparable set, but on the adjacent face.

We also compared the closest biseparable states found by the algorithm to the ones produced by Verstraete's method, and the findings are consistent with previous observations. The algorithm provides a better minimum distance for the range of θ in the vicinity of $|GHZ\rangle$ and $|GHZ^{-}\rangle$, except when the state tends towards the pure separable states. In that case, the lower distance using Verstraete's method might indicate the presence of PPT entangled states, although

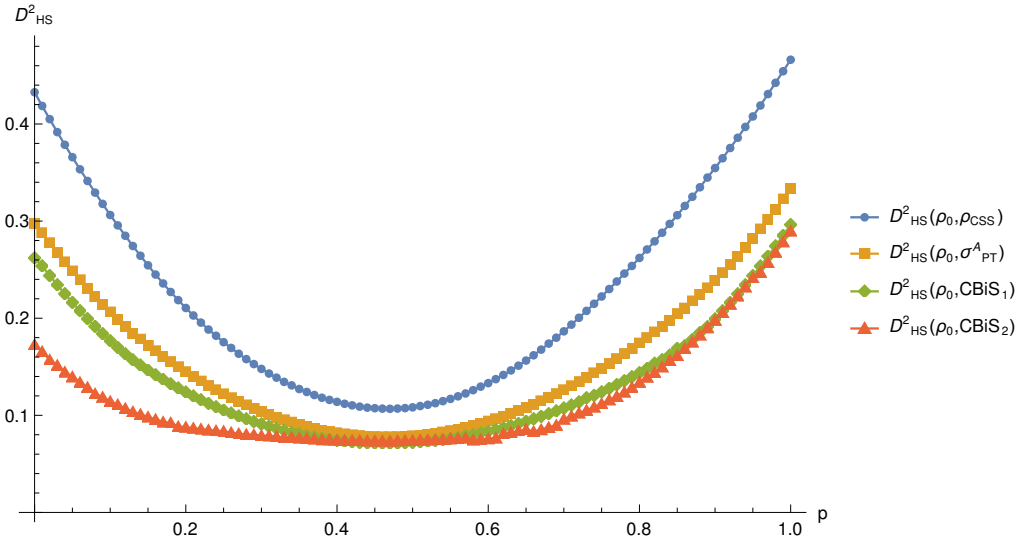


Figure 3.9: The plot shows the minimum distances of the GHZ-W mixture from the set of fully separable (Blue) and biseparable states (Red). In the middle of the two lie the distances from the closest biseparable states found using Verstraete's method, namely, the distance from the closest PPT state in any bipartition (Orange) and the distance from the convex combination of the three closest PPT states (Green).

a quick check of the minimum eigenvalue of the partial transpose dispels this notion. The remaining reason would be not enough iterations of the algorithm and also that the algorithm takes more iterations to converge in the vicinity of pure product states, due to the vertex of $\mathcal{S}$ being close by the reference state.

3.5.2 *GHZ-W line*

We considered the convex combination of the 3 qubit GHZ and W states,

$$\rho = p\,|GHZ\rangle\langle GHZ| + (1-p)\,|W\rangle\langle W|\,, \tag{3.34}$$

and ran the algorithm to find the minimum distance from the set of fully separable and biseparable states. Figure 3.9 illustrates these distances for 101 values of $p \in [0,1]$. We also compared the output from the algorithm to closest biseparable states found using Verstraete's method, to again find that while the latter finds the boundary of the set of biseparable states, it does not provide the overall closest biseparable state near the maximally entangled endpoints. Further study is required to ascertain why the convex combination of closest bisepa-

rable states in each bipartition does not result in the overall closest separable state, even though they are arranged symmetrically about the reference state. On the other hand, in the middle of the interval from about $p = 0.4$ to $p = 0.5$, Verstraete's method gives a slightly better distance. If we check the minimum Eigenvalues of the partial transpose in this range, we do not find any PPT states, indicating again that more iterations would surpass this slight difference.

On the whole, the line from GHZ to W always lies above the boundaries of both separable and biseparable state, but is closer to the boundary of the biseparable states on the W side then on the GHZ side. Using the separability conditions for biseparability introduced in [23], one can verify that the whole GHZ-W line is not biseparable for any value of p. In the middle of the plot from about $p = 0.35$ to $p = 0.6$ the distance from the biseparable states is almost constant indicating that the GHZ-W line might be parallel to the boundary in this region.

3.6 SUMMARY

In the chapter we discussed the application of simplified Gilbert's algorithm in finding the minimum Hilbert-Schmidt distance from the set of separable states and the closest separable state. We verified that the algorithm gives good results by comparing it's output to previously known results from some well known classes of states. In the process, we saw how to glean out of it, new insights about the geometry of set of separable states, especially in the case of GHZ states. We also looked at the methods to find the closest biseparable state, and proved yet again that the algorithm is capable of delivering analytical insights. We also showed that the output of the algorithm provides the true closest biseparable state, while verifying the results of Vertraete's method in individual bipartitions.

In the next chapter, we will discuss another important application of the Gilbert's algorithm, namely, the construction of Entanglement Witnesses.

4

CONSTRUCTING ENTANGLEMENT WITNESSES

In Chapter 1 we had a brief introduction to the concept of Entanglement Witnesses and their intimate relation with the geometry of the set of separable states, $\mathcal{S}$. In terms of the optimization problems discussed in Section 2.1, Entanglement Witnesses provide a solution to the `WSEP` problem, by providing a separating hyperplane. In the Hilbert space, an Entanglement witness is represented by a hermitian operator, such that it has at least one negative eigenvalue and has a positive expectation over all separable density matrices [51]. The existence of an Entanglement Witness for a given entangled states is guaranteed by the Hahn-Banach theorem due to the convexity of $\mathcal{S}$.

The importance of Entanglement Witnesses comes to light when we want to certify entanglement experimentally. In the case of qubits an Entanglement Witness being an hermitian operator it can be written in Pauli basis and then is easy to implement in a laboratory setting. There are other tasks, in addition, that help in the experimental setting, where one can try to minimize the number of measurements one needs to perform for implementing an Entanglement Witness, or finding a witness for a state that requires the least number of measurements although such an Entanglement Witness might not be optimal.

4.1 INTUITION BEHIND ENTANGLEMENT WITNESSES

Let us illustrate this connection between the hermitian operator and a separating hyperplane. A hyperplane with the normal vector $\mathbf{n}$ in $\mathbb{R}^n$ divides it into two halfspaces. If the normal vector has the base at point $\mathbf{p}$, then for every point $\mathbf{x}$ in the halfspace that is in the direction of the normal $\mathbf{n}$, satisfies $(\mathbf{x}-\mathbf{p})\cdot\mathbf{n} > 0$ and every point $\mathbf{y}$ in the other halfspace satisfies $(\mathbf{y}-\mathbf{p})\cdot\mathbf{n} < 0$. Only the points $\mathbf{z}$ lying on the plane satisfy $(\mathbf{z}-\mathbf{p})\cdot\mathbf{n} = 0$ (Figure 4.1). This intuition translates exactly to the set of quantum states, $\mathcal{Q}$. The set $\mathcal{S}$ lies inside $\mathcal{Q}$ and both are convex. Therefore, one can draw a hyperplane that cuts through $\mathcal{Q}$ such that $\mathcal{S}$ is on one side and the rest on the other side of the hyperplane. Then, if the normal of the hyperplane, W, is defined to be in the half space with $\mathcal{S}$, we get

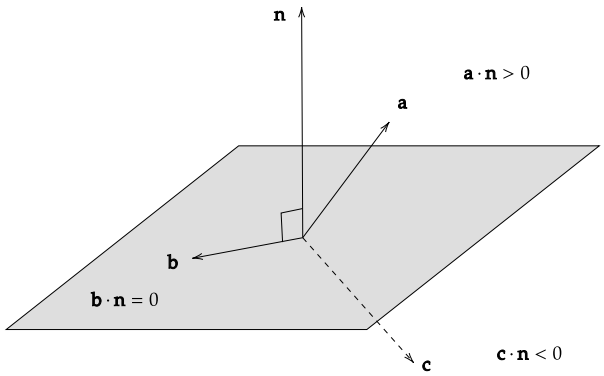


Figure 4.1: A hyperplane divides the space in two halfspaces such that the dot product of the vectors on the side where the normal (**n**) lies is positive and negative in the opposite halfspace. The plane is characterised by all vectors **b** such that $\mathbf{b} \cdot \mathbf{n} = 0$. The vectors are relative to the base point of the normal.

$\mathrm{Tr}(W\sigma) > 0$ for all $\sigma \in \mathcal{S}$, and for some $\rho \in \mathcal{Q} - \mathcal{S}$ that lie on the other side, $\mathrm{Tr}(W\rho) < 0$. The operator W is then said to witness the entanglement of such a ρ and states in its immediate neighborhood. In this picture, it has satisfied both the properties, positive expectation on all separable states and negative on subset of the entangled states. Intuitively, if the hyperplane cutting through $\mathcal{Q}$ is moved closer to the set of separable states, the space of entangled states that is witnessed by the hyperplane grows larger. Clearly, the space of entangled states witnessed will be the largest when the hyperplane becomes tangential to the set $\mathcal{S}$. When an entanglement witness W_2 detects more entangled states than another witness W_1, then W_2 is called a *finer* witness. Formally, the space of entangled states detected by W_1 is a subset of the space of entangled states detected by W_1, i. e., $D_{W_1} \subseteq D_{W_2}$, then W_2 is a *finer* witness.

So follows the definition of an *Optimal Entanglement Witness*. A witness W_{opt} is called optimal if and only if there is no other entanglement witness finer than it [10, 11, 30].

4.2 SOME PROPERTIES OF ENTANGLEMENT WITNESSES

We will briefly mention some properties and relations for Entanglement Witnesses that will help us form Optimal Entanglement witnesses using Gilbert's algorithm.

- For two entanglement witnesses, $D_{W_1} = D_{W_2}$ if and only if $W_1 = W_2$. Only equal hyperplanes can detect the same halfspace.
- If W_2 is finer than W_1 then their difference is a positive operator, P, i. e., for some $\epsilon > 0$, we can write $W_1 = (1-\epsilon)W_2 + \epsilon P$
- Using the above, a witness W_2 is *optimal*, if and only if for all $P > 0$ and $\epsilon > 0$, then $W_1 = (1-\epsilon)W_2 + \epsilon P$ is not an Entanglement Witness.
- The two properties of an optimal Entanglement Witness: $Tr(\sigma W_{\text{opt}}) \geq 0 \, \forall \sigma \in \mathcal{S}$ and that there is no other finer witness, restrict the position of the Entanglement Witness as a tangent to the set of separable states. If W_{opt} is not a tangent to $\mathcal{S}$, there will always exist a finer witness, and therefore, it is not optimal.
- On the other hand, if the hyperplane W_1 cuts across $\mathcal{S}$, then it does not satisfy the property that $Tr(\sigma W_1) \geq 0 \, \forall \sigma \in \mathcal{S}$, but it still witnesses an entangled state ρ as $\text{Tr}(\rho W_1) < 0$. See Figure 4.2. Like we moved a non-optimal Entanglement Witness to a finer witness by subtracting a positive operator, here we can do the opposite and add positive operators to move the hyperplane towards the boundary of $\mathcal{S}$ and make it optimal.
- Finding such positive operators to add to or subtract from an hyperplane to make it an optimal Entanglement Witness constitutes the optimization problem `WOPT` from Chapter 2, and therefore, is hard to do. Convexity of $\mathcal{S}$ again makes it so that we only have to search in the space of pure product vectors $\mathcal{S}_{\text{pure}}$.
- If there is a set of pure product states P_W such that $\text{Tr}\left(|\psi_i\rangle\langle\psi_i| \, W_{\text{opt}}\right) = 0$, $\forall \; |\psi_i\rangle \in P_W$ and the state ρ is detected by the optimal witness W_{opt} then so are the states $\rho + \sum_i p_i \, |\psi_i\rangle\langle\psi_i|$. The set P_W of product vectors characterises the face of the set $\mathcal{S}$ on which the the witness W_{opt} is tangent.

4.3 RELATION BETWEEN HILBERT-SCHMIDT DISTANCE AND ENTANGLEMENT WITNESSES

In [10] the authors explored the relation between the minimum Hilbert-Schmidt distance from $\mathcal{S}$ and the Optimal Entanglement Witness for a given state ρ. To

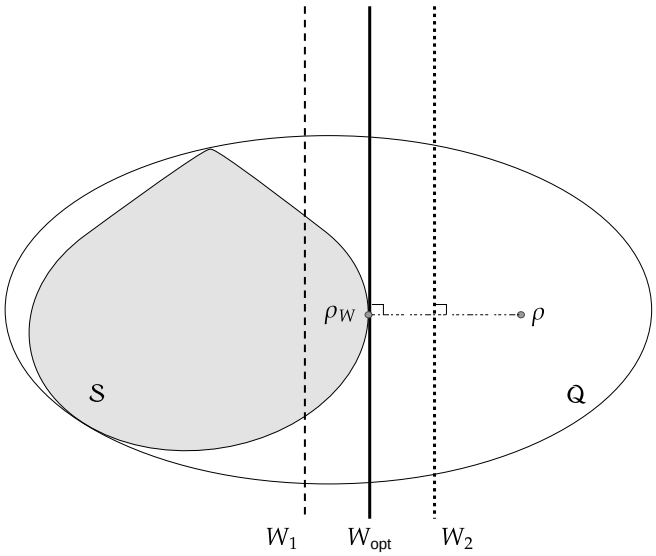


Figure 4.2: All three hyperplanes, W_1, W_{opt}, W_2, witness the state ρ, but W_1 fails to have a positive expectation over all states in $\mathcal{S}$, as some lie on the same side as ρ. The hyperplane W_2, on the other hand, is an Entanglement Witness, albeit not an optimal one. If we move the hyperplane away from ρ by subtracting a positive operator, it becomes a finer Entanglement Witness, until it becomes W_{opt}. We have $\text{Tr}(\rho_W W_{\text{opt}}) = 0$ as W_{opt} is tangent to $\mathcal{S}$ at ρ_W. Similar to how we made W_2 finer, we can move the hyperplane W_1 by adding positive operators to it till it becomes optimal.

do so they formulate Entanglement witnesses to be Generalized Bell inequalities. Generalized in the sense that they detect all entanglement, including PPT entanglement which Bell inequalities do not. We can take the example of the Clauser-Horne-Shimony-Holt (CHSH) inequality to see this similarity by writing it in the form $2\mathbb{1} - \mathcal{B}$, where $\mathcal{B}$ is the Bell operator, then for all separable and PPT states σ,

$$\text{Tr}(\sigma(2\mathbb{1} - \mathcal{B})) \geq 0. \tag{4.1}$$

Bell inequalities thus can be written in a form that is similar to non-optimal Entanglement witnesses, in the sense that the entangled states that violate the Bell inequality have a negative expectation value with the modified Bell opera-

tor and a positive expectation value on the rest. By formulating the following generalized Bell inequality for an entangled state ρ,

$$B(\rho) = \max_{W} \left(\min_{\sigma \in \mathcal{S}} \mathrm{Tr}(\sigma W) - \mathrm{Tr}(\rho W) \right) \tag{4.2}$$

where the maximum is taken over all feasible Entanglement Witnesses, W, and the minimum over all separable states, we have the following theorem known as the Bertlmann-Narnhofer-Thirring Theorem,

- `Bertlmann-Narnhofer-Thirring Theorem` The minimum Hilbert-Schmidt distance of a state ρ from the set $\mathcal{S}$ is equal to the maximal violation of the generalized Bell inequality,

$$D_{\mathrm{HS}_{\min}}(\rho) = B(\rho) = \mathrm{Tr}(\rho_{\mathrm{CSS}} W) - \mathrm{Tr}(\rho W). \tag{4.3}$$

Note that for all separable states the value of $B(\sigma) = 0 = D_{\mathrm{HS}_{\min}}(\sigma)$, and such a min-max problem is very hard to solve computationally over the set of separable states.

We'll discuss the proof in the geometric sense. The Entanglement Witness W has the expectation $\mathrm{Tr}(\sigma W)$ over $\mathcal{S}$. The minimum value of this expectation is $\min_{\sigma \in \mathcal{S}} \mathrm{Tr}(\sigma W) = \mathrm{Tr}(\sigma_W W) = 0$ if the witness W is tangential to $\mathcal{S}$ at the point σ_W. Therefore, the minimization term in Equation 4.2 is zero if we take tangential hyperplanes as witnesses. Then the maximal value of $B(\rho)$ depends on $\mathrm{Tr}(\rho W)$ and that is maximum when the normal vector is parallel to the vector $\sigma_W - \rho$. The Entanglement Witness with the normal $\sigma_W - \rho$ is a valid Entanglement Witness with positive expectation on all separable states if and only if $\sigma_W = \rho_{\mathrm{CSS}}$ is the closest separable state to ρ under the Hilbert-Schmidt measure.

We can conclude from the above analysis, that for a given state ρ the Optimal Entanglement Witness is the hyperplane tangent at the state ρ_{CSS}. This is where the significance of finding the Closest Separable states using Gilbert's algorithm becomes clear. In theory, *we are now able to provide Optimal Entanglement Witnesses for any given entangled state ρ by employing the Gilbert's algorithm to calculate ρ_{CSS}*. Then the Optimal Entanglement Witness is simple to formulate,

$$W = \frac{\rho_{\mathrm{CSS}} - \rho - Tr(\rho_{\mathrm{CSS}}(\rho_{\mathrm{CSS}} - \rho))\mathbb{1}}{\|\rho_{\mathrm{CSS}} - \rho\|_{HS}} \tag{4.4}$$

and has the requisite properties, $\mathrm{Tr}(\sigma W) \geq 0$ for all $\sigma \in \mathcal{S}$, $\mathrm{Tr}(\rho_{\mathrm{CSS}} W) = 0$ and $\mathrm{Tr}(\rho W) < 0$.

4.4 OPTIMIZING ENTANGLEMENT WITNESSES

We have discussed in Section 4.2, that when it comes to optimizing Entanglement Witnesses, there are two possibilities, first, pushing the Entanglement Witness towards the set $\mathcal{S}$ by subtracting positive operators, or second, pulling a hyperplane out to the boundary of $\mathcal{S}$ by adding positive operators. This is the function of the term $Tr(\rho_{\mathrm{CSS}}(\rho_{\mathrm{CSS}} - \rho))\mathbb{1}$ in Equation 4.4. It pushes the hyperplane radially out from the center of the hyperball if it is added and vice versa. In this work, we only need to deal with the second type, because of the fact that the Gilbert's algorithm approximates the Closest Separable state from inside the set $\mathcal{S}$ and reaches the vicinity of the boundary of the set. For this reason, if an Entanglement Witness was written in terms of its approximate CSS like so,

$$\tilde{W} = \tilde{\rho}_{\mathrm{CSS}} - \rho - Tr(\tilde{\rho}_{\mathrm{CSS}}(\tilde{\rho}_{\mathrm{CSS}} - \rho))\mathbb{1} \tag{4.5}$$

then although the hyperplane $\tilde{W}$ goes through $\tilde{\rho}_{\mathrm{CSS}}$ and witnesses ρ it is still not an Entanglement Witness, because there are separable states between it and the boundary of set $\mathcal{S}$ that have a negative expectation value. Also, because we are using an approximate CSS to form this hyperplane, it can be visualized to have a slight tilt with respect to the optimal Entanglement Witness at the exact CSS.

This is where the optimization part comes in. As the hyperplane currently cuts through $\mathcal{S}$, there are product vectors $|\psi\rangle$ such that $\mathrm{Tr}\big(\tilde{W}\,|\psi\rangle\langle\psi|\big) < 0$. We already know that it is enough to optimize this over the pure product states due to convexity of $\mathcal{S}$. The goal of the optimization problem now is to find $|\psi\rangle$ such that $\mathrm{Tr}\big(\tilde{W}\,|\psi\rangle\langle\psi|\big)$ is minimum. Once we have that, we can update the operator $\tilde{W}$,

$$\tilde{W} = \tilde{\rho}_{\mathrm{CSS}} - \rho + Tr(|\psi\rangle\langle\psi|\,(\tilde{\rho}_{\mathrm{CSS}} - \rho))\mathbb{1} \tag{4.6}$$

which is now a valid Entanglement Witness because we have ensured that the expectation of $\tilde{W}$ is positive over all product states and in turn their convex combinations with them.

This optimization problem is the WOPT problem and thus hard to solve efficiently. Solving it needs a characterization of the set, or some sort of parameterization, as we saw in former chapters. Although it grows exponentially, it is easy enough to formulate and solve in the lower dimensions and multi-qubit systems. Compared to the mixed states, parameterizing pure states and optimizing is a relatively easier task. For instance, in the case of multi-qubit systems, each qubit can be parameterized as

A qudit can be parameterized using d parameters.

$$|\theta_i, \phi_i\rangle = \begin{pmatrix} \cos\theta_i \\ e^{\iota\phi_i}\sin\theta_i \end{pmatrix} \quad \text{where} \quad i = 1, 2 \ldots, N. \tag{4.7}$$

One can then also employ gradient descent algorithms, which work well up to a point. They have a tendency of getting stuck in local minima as the parameter space is periodic, and as the parameter space grows, the probability of finding the global minimum decreases. The other approach is to randomly draw states from a Haar uniform distribution and calculate the expectation for each and pick the maximum. This suffers from the drawback that there is no way to ensure that the maximum reached is optimum. On the other hand, the convexity of the pure product states guarantees that if a local maximum is found it is also the global maximum. Let us move on to some examples and demonstrations.

4.5 ENTANGLEMENT WITNESS USING CLOSEST SEPARABLE STATE

We shall see examples of the construction of Entanglement Witnesses that we discussed in the previous section. There are certain classes of states with known analytic closest separable states and in those cases it is trivial to construct an Entanglement Witness. We will first look at the classes that were discussed in Chapter 3 and then we will discuss two classes of PPT entangled states and construction of Entanglement Witnesses for them.

4.5.1 *Bipartite Maximally Entangled states*

The bipartite maximally entangled states are defined as,

$$|\psi_d\rangle = \frac{1}{\sqrt{d}} \sum_{i=0}^{d} |i,i\rangle \,. \tag{4.8}$$

and their closest separable states are known to be,

$$\rho_{\text{CSS}} = \left(\frac{1}{d+1}\right) |\psi_d\rangle\langle\psi_d| + \left(\frac{d}{d+1}\right) \frac{\mathbb{1}}{d}, \quad 0 \leq p \leq 1. \tag{4.9}$$

Now we can define the Entanglement Witness according to the Equation 4.4 with ignoring the normalization,

$$W_d = \rho_{\text{CSS}} - |\psi_d\rangle\langle\psi_d| - Tr(\rho_{\text{CSS}}(\rho_{\text{CSS}} - |\psi_d\rangle\langle\psi_d|))\mathbb{1} \tag{4.10}$$

Substituting the expressions from above gives us,

$$W_d = \frac{d}{d+1}\left(\frac{2}{d+1}\mathbb{1} - |\psi_d\rangle\langle\psi_d|\right). \tag{4.11}$$

If we calculate the expectation of W_d with $|\psi_d\rangle\langle\psi_d|$ we get,

$$\text{Tr}(W_d\,|\psi_d\rangle\langle\psi_d|) = \frac{d(1-d)}{(d+1)^2} < 0. \tag{4.12}$$

So the operator detects the maximally entangled states, but to see if it's an optimal Entanglement Witness, we need to ensure that the expectation is zero at the closest separable state,

$$\text{Tr}(W_d\rho_{\text{CSS}}) = \frac{2d}{d+1}\left(\frac{d+1}{(d+1)^2} - \frac{1}{(d+1)}\right) = 0. \tag{4.13}$$

Therefore, the operator W_d is an optimal Entanglement Witness for the bipartite maximally entangled states of d dimensions. Furthermore, W_d is an optimal Entanglement Witness for all the states on the line defined by the d-dimensional Werner states for visibility $1/3 \leq p \leq 1$.

4.5.2 *N-Qubit GHZ states*

Now that we have $N > 2$, it is possible to provide Entanglement Witnesses that witness multiparty entanglement. For example in the 3 qubit case, an Entanglement Witness that is tangent to the set of fully separable states will detect entanglement with respect to all bipartitions as well as genuine multiparty entanglement.

We follow the same procedure here as for the previous class because we know the analytical closest fully separable state. Additionally, we will construct the Entanglement Witness that detects genuine multiparty entanglement using the closest biseparable state we found using the Gilbert's algorithm.

4.5.2.1 *Fully separable case*

The closest separable state for N-qubit GHZ states was defined as the convex combination $\rho_{\text{CSS}} = p\varrho_N + (1-p)\Delta_N$, where $p = (2^N-2)^2(4+4^N-2^{N+1})^{-1}$, in Equation 3.9.

The optimal Entanglement Witness constructed using Equation 4.4 for $N = 3$ reads,

$$W = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & -\frac{6}{13} \\ 0 & \frac{2}{13} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{2}{13} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{2}{13} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{2}{13} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \frac{2}{13} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \frac{2}{13} & 0 \\ -\frac{6}{13} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \tag{4.14}$$

and satisfies all the optimality conditions. Similar construction follows in $N > 3$ qubit GHZ states.

4.5.2.2 *Biseparable case*

For the GHZ states we conjectured based on the geometrical insights provided by the output of the Gilbert's algorithm that the closest biseparable state is the Werner state for GHZ (mixture with Identity) with the visibility $p_c = (1 - 0.5(1-2^{-N})^{-1})$. To recap, the GHZ states are defined as

$$|\text{GHZ}_N\rangle = \frac{1}{\sqrt{2}}\left(|0\rangle^{\otimes N} + |1\rangle^{\otimes N}\right) \tag{4.15}$$

and the closest biseparable state,

$$\rho_{\text{CSS}} = p_c \left|\text{GHZ}_N\right\rangle\!\left\langle\text{GHZ}_N\right| + (1 - p_c)\frac{\mathbb{1}}{2^N}. \tag{4.16}$$

Then again the Witness operator turns out to be,

$$W = (1 - p_c)\left(\frac{1 - p_c(1 - 2^N)}{2^N}\mathbb{1} - \left|\text{GHZ}_N\right\rangle\!\left\langle\text{GHZ}_N\right|\right), \tag{4.17}$$

where $p_c = (1 - 0.5(1 - 2^{-N})^{-1})$. It reduces to, $W = \frac{4}{7}(0.5\mathbb{1} - |\text{GHZ}_3\rangle\langle\text{GHZ}_3|)$ for $N = 3$. It is then easy to check, $\text{Tr}(W.\rho_{\text{CSS}}) = 0$ and if increase the visibility $p = p_c + \epsilon$ by a small amount $\epsilon > 0$, the expectation of the Entanglement Witness immediately changes sign. Also, $\text{Tr}(W\left|\text{GHZ}_N\right\rangle\!\left\langle\text{GHZ}_N\right|) = -\frac{2}{7}$. The closest PPT states calculated by Verstraete's method σ_{PT} all lie on this hyperplane and have $\text{Tr}(W\sigma_{\text{PT}}) = 0$. This verifies our findings about the geometry of the boundary of the biseparable states near GHZ (Figure 3.6).

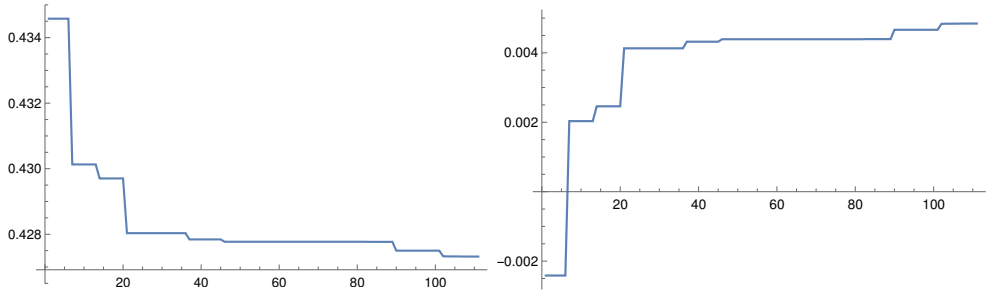
4.5.3 *W states*

Till now the optimization of the Entanglement Witnesses was not necessary as we more or less knew the closest separable states (or biseparable in case of GHZ). For W states we do not have such knowledge, instead we have a very good approximation of the closest separable state. We will, only when talking about the W states, change the notation for an Entanglement Witness, and denote it with Λ instead.

First we create the hyperplane passing through the approximate CSS denoted by $\tilde{\rho}_{\text{CSS}}$,

$$\tilde{\Lambda} = \tilde{\rho}_{\text{CSS}} - |W\rangle\langle W| - \text{Tr}(\tilde{\rho}_{\text{CSS}}(\tilde{\rho}_{\text{CSS}} - |W\rangle\langle W|))\mathbb{1}, \tag{4.18}$$

The next task is to minimize $\text{Tr}\left(\tilde{\Lambda}\left|\psi\right\rangle\!\left\langle\psi\right|\right)$ over product vectors $|\psi\rangle$. We use the parameterization from Equation 4.7 for the three subsystems, which makes total of 6 parameters, and pass it to an optimization method. The Figure 4.3a shows the distance of hyperplane as it is updated to move towards the boundary of the set $\mathcal{S}$ from the W state, which as expected decreases. On the other hand, Figure 4.3b shows the distance of hyperplane from the approximate CSS as it moves away from it.



(a) (b)

Figure 4.3: The plots showing the distance of the hyperplane $\tilde{\Lambda}$ as it is optimized to be on the boundary of $\mathcal{S}$ from (a) the W state, and (b) the $\tilde{\rho}_{\text{CSS}}$ (approximate CSS)

We find $\min \text{Tr}\left(\tilde{\Lambda}\,|\psi\rangle\langle\psi|\right)$ to be -0.00484365. We update $\tilde{\Lambda} \leftarrow \tilde{\Lambda} + 0.00484371\mathbb{1}$ and run the optimization again. The second optimization is to make sure that we have reached the optimum and we are not stuck in a local minima. The second run gave us a minimum $\min \text{Tr}\left(\tilde{\Lambda}\,|\psi\rangle\langle\psi|\right) = 7.88 \times 10^{-16}$, which is a positive value, and signifies the positive expectation of the optimized operator over all separable states. Consequentially, we are able to conclude that within computational precision the Entanglement Witness is optimal. Similarly, we were successfully able to optimize and obtain Entanglement Witnesses for $N = 4, 5$ qubit W states.

This amounts to a solution to the `WOPT` *problem.*

4.5.4 *Generalized GHZ states*

In a previous section we defined the Entanglement Witnesses for GHZ state using the closest separable and biseparable states. The Entanglement Witnesses, as we know, correspond to the faces of the separable and the biseparable set, and therefore, are optimal for all the states that have their closest separable states and closest biseparable states on those faces. Therefore, for the class of generalized GHZ states ρ_θ, the Entanglement Witnesses for $|GHZ\rangle$ are optimal for $\theta \in [0, \pi/2]$ and for $\theta \in [\pi/2, \pi]$ the Entanglement Witnesses for $|GHZ^-\rangle$ are optimal. See Figure 4.4.

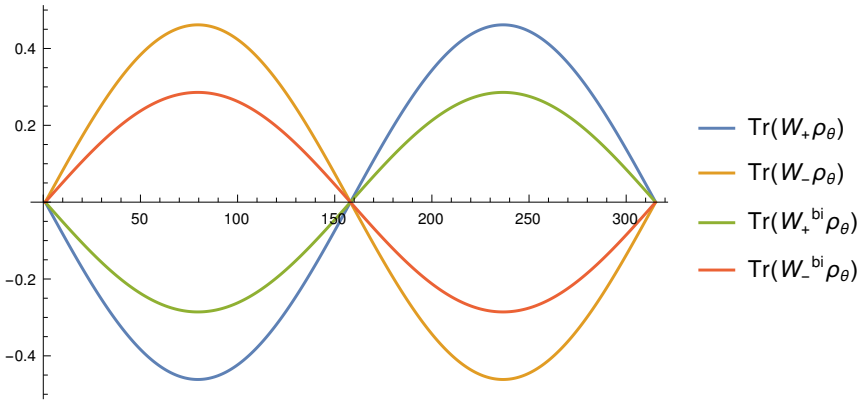


Figure 4.4: In the plot is shown the expectation values of the type $\mathrm{Tr}(W\rho_\theta)$ as a function of $\theta[\times 100]$. Here we denote the witness for $|GHZ\rangle$ from the closest separable state as W_+ and from closest biseparable state W_+^{bi}. Symmetrically define W_- and W_-^{bi} for $|GHZ^-\rangle$. This illustrates that the two faces of the set $\mathcal{S}$ witness the entanglement of the generalized GHZ states for the whole range of θ. Same holds true for the two faces of the set of biseparable sets, thus reaffirming our findings.

4.6 WITNESSING PPT ENTANGLEMENT

In Hilbert spaces of dimension $\dim \mathcal{H} > 6$ there are entangled states that have a positive partial transpose, such states are referred to as PPT entangled states. Their entanglement is special in the sense that it is very hard to detect using entanglement measures and Bell inequalities. The set of PPT states is convex and contains in it the set of all k-separable states. As a consequence of this geometry, from the point of view of applying the Gilbert's algorithm, the set of PPT states is no different from the set of entangled states as it is not reachable via convex combinations of pure state in k-separable set. In this manner, all PPT entangled states have a non-zero minimum distance from the set of separable states, and there always exists an Entanglement Witness that will witness its entanglement. See Figure 4.5 for an illustration. In what follows we will discuss special cases of PPT entanglement arising from particular constructions.

4.6.1 *Bound Entangled states from Unextendible Product Bases*

In [9], the authors presented the examples of *Unextendible Product basis* (UPB). A set of mutually orthogonal product vectors spanning a proper subspace $\mathcal{H}_{\mathrm{PB}}$ of the total Hilbert space $\mathcal{H}$ is called a *Product Basis* (PB). A product basis forms

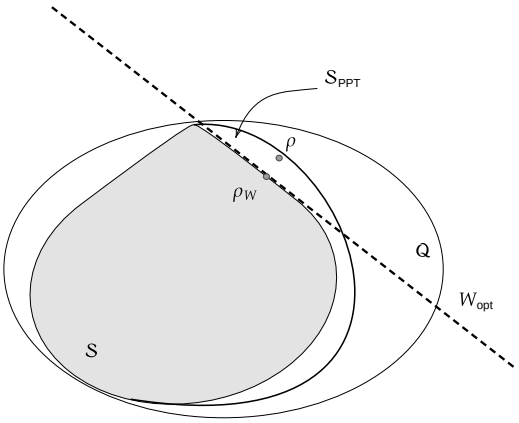
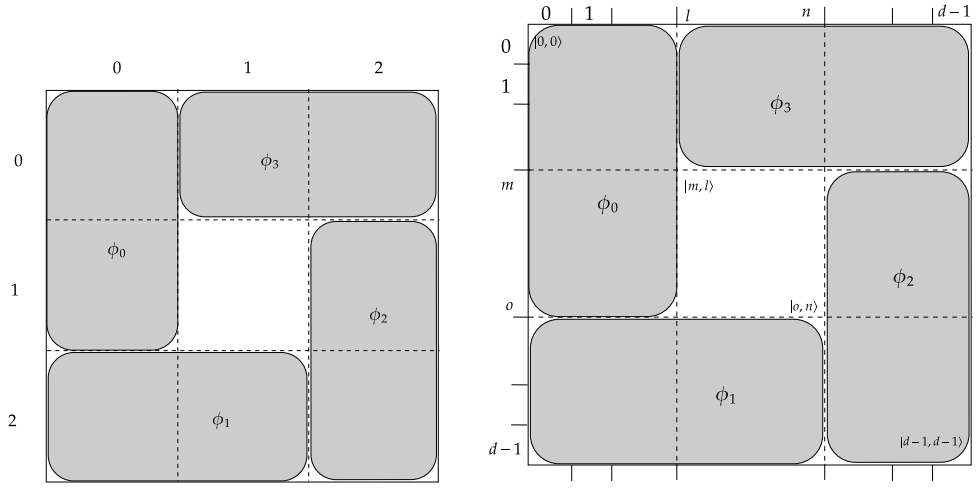


Figure 4.5: The PPT entangled states lie between the set of separable states and the set of entangled states that have a negative partial transposition. It is possible, again because of the convexity of the sets, to draw a hyperplane that separates a given PPT entangled state ρ from the set of separable states.

an unextedendible product basis if there can not be found any more product vectors orthogonal in $\mathcal{H}$ to all of the existing ones. Therefore, the subspace complementary to $\mathcal{H}_{\text{UPB}}$ does not have any product states, and a mixed state on this complementary subspace is always a *Bound entangled state.* While Bound Entangled states arise without the need of an UPB, creating a UPB provides a sufficient condition of the existence of such states. It was also proven in [16], that if we have a product vectors $|\phi_i\rangle \in \mathcal{H}_{\text{UPB}}$, $i = 1, 2 \ldots n$, then the following state is always a bound entangled state,

$$\rho_{\text{BE}} = \frac{1}{D-n}\left(\mathbb{1} - \sum_{i=1}^{n} |\phi_i\rangle\langle\phi_i|\right). \tag{4.19}$$

It is easy to check that ρ_{BE} is Bound Entangled, by partially transposing any subsystem. The identity operator in invariant under partial transposition and the product vectors just map onto another product vector thus preserving the positive-semidefiniteness of the density matrix. They also provided a bound on the number of product vectors n in the UPB, $n \geq \sum_j^N (d_j - 1) + 1$, where d_js are the dimensions of the N individual subsystems, and D is the dimension of the total Hilbert space.



(a) (b)

Figure 4.6: Visualization of the 2 qutrit UPB TILES in (a), and the generalization to the $d \times d$ case in (b) each state in the basis can be represented as individual tiles covering the space.

In [9] they give examples of three such constructions, but we are interested in the UPB called TILES. It is bipartite UPB in 3×3 dimension composed of the following states (visualized in Figure 4.6a),

$$
\begin{aligned}
|\phi_0\rangle &= \frac{1}{\sqrt{2}}\,|0\rangle\,(|0\rangle - |1\rangle), & |\phi_1\rangle &= \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)\,|2\rangle\,, \\
|\phi_2\rangle &= \frac{1}{\sqrt{2}}\,|2\rangle\,(|1\rangle - |2\rangle), & \phi_3 &= \frac{1}{\sqrt{2}}(|1\rangle - |2\rangle)\,|0\rangle\,, \\
|\phi_4\rangle &= \frac{1}{3}(|0\rangle + |1\rangle + |2\rangle)(|0\rangle + |1\rangle + |2\rangle) & & \qquad (4.20)
\end{aligned}
$$

Here $|\phi_i\rangle$ correspond to the tiles in Figure 4.6a with the exception of $|\phi_4\rangle$ which is also known as the stopper state, and it's function is to force the unextendibility of this product basis. This construction can be generalized to the $d \times d$ case [5]. We will only consider 5-tile UPBs and describe it's construction, as they are simple enough to construct and sufficient to prove that the Gilbert's algorithm is able to provide non-decomposable witnesses that detect PPT entanglement. First we choose a factorizable subspace in the Hilbert space, and from the projection on this subspace we remove the projector on the equal superposition of the basis states from the support of this subspace. The remaining

part defines a tile, e. g., in 3×3 a tile in the subspace $\{|1\rangle, |2\rangle\} \otimes \{|1\rangle, |2\rangle\}$ can be of the form,

$$\sum_{i,j=1,2} |i,j\rangle\langle i,j| - \frac{1}{2} \sum_{i,j,k,l=1}^{2} |i,j\rangle\langle k,l| . \tag{4.21}$$

Then the next thing to do is to find all such tiles so that there are no regions left that can be combined to form another tile. Once we have found all the tiles, the uniform superposition of all the states in the basis of the Hilbert space, Π_{sym} functions as the stopper state that completes the construction of the UPB. If we denote the projection on the subspace of a tile by Π_i and the projector of the equal superposition of the support of the tile's subspace by Π_i^{sym}, then the corresponding Bound Entangled state is formulated as

$$\rho_{\text{BE}} = \frac{1}{n}(\mathbb{1} - \Pi_{\text{sym}} - \sum_{i=1}^{n}(\Pi_i - \Pi_i^{\text{sym}})) \tag{4.22}$$

where n is the number of tiles in the UPB. The least number of tiles required to cover the whole Hilbert space is 5 in bipartite case and 9 in tripartite case. A 5-tile UPB with dimensions d_1 and d_2 can have the central tile of different sizes, and the number of such Bound Entangled states corresponding to the UPBs is given by $\frac{1}{4}(d_1^2 - 3d_1 + 1)(d_2^2 - 3d_2 + 1)$. If we assume $d_1 = d_2 = d$ then for $d = 3, 4, 5, 6$ we have the following number of Bound Entangled states:

3×3	1
4×4	9
5×5	36
6×6	100

That comes to a total of 146 Bound Entangled states. For these states we ran the Gilbert's algorithm and found the minimum Hilbert-Schmidt distance to be in the range 0.06 to 0.09, which signifies that they are indeed entangled, but barely so. The next step in the application would be to construct the approximate Witness operator and then optimize it. In [5], the authors provided a simple way to construct Entanglement Witnesses for theses states, that involves taking the projection over the support of the UPB and minimizing the overlap of this projector over the set of separable states. The Entanglement Witnesses so formed will be called the BGR witnesses.

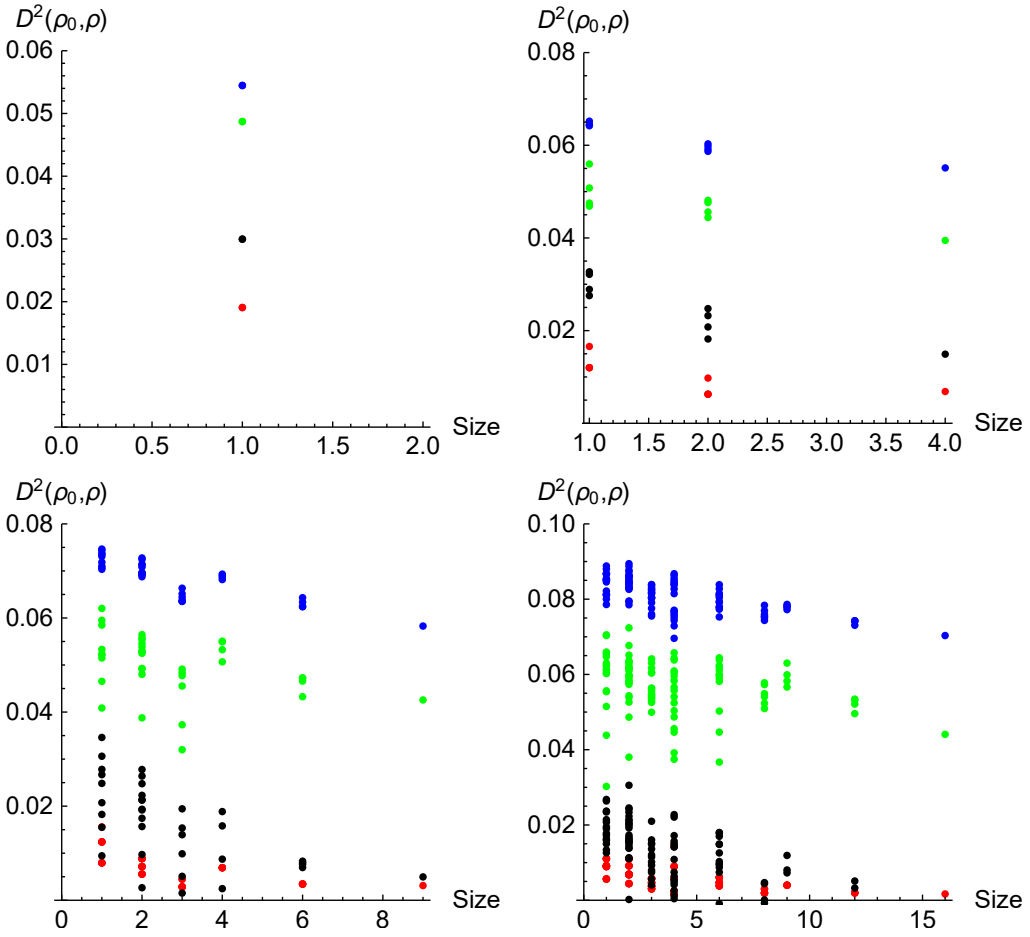


Figure 4.7: For the 146 Bound Entangled states from UPB in dimensions $d = 3, 4, 5, 6$ we compare the distance of the Entanglement Witnesses from the reference state, plotted together with the minimum Hilbert-Schmidt distance of the reference state to the set of separable states (blue) and the asymptotic minimum Hilbert-Schmidt distance obtained form linear model fitting (green). The distance from the BGR Entanglement Witnesses (red) is lower than the distance from the optimal Entanglement Witnesses found using Gilbert's algorithm, except in 4 cases in 5×5 and 6 cases in 6×6. Figure taken from our work [58]

Comparing two Entanglement Witnesses entails finding out which one is *finer* and as learnt before, an Entanglement Witness is finer if it detects a larger space of entangled states and the optimal Entanglement Witness is the finest possible. Therefore, to compare the Entanglement Witnesses formed using Gilbert's algorithm and the BGR witnesses, we compare the distance of the hyperplanes representing the Entanglement Witnesses from the reference states they are supposed to detect. A larger distance implies a finer Entanglement Witness. Now using the Gilbert's algorithm we have three such distances, 1) distance of the reference state from the approximate closest separable state, 2) Asymptotic minimum distance of the reference state to the set of separable states, 3) distance of the reference state from the optimized Entanglement Witness from Gilbert's algorithm. Add to this list, the distance of the reference state from the BGR witness as the fourth, and we compare these four in Figure 4.7. The distances 1) provides an upper bound, 2) provides a lower estimate of the distance, 3) provides the lower bound on the distance to the separable states. In Figure 4.7 1) is in blue, 2) is in green, 3) is in black and the fourth, the distance to BGR witnesses is in red. The four subplots correspond to the dimensions $d = 3, 4, 5, 6$ and the states are arranged based on their central tile size. Straightaway, we notice that in almost all cases the red points are below the black, implying that the optimized Entanglement Witnesses obtained from Gilbert's algorithm are finer, and as we have seen before, after optimization they are extremely close to being optimal Entanglement Witness.

4.7 SUMMARY

The geometric picture we developed in Chapter 2 is carried over to Chapter 3 to find the closest separable states, and then we saw how to construct Optimal Entanglement Witnesses for a given state using its closest separable state. We also showed that the Entanglement Witness for the reference state can only be optimal if it is a tangent to the set of separable states, precisely at the closest separable state, and we proved with examples that the Gilbert's algorithm is capable of giving us close to optimal Entanglement Witnesses. We say close to optimal, because the optimization problem involved is equivalent to the NP-HARD problem of `WOPT`, and the algorithms provide a result within some margin of precision. As such we can call such Entanglement Witnesses *weakly Optimal Entanglement Witnesses*, because their minimum expectation over the set

of separable states $0 \leq \min \mathrm{Tr}(W |\psi\rangle\langle\psi|) < \delta$, where δ is a very small positive number and $|\psi\rangle \in \mathcal{S}$.

5

SELF-TESTING GENUINE MULTIPARTY ENTANGLEMENT

We have concerned ourselves till this point, about how to certify a state that is known to us, as entangled or separable. We discussed the entanglement measures and criteria for separability that are a function of the state. We also discussed some algorithms that take the given state as input and as output provide information that can be used to conclude if a state is entangled or not. We learned that Entanglement Witnesses insofar as they detect entangled states, also detect PPT entangled states (or Bound Entangled states). The other point of view has to do with the measurement statistics of states and correlations arising therein. The correlations that manifest themselves in PPT entangled states can always be simulated using classical theories called the *Local Hidden Variable* theories. The states that exhibit correlations which cannot be explained or modeled by using Local Hidden variable theories are called *non-local* correlations.

Non-local correlations have been found to have numerous applications, most significantly in the field of Quantum Cryptography, where they enable secure Quantum Key Distribution scenarios. While Classical Cryptographic protocols rely on computational problems that are NP (non-deterministic polynomial time) for security (such as large prime factorization problem), the protocols in Quantum cryptography use the fundamental laws of physics to secure information against eavesdropping. BB84 (Bennett-Brassard-84) protocol [8] for quantum key distribution was the first such protocol proposed that used indistinguishability of non-orthogonal states to its advantage. While it has been proven to be insecure, it opened the gateway for further research in Quantum Cryptography. The usefulness of non-local correlations was highlighted when Ekert's protocol [20] was presented with its essential use of a bipartite maximally entangled state shared between the two parties trying to establish a private key. It provided a natural level of security using monogamy of entanglement [15, 37]. It became even clearer that non-local correlations are indispensable in such protocols with the introduction of *Device Independent* certification of entanglement in security proofs [1, 2]. Therefore, it is of interest in such a setting to detect entanglement that will be useful for the particular task, and

the requisite conditions are provided by the so called *Bell Inequalities*, first discussed in the celebrated paper dispelling local-realism, [7]. Considering for a moment, a bipartite scenario where two parties share a correlated state, the Bell inequalities provide a constraint on the statistics of the measurement outcomes performed by both the parties. The violation of this constraint is proof of non-local correlations shared between the two parties. While there are several ways to provide a security proof of a cryptographic protocol, we will focus on Device Independent certification and Self-Testing which is a stricter form of Device Independent certification, for the reason that they require minimal assumptions and therefore are applicable to a wide variety of scenarios.

5.1 DEVICE INDEPENDENCE CERTIFICATION

In a basic scenario in Quantum Key Distribution, the goal is to establish a secure private key among the parties, without letting a malicious outsider gaining knowledge about the key. Secure key distribution is the first step in establishing any kind of secure communication channel, and arguably also the most important one. The two parties, conventionally Alice and Bob, share a correlated state that is distributed to them by a source that they may or may not control. Subsequently they perform measurements on their subsystems and obtain outcomes on which they apply some predetermined transformation based on classical communication to obtain the private key.

In cryptography, for any such protocol, we require proofs of security that ensure that an eavesdropper, Eve, cannot glean information about the key, measurements and outcomes of the parties, or the state being generated by the source by performing measurements on the quantum states being sent via the shared quantum channel. Usually when giving a security proof, one assumes that the source distributing the correlated state is controlled by the parties, Alice and Bob. For some protocols, like BB84 (Bennett-Brassard-84) protocol [8], it is an absolutely essential assumption [1].

A security proof for a protocol that does not make the assumption that the source of the correlated state being distributed to the parties is trustworthy and essentially treats it like a black box, is called *Device Independent* [2]. The minimal assumptions for such a proof are that Quantum Mechanics is a valid theory, and that the parties participating in the protocol do not leak any information about measurements and outcomes via classical communication. The basis of device

independent security is that the parties receiving the distributed state upon collecting measurement statistics can find evidence of non-local correlations that cannot be reproduced by any local correlations between the parties and the eavesdropper. Let's understand this point with an simple example in the bipartite case. When we do not trust the source and do not know the underlying state coming from the source, all we have to rely on are the measurement statistics comprised of joint probabilities $P(a, b|x, y)$ which are commonly termed in the literature as *correlations*. Each time the source sends the state to Alice and Bob, they conduct measurements with settings x and y and obtain outcomes a and b respectively, and gather the results to form the joint probabilities. They can subsequently test the Clauser-Horne-Shimony-Holt inequality [14],

$$P(a_0 = b_0) + P(a_0 = b_1) + P(a_1 = b_0) + P(a_0 \neq b_0) \leq 3. \tag{5.1}$$

where a_0, a_1 are Alice's outcomes and b_0, b_1 are Bob's outcomes. If the correlations show violation of the above inequality, we have certified that the source is distributing non-local correlations, hence the eavesdropper must be uncorrelated from each of the parties. This follows from the *monogamy* of non-local correlations [6], which simply states that if two parties share non-local correlations between them then they are completely uncorrelated with any other third party.

Consequently, the Device Independent certification of non-local correlations shared among the parties participating in any kind of Quantum Cryptographic protocol ensures security of the protocol without any knowledge whatsoever of the underlying device generating the correlations.

5.2 SELF-TESTING AS A FORM OF DEVICE INDEPENDENCE

The goal in Device independent certification was to make sure that the parties actually share entanglement, without knowing anything about the apparatus. If we wanted to go one step further and in addition to verifying presence of entanglement, also wanted to know exactly the state produced by the source that we take as a black box, then we can only rely on the maximal violation of the Bell Inequality by the correlations [33, 34]. This verification of the state is called Device Independent Self-Testing of the state. Often such a maximal violation of the Bell Inequality also allows us to self-test the measurements used by the

black box, corresponding to the measurement settings and outputs of the parties. Note however, we get to know the state or the measurements up to local unitary transformations. As the authors point out in [50], the correlations are a result of applying the Born rule, which has no inverse as combinations of the state and measurements acting upon it are not unique for a particular conditional probability $P(a|b)$. Therefore, Self-Testing picks probability distributions from all the set of quantum probability distributions such that it has an inverse (unique combination of state and measurements), which only happens at the extreme point of the set of quantum correlations. We shall first go through an example in the bipartite case to set up the prerequisites for the multipartite scenario.

5.2.1 *Setting up Self-Testing in the bipartite scenario*

Given two party correlations $P(a, b|x, y)$ the task is to find a state $|\psi\rangle$ and measurement operators $\{M_A, M_B\}$ such that using the born rule,

$$P(a, b|x, y) = \mathrm{Tr}(|\psi\rangle\langle\psi| \, M_A \otimes M_B), \tag{5.2}$$

where the state $|\psi\rangle$ and the measurement operators $\{M_A, M_B\}$ are collectively called the *reference experiment*, that is equivalent up to *local isometric embeddings* to the underlying physical state and measurement operators of the source. This is known as the Mayers-Yao criterion [33]. We can assume that the shared state $|\psi\rangle$ is pure and the measurement operators are projective (i. e., $A^2 = \mathbb{1}$), because there is no characterization available of the total Hilbert space. Here local isometric embeddings mean there exists a local unitary that expands the Hilbert space locally,

$$\Phi = \Phi_A \otimes \Phi_B \, : \, \mathcal{H}_A \otimes \mathcal{H}_B \to \mathcal{H}_A \otimes \mathcal{H}_B \otimes \mathcal{H}_{A'} \otimes \mathcal{H}_{B'}. \tag{5.3}$$

By defining local isometric embeddings we take care of the unitary invariance of the trace, i. e.,

$$P(a, b|x, y) = \mathrm{Tr}(|\psi\rangle\langle\psi| \, M_A \otimes M_B) = \mathrm{Tr}\Big(|\psi\rangle\langle\psi| \, (UM_AU^\dagger) \otimes (VM_BV^\dagger)\Big),$$

as well as supplementary degrees of freedom so that if the state was $|\psi\rangle \otimes |\zeta\rangle$ with the measurement operators $\{M_A \otimes \mathbb{1}, M_B \otimes \mathbb{1}\}$ then in spite of the added

local Hilbert spaces we obtain the same correlation. Usually such embeddings are realized using an auxiliary system $|00\rangle_{A'B'}$ so that the local isometry can be written as the following unitary operations on the extended local Hilbert spaces,

$$\Phi_A \otimes \Phi_B \left[|\psi\rangle\langle\psi|\right] = U_{AA'} \otimes V_{BB'} \left(|\psi\rangle \otimes |00\rangle_{A'B'} \langle\psi| \otimes \langle 00|_{A'B'}\right) U^\dagger_{AA'} \otimes V^\dagger_{BB'}. \quad (5.4)$$

Then the self-testing of the state ψ and the measurements $\{M_A, M_B\}$ is defined on the basis of the existence of such an local isometry as follows [50].

- `Self-testing of a state and measurements`:
  For a state ρ_{AB}, its purification $|\psi\rangle_{ABC}$, measurements $\{M_A, M_B\}$ and a given correlation $P(a,b|x,y)$, such that $P(a,b|x,y) = \mathrm{Tr}(M_A \otimes M_B \rho_{AB})$ and $\mathrm{Tr}_C(|\psi\rangle_{ABC}) = \rho_{AB}$, we say that the correlation $P(a,b|x,y)$ self-tests the state $|\psi'\rangle_{A'B'}$ and the measurements $\{M'_A, M'_B\}$ if there exists a local isometry, $\Phi_A \otimes \Phi_B$:

$$\Phi_A \otimes \Phi_B \otimes \mathbb{1}_C \left[M_A \otimes M_B \otimes \mathbb{1}_C |\psi\rangle_{ABC} \otimes |00\rangle_{A'B'}\right] = |\zeta\rangle_{ABC} \otimes \left(M_{A'} \otimes M_{B'} |\psi'\rangle_{A'B'}\right) \quad (5.5)$$

Here the state to be self-tested is *extracted* from ρ_{AB} into the auxiliary system via the unitary operations, and the remains are left in $|\zeta\rangle$ which is often called the *junk state*. We will take the two qubit maximally entangled state, $|\Psi\rangle$, as an example to demonstrate this procedure.

$$|\Psi\rangle = \frac{1}{\sqrt{2}} \left(|00\rangle + |11\rangle\right) \quad (5.6)$$

We'll assume a simple scenario, with two inputs $(x,y) = \{0,1\}$ and two outputs $(a,b) = \{+1,-1\}$ for both Alice and Bob. Such measurement operators that have only two outcomes are called *dichotomic*. Let $M_{x|a}$ (and similarly $N_{y|b}$) denote the projective operator with the measurement setting x (y) and it's outcome a (b), then the observables in the above scenario for Alice and Bob can be written as,

$$A_x = M_{+|x} - M_{-|x} \quad \text{and} \quad B_y = N_{+|y} - N_{-|y} \quad (5.7)$$

It also follows from the construction that they are Hermitian and Unitary, thus $A_x^\dagger = A_x$, $A_x^2 = \mathbb{1}$, $B_y^\dagger = B_y$, and $B_y^2 = \mathbb{1}$.

In the 2 qubit scenario the Bell inequality we use for self testing is the CHSH inequality, which in terms of expectations of observables is the following,

$$\mathcal{B}_{\text{CHSH}} = \langle A_0 B_0 \rangle + \langle A_1 B_0 \rangle + \langle A_0 B_1 \rangle - \langle A_1 B_1 \rangle \leq 2. \tag{5.8}$$

The inequality above is satisfied when Alice and Bob share local correlations whereas non-local correlations violate it. In fact, if Alice and Bob share the 2 qubit maximally entangled state and use anti-commuting observables, $A_0 = \sigma_x$, $A_1 = \sigma_z$, $B_0 = (\sigma_x + \sigma_z)/\sqrt{2}$ and $B_1 = (\sigma_x - \sigma_z)/\sqrt{2}$, they obtain the maximal violation of CHSH inequality for quantum correlations, $2\sqrt{2}$. Note that expectations and correlations are related in the following manner,

$$\langle A_x B_y \rangle = \sum_{a,b} a\, b\, P(a,b|x,y). \tag{5.9}$$

The self-testing statement in this scenario is that the maximal violation of the CHSH inequality can only be achieved by measurements on the 2 qubit maximally entangled state. The main step to prove this self-testing statement for given correlations that violate the CHSH inequality maximally is to show that the local observables of Alice and Bob anti-commute. To show this we consider the CHSH operator shifted by its maximal violation and write it as a *Sum of Squares* (SOS) [41],

Sum of Squares decomposition is an essential part of most Self-Testing proofs.

$$2\sqrt{2}\mathbb{1} - \mathcal{B}_{\text{CHSH}} = \frac{1}{\sqrt{2}}\left[\left(\frac{A_0 + A_1}{\sqrt{2}} - B_0\right)^2 + \left(\frac{A_0 - A_1}{\sqrt{2}} - B_1\right)^2\right], \tag{5.10}$$

and if the purification $|\psi\rangle_{ABP}$ violates the Bell Inequality maximally then

$$\text{Tr}\Big((2\sqrt{2}\mathbb{1} - \mathcal{B}_{\text{CHSH}})\,|\psi\rangle\langle\psi|\Big) = 0,$$

which leads us to the relations,

$$\frac{A_0 + A_1}{\sqrt{2}}\,|\psi\rangle = B_0\,|\psi\rangle \quad \text{and} \quad \frac{A_0 - A_1}{\sqrt{2}}\,|\psi\rangle = B_1\,|\psi\rangle\,. \tag{5.11}$$

Here the $\mathbb{1}$ acting on the second subsystem is implicit. Then using the above relation together with the properties that $A_x^2 = B_y^2 = \mathbb{1}$ it is very easy to show

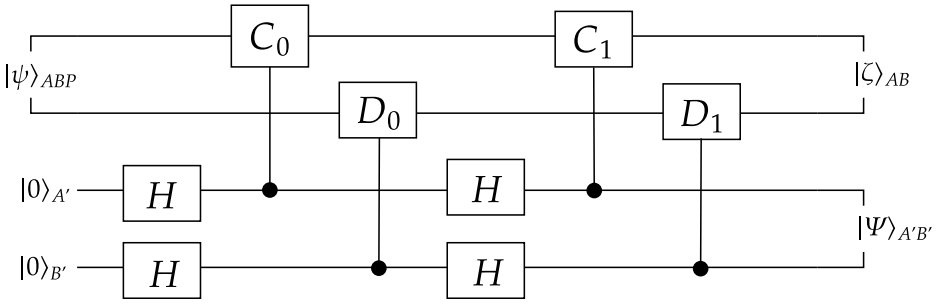


Figure 5.1: Circuit diagram for applying the local isometry equivalent to the swap operation that extracts the state $|\Psi\rangle$ to the auxiliary system, $|00\rangle_{A'B'}$, which is brought to an equal superposition by the Hadamard gate H, after which its value controls the unitary operations on the reference state $|\psi\rangle$.

that the local observables of both Alice and Bob anti-commute, $\{A_0, A_1\} = \{B_0, B_1\} = 0$.

The next step in the proof is to work out the local isometry that will extract the maximally entangled state into the auxiliary space. It turns out that in most cases, the found local isometry is equivalent to the swap operation. The method calls for new Unitary operators C_0, C_1 for Alice and D_0, D_1 for Bob that act on $|\psi\rangle$ and are expected to do the extraction to the auxiliary space. We also want them to act on $|\psi\rangle$ in the same way as A_x and B_y, therefore we define them as,

$$C_0 = \frac{1}{\sqrt{2}}(A_0 + A_1), \qquad C_1 = \frac{1}{\sqrt{2}}(A_0 - A_1) \tag{5.12}$$

for Alice and for Bob we have,

$$D_0 = B_0, \qquad D_1 = B_1. \tag{5.13}$$

Straightaway, from Equation 5.11 we have $C_0|\psi\rangle = D_0|\psi\rangle$ and $C_1|\psi\rangle = D_1|\psi\rangle$. Using this and the properties of A_x and B_y it is possible to verify that $\{C_0, C_1\} = \{D_0, D_1\} = 0$. All that is left now is to apply the local isometry on $|\psi\rangle$ with C_x and D_x as the *controlled* unitary operators in Equation 5.4, i. e., the auxiliary systems of Alice and Bob control the unitary operation that is applied on the reference state $|\psi\rangle$. This procedure is illustrated by the quantum circuit diagram in Figure 5.1.

After the application of the local isometry and a little simplification we obtain,

$$\begin{aligned}\Phi[|00\rangle_{A'B'}|\psi\rangle_{ABP}] =& \frac{1}{4}(|00\rangle_{A'B'} \otimes (\mathbb{1}+C_0)(\mathbb{1}+D_0)|\psi\rangle_{ABP} \\ &+ |01\rangle_{A'B'} \otimes (\mathbb{1}+C_0)D_1(\mathbb{1}+D_0)|\psi\rangle_{ABP} \\ &+ |10\rangle_{A'B'} \otimes C_1(\mathbb{1}+C_0)(\mathbb{1}+D_0)|\psi\rangle_{ABP} \\ &+ |11\rangle_{A'B'} \otimes C_1(\mathbb{1}+C_0)D_1(\mathbb{1}+D_0)|\psi\rangle_{ABP}). \end{aligned} \tag{5.14}$$

A little manipulation using properties of the observables C_x and D_y, leave us with the following,

$$\Phi[|00\rangle_{A'B'}|\psi\rangle_{ABP}] = |\Psi\rangle_{A'B'} \otimes \left(\underbrace{\sqrt{2}(\mathbb{1}+C_0)(\mathbb{1}+D_0)|\psi\rangle_{ABP}}_{|\zeta\rangle} \right), \tag{5.15}$$

where the state $|\Psi\rangle$ has been extracted to the auxiliary system, and thus we have successfully self-tested the state.

The same methodology can be applied to self-testing the measurements in this scenario. Consider the state $D_0|\psi\rangle$, applying the local isometry discussed above and simplifying, we obtain

$$\Phi(|00\rangle_{A'B'} \otimes D_0|\psi\rangle_{ABP}) = (\mathbb{1}\otimes\sigma_z)|\Psi\rangle_{A'B'}|\zeta\rangle_{ABP} \tag{5.16}$$

and again we extracted the state $|\Psi\rangle$ with the operator σ_z acting on the support of $|\psi\rangle$. Similarly all the observables used in the maximal violation of the CHSH inequality can be self-tested.

Then the above two self-testing statements together provide a complete Self-testing of the state and measurements for the maximal violation of the CHSH inequality.

Similar methods have been applied to a variety of scenarios, where the maximal violation of a Bell inequality requires particular entangled states and measurement settings. Indeed, all bipartite pure entangled states can be self-tested. But the outlook becomes considerably more difficult in the multipartite scenario, as we shall discuss in the next section.

5.3 SELF-TESTING IN MULTIPARTITE SCENARIOS

Self-testing becomes more demanding in the multipartite scenario as there is now a requirement for space-like separation between multiple measurement devices. Similar to the case of entanglement detection where increasing number of subsystems raised the complexity of the problem significantly, there are similar challenges in Self-testing. A significant factor here is the non-applicability of the Schmidt Decomposition in the multipartite scenario. There are a few major methods that are widely used in the literature for self-testing in the multipartite scenario. For example, self-testing using stabilizer formalisms has been employed for graph states [35], Genuinely Entangled subspaces [32] and partially entangled GHZ states [3]. Another way is to tailor the Bell inequalities to self-test particular states, for example for GHZ states, and in general this can be done using linear programming [4, 47, 60], however, the complexity cost to check self-testing statements is severe . Another way to go is to reduce the multipartite scenario to a set of bipartite scenarios [49] where the self-testing is easier and similar to what we saw in the previous section. The self-testing statements for the classes of Dicke states [21], including W states [59] were obtained analytically in this manner.

In our work [40], we prove Mayer-Yao like self-testing statements for multipartite Bell inequalities without relying on the widely used sum of squares decomposition of the Bell operator. It applies in the general N-party case where the spatially separated parties use only dichotomic measurement operators.

In what follows we will prove self testing statements for not only Bell inequalities that are linear functionals of the observed correlations, but also for quadratic Bell inequalities, that provide a stronger certification of genuine multipartite nonlocality. The importance of our proofs of self-testing in these scenarios stems from the fact that we make absolutely no assumptions about the dimensions and the number of subsystems in the shared state.

We shall first briefly discuss the multipartite Bell inequalities for which we will give self-testing statements. Then we will prove that observables of each party in a general Bell scenario with two measurement settings and two outcomes, can always be simultaneously transformed to anti-diagonal form, and then we'll use them to derive self-testing statements for the maximal violation of the considered Bell Inequalities. In the description of the Bell inequalities in

the following sections, we will have the labels of the parties to be $j = 1, 2, \ldots, N$ and each of them have two dichotomic observables $A^{(j)}$ and $A'^{(j)}$.

5.3.1 *Linear Bell inequalities*

Most widely used Bell inequalities are linear functions of the correlations, where the N-party Bell operator is of the form $\bigotimes_{i=1}^{N} O^{(i)}$ where $O^{(i)} \in \{A^{(j)}, A'^{(j)}\}$. The first family of Bell inequalities we consider is the *Werner-Wolf-Weinfurter-Żukowski-Brukner* (WWWŻB) inequalities [56, 57, 61]. The Bell operator has the following form,

$$\mathcal{W}_N = \frac{1}{2^N} \sum_{s_1,\ldots,s_N=\pm 1} S(s_1, \ldots, s_N) \bigotimes_{j=1}^{N} (A^{(j)} + s_j A'^{(j)}), \tag{5.17}$$

where $S(s_1, \ldots, s_N) = \pm 1$. For all local correlations the inequality $\langle \mathcal{W}_N \rangle \leq 1$ is satisfied.

Another family of Bell inequalities important for this work is the *Mermin-Ardehali-Belinskii-Klyshko* (MABK) inequalities where the bell operator has the following recursive form,

$$\begin{aligned} \mathcal{M}_N &= \frac{1}{2} \left(\mathcal{M}_{N-1} \otimes \left(A^{(j)} + A'^{(j)} \right) + \mathcal{M}'_{N-1} \otimes \left(A^{(j)} - A'^{(j)} \right) \right) \\ &= \frac{1}{2^N} \left((1-\imath)^{N-1} \bigotimes_{j=1}^{N} \left(A^{(j)} + \imath A'^{(j)} \right) + (1+\imath)^{N-1} \bigotimes_{j=1}^{N} \left(A^{(j)} - \imath A'^{(j)} \right) \right), \end{aligned} \tag{5.18}$$

where $\mathcal{M}'_N$ is defined as $\mathcal{M}_N$ but with $A^{(j)}$ and $A'^{(j)}$ interchanged. The inequality is defined as $\langle \mathcal{M}_N \rangle \leq 1$ for all local correlations, $\langle \mathcal{M}_N \rangle \leq \sqrt{2}^{N-2}$ for all biseparable quantum correlations and $\langle \mathcal{M}_N \rangle \leq \sqrt{2}^{N-1}$ for all quantum correlations. The maximum value is attained for the MABK inequalities when the parties share an N-qubit GHZ state and use anti-commuting observables.

A special case of the WWWŻB inequalities is the Svetlichny inequality, whose operator can be defined using the MABK operators as follows,

$$\mathcal{S}_N^{\pm} = \begin{cases} 2^{k-1}(-1)^{k(k\pm1)/2}\mathcal{M}_N^{\pm} & N = 2k \\ 2^{k\pm1}\left((-1)^{k(k\pm1)/2}\mathcal{M}_N \mp (-1)^{k(k\mp1)/2}\mathcal{M}'_N\right) & N = 2k+1 \end{cases} \tag{5.19}$$

Here $\mathcal{M}_N$ is relabeled as $\mathcal{M}_N^+$ and $\mathcal{M}'_N$ as $\mathcal{M}_N^-$. Here the Bell inequality reads $\langle \mathcal{S}_N^{\pm} \rangle \leq 2^{N-1}$ for biseparable quantum correlations and $\langle \mathcal{S}_N^{\pm} \rangle \leq 2^{N-1/2}$ for genuine multipartite correlations.

5.3.2 *Quadratic Bell inequalities*

In the bipartite scenario, the correlations $P(a, b|x, y)$ form a convex set, wherein the set of local correlations is a convex polytope. The Bell inequalities that distinguish between local and non-local correlations can be seen as the separating hyperplanes in this setting, and it is enough to consider linear Bell inequalities. On the other hand, in the multipartite scenario, while the local set still forms a convex polytope, the set of biseparable correlations is convex but not a polytope. Therefore, the consideration that non-linear functionals of the correlations would give a tighter witness for genuine multipartite correlations holds true. An example of such a Bell inequality is Uffink's quadratic Bell inequality. Uffinks' Bell inequality has two subfamilies, one based on the MABK inequalities and the other on the Svetlichny inequalitis, denoted by $\mathcal{U}_N^{\mathcal{M}}$ and $\mathcal{U}_N^{\mathcal{S}}$ respectively. They are both defined as in the following equations,

$$\mathcal{U}_N^{\mathcal{M}} = \langle \mathcal{M}_N \rangle^2 + \langle \mathcal{M}'_N \rangle^2 \tag{5.20}$$

$$\mathcal{U}_N^{\mathcal{S}} = \langle \mathcal{S}_N \rangle^2 + \langle \mathcal{S}'_N \rangle^2 \tag{5.21}$$

The set of biseparable correlations satisfy $\langle \mathcal{U}_N^{\mathcal{M}} \rangle \leq 2^{N-2}$ and $\langle \mathcal{U}_N^{\mathcal{S}} \rangle \leq 2^{2N-2}$, and the maximal violation exhibited by some non-local correlation is $\langle \mathcal{U}_N^{\mathcal{M}} \rangle \leq 2^{N-1}$ and $\langle \mathcal{U}_N^{\mathcal{S}} \rangle \leq 2^{2N-1}$.

5.3.3 *Anti-commutation of Local Observables*

Earlier in the example of self-testing in the bipartite case, we saw how the decomposition of the Bell operator into the sum of squares let us prove that the

local observables of both Alice and Bob had to be anti-commuting if the CHSH inequality was violated maximally. We will similarly characterize dichotomic observables A and A' acting on an arbitrary Hilbert space for any given party, with minimal assumptions and without relying on the sum of squares decomposition.

The first assumption we take using Naimark's dilation theorem is that the local observables A and A' are projective, i. e., $A^2 = A'^2 = \mathbb{1}$. A consequence of this is that the state that gives the maximum expectation value of a Bell operator constructed from projective observables is the eigenvector of the Bell operator corresponding to the maximum eigenvalue, therefore the shared state can always be taken as pure.

Lemma 5.3.1. *Given any two dichotomic projective observables A and A', A' can be decomposed as the sum of two observables $-\mathbb{1} \leq A'_- \leq \mathbb{1}$ and $-\mathbb{1} \leq A'_+ \leq \mathbb{1}$, such that $[A, A'_+] = 0$, $\{A, A'_-\} = 0$, $\{A'_+, A'_-\} = 0$ and $(A'_+)^2 + (A'_-)^2 = \mathbb{1}$.*

Proof. The observable A being projective, it also has to be Hermitian and Unitary, and thus can always be diagonalized with eigenvalues ± 1 grouped together according to sign,

$$A = \begin{pmatrix} \mathbb{1}_m & 0 \\ 0 & -\mathbb{1}_n \end{pmatrix}. \tag{5.22}$$

The observable A' can have the following generic block representation,

$$A' = \begin{pmatrix} D_1 & D_2 \\ D_2^\dagger & D_3 \end{pmatrix}. \tag{5.23}$$

Because A' is also Projective, Hermitian and Unitary, in the above representation we have $D_1^\dagger = D_1$ and $D_3^\dagger = D_3$. Then we can obviously write A' as the sum of the following two observables,

$$A'_- = \begin{pmatrix} 0 & D_2 \\ D_2^\dagger & 0 \end{pmatrix}, \qquad A'_+ = \begin{pmatrix} D_1 & 0 \\ 0 & D_3 \end{pmatrix}. \tag{5.24}$$

Now, A'_+ commutes with A, and A'_- anti-commutes with A and A'_+, therefore $[A, A'_+] = 0$, $\{A, A'_-\} = 0$, $\{A'_+, A'_-\} = 0$. For the last part of the lemma, consider A'^2,

$$A'^2 = (A'_+)^2 + (A'_-)^2 + \{A'_+, A'_-\} \\ = \begin{pmatrix} D_1^2 + D_2 D_2^\dagger & D_1 D_2 + D_2 D_3 \\ D_2^\dagger D_1 + D_3 D_2^\dagger & D_2^\dagger D_2 + D_3^2 \end{pmatrix} = \mathbb{1}. \quad (5.25)$$

Therefore the off-diagonal blocks of A'^2 are zero, and because they correspond to the anticommutator, $\{A'_+, A'_-\} = 0$, which leaves us with the sum of squares $(A'_+)^2 + (A'_-)^2 = \mathbb{1}$. □

The above lemma gives as a byproduct the relation between the corresponding eigenvalues of A', A'_+ and A'_-,

$$\lambda^i_{A'} = \sqrt{(\lambda^i_{A'_+})^2 + (\lambda^i_{A'_-})^2} = \pm 1, \quad (5.26)$$

where λ^i denotes the i^{th} eigenvalue. Without loss of generality we can take the eigenvalues of the observables A'_+ and A'_- to be $\pm \sin\theta_i$ and $\pm \cos\theta_i$. The next lemma shows the consequence of anti-commutation on the eigenspaces of the observables.

Lemma 5.3.2. *Given any two anti-commuting observables A and A'_-, their spectra are symmetric (i.e., if λ is an eigenvalue of one of the operators, then so is $-\lambda$), moreover, A'_- (A) is a linear map between positive and negative eigenspaces of A (A'_-) or nullifies its eigenvectors.*

Proof. If $|\psi\rangle$ is an eigenvector of A with the eigenvalue λ , then $A|\psi\rangle = \lambda|\psi\rangle$, and

$$A(A'_-)|\psi\rangle = -A'_- A|\psi\rangle = -\lambda A'_-|\psi\rangle. \quad (5.27)$$

Therefore, $A'_-|\psi\rangle$ is either a null vector or an eigenvector of A with the sign of the eigenvalue changed. Similarly for A'_-. □

We have further observations from the above lemma as $A^2 = \mathbb{1}$.

Lemma 5.3.3. *Given any two anti-commuting observables A and A'_-, such that $A^2 = \mathbb{1}$, then A provides a bijective mapping between the eigenspaces $E_\lambda^{A'_-}$ and $E_{-\lambda}^{A'_-}$*

corresponding to the eigenvalues $\lambda^{A'_-}$ and $-\lambda^{A'_-}$, respectively. Moreover, the direct sum of the pair of eigenspaces $E^{A'_-}_{\pm\lambda_i} = E^{A'_-}_{\lambda_i} \oplus E^{A'_-}_{-\lambda_i}$ is even dimensional, and so is the subspace $\bigoplus_i E^{A'_-}_{\pm\lambda_i}$.

Proof. Using the eigenvalue equation from the previous lemma,

$$A(A'_-)\,|\psi\rangle = -A'_- A\,|\psi\rangle = -\lambda^{A'_-} A\,|\psi\rangle\,, \tag{5.28}$$

and operating from the left with A,

$$-AA'_- A\,|\psi\rangle = \lambda^{A'_-} A^2\,|\psi\rangle = \lambda^{A'_-}\,|\psi\rangle\,. \tag{5.29}$$

Here the second inequality uses $A^2 = \mathbb{1}$, and notice that we have recovered the original eigenvector. Therefore, as A is it's own inverse, the mapping between the eigenspaces is one to one and invertible, i. e., bijective. Furthermore, A maps the direct sum of these subspaces $E^{A'_-}_{\lambda_i} \oplus E^{A'_-}_{-\lambda_i}$ onto itself. This pairing of eigenvectors in the subspace allows us to conclude that the subspace is even dimensional.

The diagonalizability of A'_- implies that the Hilbert space can be written as a direct sum of such pairs of eigenspaces and the kernel,

$$\mathcal{H} = E^{A'_-}_{\lambda_1} \oplus E^{A'_-}_{-\lambda_1} \cdots \oplus E^{A'_-}_{\lambda_r} \oplus E^{A'_-}_{-\lambda_r} \oplus \ker(A'_-). \tag{5.30}$$

Also notice that the anti-commutation relation between A and A'_- can be written as $AA'_-A^{-1} = -A'_-$, and $\det(A'_-) = (-1)^d \det(A'_-)$, where $d = n + m$ is the dimension of the Hilbert space on which these observables act. This is only possible if either the determinant $\det(A'_-) = 0$ or the observable A'_- is even-dimensional. In the case when the determinant is zero the kernel is non-trivial but still cannot have any possible contribution to the violation of any Bell inequality. Therefore, it can be effectively dropped from the system, which leave us an even dimensional subspace. Lastly, as $A' = A'_+ + A'_-$, the three matrices can be simultaneously restricted to this even dimensional subspace. □

As we discussed before, the observables $A'_\pm$ have the eigenvalues $\pm\sin\theta_i$ and $\pm\cos\theta_i$ and we denote the corresponding eigenspaces as $E^{A'_+}_{\pm\theta_i}$ and $E^{A'_-}_{\pm\theta_i}$. With the previous lemmas in mind we prove the following lemma that allows us to further characterize the local observables.

Corollary 5.3.3.1. *Given two even-dimensional anti-commuting operators A_+ and A_- such that $(A'_+)^2 + (A'_-)^2 = \mathbb{1}$, they can be written in the form,*

$$A'_{+|\theta_i} = \sin\theta_i B_{+|\theta_i} \qquad A'_{-|\theta_i} = \cos\theta_i B_{-|\theta_i} \tag{5.31}$$

when restricted to the subspace $E^{A'_-}_{+\theta_i} \oplus E^{A'_-}_{-\theta_i}$ with corresponding eigenvalues $\pm\sin(\theta_i)$ for A_+ and $\pm\cos(\theta_i)$ for A_-, such that the operators $B_{\pm|\theta_i}$ are traceless and projective and anti-commuting.

Proof. We already know that the subspace $E^{A'_-}_{+\theta_i} \oplus E^{A'_-}_{-\theta_i}$ is even dimensional. Restricted to this subspace, the operators $A'_{+|\theta_i}$ and $A'_{-|\theta_i}$ only have eigenvalues $\pm\sin\theta_i$ and $\pm\cos\theta_i$, while still satisfying $(A'_{+|\theta_i})^2 + (A'_{-|\theta_i})^2 = \mathbb{1}$. The two restricted operators can be scaled by the inverse of their eigenvalues to obtain another two observables which will then satisfy,

$$\cos^2\theta_i B^2_{+|\theta_i} + \sin^2\theta_i B^2_{-|\theta_i} = \mathbb{1}. \tag{5.32}$$

As a result of the scaling, the operators only have eigenvalues ± 1 that occur in pairs, as the subspace is even-dimensional, so that $\mathrm{Tr}\big(B_{\pm|\theta_i}\big) = 0$, and $B^2_{\pm|\theta_i} = \mathbb{1}$. That $B_{\pm|\theta_i}$ anti-commute with each other follows directly from the anti-commutation of $A'_{\pm|\theta_i}$. □

Then in such an even dimensional subspace we have the following lemma.

Lemma 5.3.4. *Given any two traceless and projective anti-commuting observables B_+ and B_-, the observable $\cos\alpha B_+ + \sin\alpha B_-$ is also traceless and projective.*

Proof. We start by expanding the square,

$$\begin{aligned}&(\cos\alpha B_+ + \sin\alpha B_-)^2\\ &\quad= \cos^2\alpha B_+^2 + \sin^2\alpha B_-^2 + \cos\alpha\sin\alpha\{B_+, B_-\}\\ &\quad= \mathbb{1}.\end{aligned} \tag{5.33}$$

And then the trace is, $\cos\alpha\,\mathrm{Tr}(B_+) + \sin\alpha\,\mathrm{Tr}(B_-) = 0$. □

We started with two dichotomic observables that were Hermitian, Unitary and Projective, and using their properties that arise organically without further assumptions, we characterised their behavior. The results above indicate that when a Bell inequality is maximally violated, the dimension of the local Hilbert

spaces must be even dimensional. This brings us to the most important theorem in our work, that proves that all the local observables can be represented in anti-diagonal form.

Theorem 5.3.5. *Given any three dichotomic traceless and projective observables A, B_+, and B_-, such that $[A, B_+] = 0$, $\{A, B_-\} = 0$, and $\{B_+, B_-\} = 0$, then these operators have a simultaneous anti-diagonal matrix representation.*

Proof. As $[A, B_+] = 0$ they are diagonal in the same basis and because $\{A, B_-\} = 0$, the operator B_- will have an anti-diagonal form in the same basis. We take the dimension of the subspace for which the eigenvalues of A and B_+ are equal to be $2d_1$, and $2d_2$ for the subspace where the eigenvalues differ. Using the previous Lemmas, the operators A, B_+, and B_- will have the following form,

$$A = \begin{pmatrix} \mathbb{1}_{d_1} & \cdot & \cdot & \cdot \\ \cdot & \mathbb{1}_{d_2} & \cdot & \cdot \\ \cdot & \cdot & -\mathbb{1}_{d_2} & \cdot \\ \cdot & \cdot & \cdot & -\mathbb{1}_{d_1} \end{pmatrix},$$

$$B_+ = \begin{pmatrix} \mathbb{1}_{d_1} & \cdot & \cdot & \cdot \\ \cdot & -\mathbb{1}_{d_2} & \cdot & \cdot \\ \cdot & \cdot & \mathbb{1}_{d_2} & \cdot \\ \cdot & \cdot & \cdot & -\mathbb{1}_{d_1} \end{pmatrix},$$

$$B_- = \begin{pmatrix} \cdot & \cdot & \cdot & U_1 \\ \cdot & \cdot & U_2 & \cdot \\ \cdot & U_2^\dagger & \cdot & \cdot \\ U_1\dagger & \cdot & \cdot & \cdot \end{pmatrix}. \tag{5.34}$$

Since $B_-^2 = \mathbb{1}$, U_1 and U_2 must be unitary. Thus, without altering A or B_+ we can take four unitaries V_1, V_2, V_3, V_4, each acting on a different block, such that $V_1 U_1 V_4^\dagger = J_{d_1}$ and $V_2 U_2 V_3^\dagger = J_{d_2}$, where J_d is the row-reversed $d \times d$ identity matrix.

Consequently, we can now restrict ourselves to considering any one of the $d_1 + d_2$ two-dimensional subspaces, on which A and B_+ are represented by $\pm\sigma_z$,

while B_- is represented by σ_x. Upon applying in each of these two-dimensional subspaces, a rotation by $\frac{2\pi}{3}$ with respect to axis $(1,1,1)$ yields,

$$U = \frac{1}{2}\begin{pmatrix} -1+\iota & 1+\iota \\ -1+\iota & -1-\iota \end{pmatrix}, \tag{5.35}$$

which essentially is the transformation $\sigma_z \to \sigma_x \to \sigma_y$. Hence bringing all three operators to strictly anti-diagonal form. □

The significance of the above results lies in the fact that we are now able to take, for each party j, the observables σ_x and $\cos\theta_j\sigma_x + \sin\theta_j\sigma_y$ as $A^{(j)}$ and $A'^{(j)}$. Then, the Bell operator itself is in a strictly anti-diagonal form, and to achieve maximal violation of the such a Bell inequality one necessarily needs N-qubit GHZ states as they have maximal anti-diagonal elements.

5.3.4 *Self-Testing statements for Linear Bell inequalities*

We shall now use the results obtained in the previous section to obtain self-testing statements for the MABK and WWWŻB inequalities.

5.3.4.1 *Self-Testing N-party MABK inequalities*

Theorem 5.3.6. *To achieve maximal quantum violation of an N-party MABK inequality, $\langle \mathcal{M}_N \rangle = \sqrt{2}^{N-1}$, the parties must share an N qubit GHZ state $|GHZ_N\rangle = \frac{1}{\sqrt{2}}(|0\rangle^{\otimes N} + e^{\iota\phi_N}|1\rangle^{\otimes N})$ and perform maximally anti-commuting projective measurements $A^{(j)} = \sigma_x$ and $A'^{(j)} = \sigma_y$ (upto local auxiliary systems and local isometries).*

Proof. As we have already established the local observables of any party j can be taken to be,

$$\begin{aligned} A^{(j)} &= \sigma_x, \\ A'^{(j)} &= \cos\theta_j\sigma_x + \sin\theta_j\sigma_y, \end{aligned} \tag{5.36}$$

effectively acting on the two dimensional subspaces. Substituting these in the MABK operator in Equation 5.18, we get an antidiagonal operator with the values on the antidiagonal from top right to bottom left,

$$\mathcal{M}_N = \text{adiag}\begin{pmatrix} \frac{1}{2}\left(\left(\frac{1-\iota}{2}\right)^{N-1}\prod_{j=1}^{N}(1+\iota e^{-\iota\theta_j}) + \left(\frac{1+\iota}{2}\right)^{N-1}\prod_{j=1}^{N}(1-\iota e^{-\iota\theta_j})\right) \\ \vdots \\ \frac{1}{2}\left(\left(\frac{1-\iota}{2}\right)^{N-1}\prod_{j=1}^{N}(1+\iota e^{\iota\theta_j}) + \left(\frac{1+\iota}{2}\right)^{N-1}\prod_{j=1}^{N}(1-\iota e^{\iota\theta_j})\right) \end{pmatrix} \tag{5.37}$$

When for a party j, $\theta_j = \pm\pi/2$ one of the antidiagonal entries becomes maximal and has the value $\sqrt{2}^{N-1}$, while all the others become zero.

If we consider a state $|\psi\rangle = a\,|0\rangle^{\otimes N} + b\,|1\rangle^{\otimes N}$, the value of the inner product,

$$\begin{aligned}\langle\psi_N|\mathcal{M}_N|\psi_N\rangle = 2\,\text{Re}\Bigg[\overline{\alpha}\frac{1}{2}\Bigg(&\left(\frac{1-\iota}{2}\right)^{N-1}\prod_{j=1}^{N}(1+\iota e^{-\iota\theta_j}) \\ &+ \left(\frac{1+\iota}{2}\right)^{N-1}\prod_{j=1}^{N}(1-\iota e^{-\iota\theta_j})\Bigg)\beta\Bigg],\end{aligned} \tag{5.38}$$

is essentially the weighted sum of the entries at top right and bottom left. The expression can be upper bounded [40],

$$\langle\psi_N|\mathcal{M}_N|\psi_N\rangle \leq 2^{\frac{N-1}{2}}\left(\prod_{j=1}^{N}\left|\cos\frac{\pi}{4}+\frac{\theta_j}{2}\right| + \prod_{j=1}^{N}\left|\sin\frac{\pi}{4}+\frac{\theta_j}{2}\right|\right). \tag{5.39}$$

Where we can further drop the positive terms in the product for arbitrarily chosen $N-2$ parties, and get a simplified expression,

$$\langle\psi_N|\mathcal{M}_N|\psi_N\rangle \leq \tag{5.40}$$

$$2^{\frac{N-1}{2}}\left(\left|\sin\left(\frac{2\theta_i+\pi}{4}\right)\sin\left(\frac{2\theta_j+\pi}{4}\right)\right| + \left|\cos\left(\frac{2\theta_i+\pi}{4}\right)\cos\left(\frac{2\theta_j+\pi}{4}\right)\right|\right)$$

$$= 2^{\frac{N-1}{2}}\max\left\{\left|\cos\left(\frac{\theta_i+\theta_j+\pi}{2}\right)\right|, \left|\cos\left(\frac{\theta_i-\theta_j}{2}n\right)\right|\right\} \tag{5.41}$$

$$\leq 2^{\frac{N-1}{2}}, \tag{5.42}$$

Therefore, the inequality is only satisfied when for all the parties $\theta_j = \pm\pi/2$, so that the local observables of each party anti-commute maximally, and the state in the space where the measurements act non-trivially is the GHZ state,

upto local isometries and auxiliary degrees of freedom, which concludes the self-testing statement. □

5.3.4.2 *Self-Testing of WWWŻB inequalities*

Using the same methods as above the Svetlichny inequality can be self-tested as the bell operator again comprises of MABK operators. In the case of tripartite WWWŻB inequalities there are a few different equivalence classes of the inequalities.

The first class is equivalent to the Mermin's inequality (up to relabeling),

$$\langle \mathcal{M}_3 \rangle = \frac{1}{2}\Big(\Big\langle (A^{(1)}A^{(2)}A'^{(3)}\Big\rangle + \Big\langle A^{(1)}A'^{(2)}A^{(3)}\Big\rangle + \Big\langle A'^{(1)}A^{(2)}A^{(3)}\Big\rangle - \Big\langle A'^{(1)}A'^{(2)}A'^{(3)})\Big\rangle\Big) \leq 1. \quad (5.43)$$

The proof of Theorem 5.3.6 directly apply, and we retrieve the three qubit GHZ state as well as maximally anti-commuting observables for the maximal violation of this inequality. The second equivalence class is a group of *tilted* Bell inequalities.

$$\Big(3\Big\langle A^{(1)}A^{(2)}A^{(3)}\Big\rangle + \Big\langle A^{(1)}A^{(2)}A'^{(3)}\Big\rangle + \Big\langle A^{(1)}A'^{(2)}A^{(3)}\Big\rangle + \Big\langle A'^{(1)}A^{(2)}A^{(3)}\Big\rangle - \Big\langle A^{(1)}A'^{(2)}A'^{(3)}\Big\rangle - \Big\langle A'^{(1)}A^{(2)}A'^{(3)}\Big\rangle - \Big\langle A'^{(1)}A'^{(2)}A^{(3)}\Big\rangle + \Big\langle A'^{(1)}A'^{(2)}A'^{(3)}\Big\rangle\Big) \leq 4. \quad (5.44)$$

In this case the operator is strictly anti-diagonal matrix, where unlike the MABK case, the maximal value is reached for $\cos\theta_j = -1/3$. Here the GHZ state is recovered but the observables are not required to be maximally anti-commuting.

The other two equivalence classes are CHSH like inequalitites,

$$\frac{1}{2}\Big(\Big\langle A^{(1)}A^{(2)}A^{(3)}\Big\rangle + \Big\langle A'^{(1)}A^{(2)}A^{(3)}\Big\rangle + \Big\langle A^{(1)}A'^{(2)}A'^{(3)}\Big\rangle - \Big\langle A'^{(1)}A'^{(2)}A'^{(3)}\Big\rangle\Big) \leq 1 \quad (5.45)$$

and the other one is

$$\frac{1}{2}\left(\left\langle A^{(1)}A^{(2)}A^{(3)}\right\rangle + \left\langle A'^{(1)}A^{(2)}A^{(3)}\right\rangle + \left\langle A^{(1)}A'^{(2)}A^{(3)}\right\rangle - \left\langle A'^{(1)}A'^{(2)}A^{(3)}\right\rangle\right) \leq 1. \quad (5.46)$$

In both the cases the operator is anti-diagonal and the maximal violation of these inequalities is recovered for the GHZ state and for all parties $\theta_j = \pm\pi/2$ so that they have maximally anti-commuting observables.

The above results demonstrate that this technique of self-testing proofs that relies on the antidiagonal form of the Bell operator, can be generally applied to any Bell inequality that is linear in the observed correlations and thus its operator is written in terms of local observables of the form Equation 5.36.

5.3.5 *Self-Testing Uffink's quadratic inequalities*

The Uffink's inequality in Equation 5.20 and Equation 5.21 is quadratic in correlations and thus forms a better witness of the genuine multipartite non-locality for $N \geq 3$, as compared to linear Bell inequalities. To provide a self-testing statement for Uffink's inequality, we use the simple property of two real numbers, $a, b \in \mathbb{R}$ that $a^2 + b^2 = |a + \imath b|^2$, to linearize the Uffink's Bell inequality for $N \geq 3$, which transforms the problem to self-testing of anti-diagonal representation of a non-Hermitian matrix.

Theorem 5.3.7. *To achieve the maximal violation of an N-party Uffink's inequality given by, $\langle \mathcal{U}_N \rangle = 2^{N+1}$, the parties must share an N-qubit GHZ state and perform maximally anti-commuting projective measurements, $A^{(j)} = \sigma_x$ and $A'^{(j)} = \sigma_y$, up to local isometries.*

Proof. The Uffink's inequality is either defined with the squares of the mean values of MABK operators or with operators from the Svetlichny Bell Inequality. Svetlichny Bell operators, again can be defined in terms of MABK operators, therefore we consider two MABK operators $\mathcal{U}_N$ and $\mathcal{U}'_N$ which have the following relation,

$$\mathcal{U}_N^{\mathcal{M}} = \langle \mathcal{U}_N \rangle^2 + \langle \mathcal{U}'_N \rangle^2 = \left| \langle \mathcal{U}_N \rangle + \imath \langle \mathcal{U}'_N \rangle \right|^2. \quad (5.47)$$

Then we can define the non-Hermitian operator $\tilde{\mathcal{U}}_N$ as

$$\begin{aligned}\tilde{\mathcal{U}}_N =&(\mathcal{U}_N + \imath\mathcal{U}'_N)\\ =&\left(\frac{1-\imath}{2}\right)^{N-1}\bigotimes_{j=1}^{N}(A^{[j]} + \imath A'^{[j]}).\end{aligned} \tag{5.48}$$

The operator $\tilde{\mathcal{U}}_N$ in terms of the general local observables $A^{(j)} = \sigma_x$ and $A'^{(j)} = \cos\theta_j\sigma_x + \sin\theta_j\sigma_y$, takes the anti-diagonal form,

$$\tilde{\mathcal{U}}_N = \text{adiag}\begin{pmatrix}\left((\frac{1-\imath}{2})^{N-1}\Pi_{j=1}^{N}(1+\imath e^{-\imath\theta_j})\right)\\ \vdots \\ \left((\frac{1-\imath}{2})^{N-1}\Pi_{j=1}^{N}(1-\imath e^{\imath\theta_j})\right)\end{pmatrix}, \tag{5.49}$$

by a similar argument to what we previously used in the linear inequalities, the maximum expectation value of the operator is reached only if all the parties have $\theta_j = \pm\pi/2$ so that their observables are maximally anti-commuting and they share an N-qubit GHZ state.

Important difference here is that the quadratic nature of the inequalities makes it so that any state that is LOCC equivalent to the GHZ state maximally violates the Uffink's Bell inequality. Therefore instead of a single state that is self-tested, we have an LOCC orbit of the GHZ state with maximally anti-commuting observables that is self-tested using maximal violation of the Uffink's quadratic Bell inequality. □

5.4 SUMMARY

In the chapter we discussed the application of correlations that are non-local and thus cannot be modeled by classical Hidden Variable theories. These non-local correlations have found an irrevocably essential role in Quantum cryptographic security. The paradigm of Device independent certification of non-locality shared among participants of a quantum cryptographic protocol provides unconditional security. Device independence uses the impossibility of violation of Bell inequalities via local correlations to certify the state shared between the participants of the protocol as local or non-local without relying on any information about the physical systems involved. The only information

required are the probability statistics of the measurements performed by the participants on their own local systems. The Device Independent paradigm of Self-Testing goes one step further and allows the participants to use the maximal violation of a Bell inequality to not only ascertain the state that they share among them, but also the measurement settings that result in the maximal violation of the said Bell inequality. We also discussed the usual methods of self-testing states and measurements in the bipartite and multipartite scenario.

We then provided a new technique to achieve self-testing statements for linear and more importantly quadratic Bell inequalities. Our technique is valid in the most general case of arbitrary dimensional systems with any number of parties, and does not rely on any assumptions except that all the participating subsystems are spatially separated. We prove that given to d dimensional projective operators, we can always transform them simultaneously to an anti-diagonal form, and thus any linear Bell operator constructed with the observables is also anti-diagonal, which directly leads to the maximally entangled states, N-qubit GHZ states, as the states that provide maximal Bell violation and the self-testing statements follow.

6

CONCLUSIONS

We started in Chapter 1 with describing the long standing problem of Separability, which is the main focus of this work. There are a few different avenues of tackling the separability problem that are all state specific, and rely on its properties and the properties of its subsystems. As classification of a state as separable or entangled is of great importance for a lot of applications that use entanglement as a resource, there has been a lot of research dedicated to finding separability criteria and measures of entanglement. Numerous as they are, these tools are best applied in lower dimensional systems, for example the bipartite scenario with Hilbert space dimensions 4 and 6 is completely characterized by the PPT criterion, which was proven to be a necessary and sufficient condition. Although, in higher local dimensions and number of subsystems, the PPT criterion only provides a necessary condition. This difficulty is shared among several separability criteria.

Then there are Entanglement measures that quantify the amount of entanglement defined for both bipartite and multipartite scenarios, although the calculation is only easy for the set of pure states. Usually the entanglement measures are first defined for pure states and then extended to the set of mixed states by convex roof constructions. Such a construction entails optimization over all pure state decompositions of a given mixed state. Such optimizations are extremely hard to solve efficiently, as the dimension of the optimization problem grows exponentially with the dimension of Hilbert space. This seems to be a common theme in all attempts to solve the separability problem, there are always optimization problems that arise and are exponentially hard to solve. So it follows that the difficulty is intrinsic to the set of quantum states. Then it was proved to be so, by the result that in an arbitrary Hilbert space, the task of classifying a given state to be separable or entangled is in general NP-HARD.

We discussed several attempts that approach the problem algorithmically, but they suffer from the same curse of exponential complexity. Accepting that, the algorithms usually provide a one-way criterion, such that the algorithm is guaranteed to halt on the given input state if it was, let's say, entangled, and

would not stop indefinitely if the state was separable, or vice versa. Although two of them could be combined to provide a criteria that detects both entanglement or separability. Despite the exponential complexity of the algorithms, for lower dimensions the solution is usually found relatively easily. In higher dimensions and multiple subsystems, the trouble is not just complexity but also the different types of separability that arise with an increase in the number of subsystems. The subsystems can be entangled in all kinds of combinations but to understand them they can be arranged in an hierarchy with levels that start from full separability, where all the subsystems are separate, and ends on genuine multipartite entanglement, where no subset of the systems is separable. This concept is called k-separability.

Another approach to attempt a classification of states as separable and entangled is to use the convex geometry of the sets of separable states which is a subset of the convex set of all quantum states. Particularly, the geometry imposed by the Hilbert-Schmidt distance norm on the set of quantum states is of interest. The set of all states under the Hilbert-Schmidt norm forms a hyperball, sometimes called the generalized bloch ball, where the pure states all lie on the hyper-surface of the hyperball, while the mixed states populate the interior. For a single qubit, this geometry takes the shape of a real three dimensional ball, and every point in and on the ball is a valid quantum state. This is not the case in higher dimensional systems. Nevertheless, the set of separable states constitutes a convex set with pure product states as it's extreme points, which lie on the hyper-surface.

Using the geometric approach, the problem of state classification is equivalent to the Separation problem in optimization, where one wants to find the hyperplane separating a point from a convex set, or certify that the point is inside the set. This problem can be reformulated to the problem of finding the minimum distance of the given point from the convex set. The point in the convex set that provides this minimum distance must lie on the boundary, and thus a hyperplane tangent to the convex set at this point naturally separates the given point and the convex set. When applied to the geometry of quantum states, this translates exactly into the problem of finding the Closest Separable State to a given reference state and constructing the tangent hyperplane through the closest separable state such that the reference state and the convex set of separable states are on the opposite sides of the hyperplane. This hyperplane is called an Optimal Entanglement Witness.

We employed the algorithm proposed by Gilbert in 1966 for minimizing quadratic functions on a convex set and modified it for minimal complexity by removing an optimization step from each iteration of the algorithm. Instead we perform such an optimization only once to optimize the Entanglement Witnesses constructing the approximate Closest Separable states, which are the output of the algorithm. This approach with Gilbert's algorithm has been shown to be working successfully for a variety of classes of states. Moreover, due to the geometry, there is always a non-zero distance from the set of separable states for any entangled state, which becomes an excellent quantifier of the entanglement of a state. Gilbert's algorithm is quite versatile in it's applications as it does not make any assumptions about the convex set over which the minimization is done. This allows us in principle to calculate the minimum Hilbert-Schmidt distance of a reference state from the set of k-separable states, by simply switching the search space. The output of the algorithm is the minimum distance of the tested state from the convex set of choice and also a close approximation to the closest state in this set to the tested state. Therefore, it is a simple matter of defining the hyperplane that is tangent to the convex set to get an Entanglement witness for the tested state. This can be done by defining the hyperplane using the approximate closest separable state and then optimizing over the set of pure product states to find a positive operator that when added to this hyperplane, will move it to the boundary of the convex set. This state specific generation of close to optimal Entanglement Witnesses allows us to test full separability to genuine entanglement all using one algorithm. We also gave examples where the algorithm was able to detect Bound Entangled states and provide witnesses to certify their entanglement.

The above analysis of entanglement is possible when the density matrix is known. On the other hand, certifying that a state distributed to some parties that want to use the advertised correlation of the state for various protocols, does indeed possess those correlations becomes a more difficult problem as each party only has access to their own subsystems. The useful information that can be gathered from an individual subsystem is comprised of measurement statistics, and as it turns out there are methods of using these measurement statistics construct a joint conditional probability distributions. If these conditional probabilities violate correlation Bell inequalities, it automatically serves as certification of the state as entangled. This type of verification of entanglement or more appropriately, non-local correlations that does not rely on the un-

derlying systems are termed as Device Independent. Device Independent verification of non-local correlations using a Bell inequality constitutes a proof of unconditional security for any protocol, and as such these methods are of great importance in the field of Quantum Cryptography. The paradigm of Self-testing is the most complete form of device independence known, as it not only verifies the underlying state, it also verifies the measurements performed on that state that produced the measurement statistics. This is usually only possible when the measurement statistics produce correlations that maximally violate a Bell inequality. While there are numerous ways to give self-testing statements or proofs, we proposed a novel approach relying on minimal assumptions of no-signalling and local projective measurements. The approach, independently of the dimension or number of subsystems, lets us transform any linear Bell operator into a strictly anti-diagonal form from which the maximally entangled N-qubit GHZ states can be self-tested, accompanied with self-testing of the measurements as anti-commuting measurement operators. We also demonstrated that the same technique of self-testing proof can be applied to quadratic Bell inequalities, by taking the example of the Uffink's Bell inequality. Non-linear Bell inequalities are of interest because their non-linear nature makes them a tighter witness of non-locality compared to the linear Bell inequalities.

a

APPENDIX

A.1 PAULI MATRICES AND GENERALIZED GELL-MANN MATRICES

We recount the definitions of the extremely useful Pauli matrices and a way to obtain d-dimensional generalized Gell-Mann matrices.

A.1.1 *Pauli Matrices*

$$\sigma_x = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad \sigma_y = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 & \imath \\ -\imath & 0 \end{pmatrix} \quad \text{and} \quad \sigma_z = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{a.1}$$

A.1.2 *Generalized Gell-Mann matrices*

The d-dimensional Generalized Gell-Mann matrices can be obtained using the following procedure. Denote by $E_{j,k}$ the matrix such that the matrix element $(E_{j,k})_{j,k} = 1$ and the rest are zeroes, then there are three groups of matrices that form the set Generalized Gell-Mann matrices, symmetric, antisymmetric and diagonal.

- Symmetric Matrices are given by:

$$\Lambda^s_{j,k} = E_{j,k} + E_{k,j} \quad \text{for} \quad 1 \leq j < k \leq d \tag{a.2}$$

- Antisymmetric matrices:

$$\Lambda^s_{j,k} = \imath(E_{j,k} - E_{k,j}) \quad \text{for} \quad 1 \leq j < k \leq d \tag{a.3}$$

- Diagonal matrices:

$$\Lambda_l^s = \sqrt{\frac{2}{l(l+1)}} \left(\sum_{j=1}^{l} E_{j,j} - l E_{l+1,l+1} \right) \quad \text{for} \quad 1 \leq l \leq n-1. \tag{a.4}$$

BIBLIOGRAPHY

[1] Antonio Acín, Nicolas Gisin, and Lluis Masanes. "From Bell's Theorem to Secure Quantum Key Distribution." In: *Phys. Rev. Lett.* 97 (12 Sept. 2006), p. 120405. DOI: 10.1103/PhysRevLett.97.120405. URL: https://link.aps.org/doi/10.1103/PhysRevLett.97.120405.

[2] Antonio Acín et al. "Device-Independent Security of Quantum Cryptography against Collective Attacks." In: *Phys. Rev. Lett.* 98 (23 June 2007), p. 230501. DOI: 10.1103/PhysRevLett.98.230501. URL: https://link.aps.org/doi/10.1103/PhysRevLett.98.230501.

[3] F. Baccari et al. "Scalable Bell Inequalities for Qubit Graph States and Robust Self-Testing." In: *Phys. Rev. Lett.* 124 (2 Jan. 2020), p. 020402. DOI: 10.1103/PhysRevLett.124.020402. URL: https://link.aps.org/doi/10.1103/PhysRevLett.124.020402.

[4] Jean-Daniel Bancal et al. "Physical characterization of quantum devices from nonlocal correlations." In: *Phys. Rev. A* 91 (2 Feb. 2015), p. 022115. DOI: 10.1103/PhysRevA.91.022115. URL: https://link.aps.org/doi/10.1103/PhysRevA.91.022115.

[5] Somshubhro Bandyopadhyay, Sibasish Ghosh, and Vwani Roychowdhury. "Non-full-rank bound entangled states satisfying the range criterion." In: *Phys. Rev. A* 71 (1 Jan. 2005), p. 012316. DOI: 10.1103/PhysRevA.71.012316. URL: https://link.aps.org/doi/10.1103/PhysRevA.71.012316.

[6] Jonathan Barrett et al. "Nonlocal correlations as an information-theoretic resource." In: *Phys. Rev. A* 71 (2 Feb. 2005), p. 022101. DOI: 10.1103/PhysRevA.71.022101. URL: https://link.aps.org/doi/10.1103/PhysRevA.71.022101.

[7] J. S. Bell. "On the Einstein Podolsky Rosen paradox." In: *Physics Physique Fizika* 1 (3 Nov. 1964), pp. 195–200. DOI: 10.1103/PhysicsPhysiqueFizika.1.195. URL: https://link.aps.org/doi/10.1103/PhysicsPhysiqueFizika.1.195.

[8] Charles H. Bennett and Gilles Brassard. "Quantum cryptography: Public key distribution and coin tossing." In: *Theoretical Computer Science* 560 (2014). Theoretical Aspects of Quantum Cryptography – celebrating 30 years of BB84, pp. 7–11. ISSN: 0304-3975. DOI: https://doi.org/10.1016/j.tcs.2014.05.025. URL: https://www.sciencedirect.com/science/article/pii/S0304397514004241.

[9] Charles H. Bennett et al. "Unextendible Product Bases and Bound Entanglement." In: *Phys. Rev. Lett.* 82 (26 June 1999), pp. 5385–5388. DOI: 10.1103/PhysRevLett.82.5385. URL: https://link.aps.org/doi/10.1103/PhysRevLett.82.5385.

[10] R. A. Bertlmann, H. Narnhofer, and W. Thirring. "Geometric picture of entanglement and Bell inequalities." In: *Phys. Rev. A* 66 (3 Sept. 2002), p. 032319. DOI: 10.1103/PhysRevA.66.032319. URL: https://link.aps.org/doi/10.1103/PhysRevA.66.032319.

[11] Reinhold A. Bertlmann et al. "Optimal entanglement witnesses for qubits and qutrits." In: *Phys. Rev. A* 72 (5 Nov. 2005), p. 052331. DOI: 10.1103/PhysRevA.72.052331. URL: https://link.aps.org/doi/10.1103/PhysRevA.72.052331.

[12] Stephen Brierley, Miguel Navascués, and Tamás Vértesi. "Convex separation from convex optimization for large-scale problems." In: *arXiv: Quantum Physics* (2016).

[13] Kai Chen and Ling-An Wu. "A Matrix Realignment Method for Recognizing Entanglement." In: *Quantum Info. Comput.* 3.3 (May 2003), pp. 193–202. ISSN: 1533-7146.

[14] John F. Clauser et al. "Proposed Experiment to Test Local Hidden-Variable Theories." In: *Phys. Rev. Lett.* 23 (15 Oct. 1969), pp. 880–884. DOI: 10.1103/PhysRevLett.23.880. URL: https://link.aps.org/doi/10.1103/PhysRevLett.23.880.

[15] Valerie Coffman, Joydip Kundu, and William K. Wootters. "Distributed entanglement." In: *Phys. Rev. A* 61 (5 Apr. 2000), p. 052306. DOI: 10.1103/PhysRevA.61.052306. URL: https://link.aps.org/doi/10.1103/PhysRevA.61.052306.

[16] David P DiVincenzo et al. "Unextendible product bases, uncompletable product bases and bound entanglement." In: *Communications in Mathematical Physics* 238.3 (2003), pp. 379–410.

[17] Andrew C. Doherty, Pablo A. Parrilo, and Federico M. Spedalieri. "Detecting multipartite entanglement." In: *Phys. Rev. A* 71 (3 Mar. 2005), p. 032333. DOI: 10.1103/PhysRevA.71.032333. URL: https://link.aps.org/doi/10.1103/PhysRevA.71.032333.

[18] W. Dür and J. I. Cirac. "Classification of multiqubit mixed states: Separability and distillability properties." In: *Phys. Rev. A* 61 (4 Mar. 2000), p. 042314. DOI: 10.1103/PhysRevA.61.042314. URL: https://link.aps.org/doi/10.1103/PhysRevA.61.042314.

[19] W. Dür, G. Vidal, and J. I. Cirac. "Three qubits can be entangled in two inequivalent ways." In: *Phys. Rev. A* 62 (6 Nov. 2000), p. 062314. DOI: 10.1103/PhysRevA.62.062314. URL: https://link.aps.org/doi/10.1103/PhysRevA.62.062314.

[20] Artur K. Ekert. "Quantum cryptography based on Bell's theorem." In: *Phys. Rev. Lett.* 67 (6 Aug. 1991), pp. 661–663. DOI: 10.1103/PhysRevLett.67.661. URL: https://link.aps.org/doi/10.1103/PhysRevLett.67.661.

[21] Matteo Fadel. *Self-testing Dicke states*. 2017. arXiv: 1707.01215 [quant-ph].

[22] Elmer G. Gilbert. "An Iterative Procedure for Computing the Minimum of a Quadratic Form on a Convex Set." In: *SIAM Journal on Control* 4.1 (1966), pp. 61–80. DOI: 10.1137/0304007. eprint: https://doi.org/10.1137/0304007. URL: https://doi.org/10.1137/0304007.

[23] Otfried Gühne and Michael Seevinck. "Separability criteria for genuine multiparticle entanglement." In: *New Journal of Physics* 12.5 (May 2010), p. 053002. DOI: 10.1088/1367-2630/12/5/053002. URL: https://doi.org/10.1088/1367-2630/12/5/053002.

[24] Otfried Gühne and Géza Tóth. "Entanglement detection." In: *Physics Reports* 474.1 (2009), pp. 1–75. ISSN: 0370-1573. DOI: https://doi.org/10.1016/j.physrep.2009.02.004. URL: https://www.sciencedirect.com/science/article/pii/S0370157309000623.

[25] Pawel Horodecki. "Separability criterion and inseparable mixed states with positive partial transposition." In: *Physics Letters A* 232.5 (1997), pp. 333–339. ISSN: 0375-9601. DOI: https://doi.org/10.1016/S0375-9601(97)00416-7. URL: https://www.sciencedirect.com/science/article/pii/S0375960197004167.

[26] F Hulpke and D Bruß. "A two-way algorithm for the entanglement problem." In: *Journal of Physics A: Mathematical and General* 38.24 (June 2005), pp. 5573–5579. DOI: 10.1088/0305-4470/38/24/011. URL: https://doi.org/10.1088/0305-4470/38/24/011.

[27] Lawrence M. Ioannou. "Computational Complexity of the Quantum Separability Problem." In: *Quantum Info. Comput.* 7.4 (May 2007), pp. 335–370. ISSN: 1533-7146.

[28] Lawrence M. Ioannou, Benjamin C. Travaglione, and Donny Cheung. "Convex Separation from Optimization via Heuristics." In: *ArXiv* abs/cs/0603089 (2006).

[29] Sinisa Karnas and Maciej Lewenstein. "Separable approximations of density matrices of composite quantum systems." In: *Journal of Physics A: Mathematical and General* 34.35 (Aug. 2001), pp. 6919–6937. DOI: 10.1088/0305-4470/34/35/318. URL: https://doi.org/10.1088/0305-4470/34/35/318.

[30] M. Lewenstein et al. "Optimization of entanglement witnesses." In: *Phys. Rev. A* 62 (5 Oct. 2000), p. 052310. DOI: 10.1103/PhysRevA.62.052310. URL: https://link.aps.org/doi/10.1103/PhysRevA.62.052310.

[31] Maciej Lewenstein and Anna Sanpera. "Separability and Entanglement of Composite Quantum Systems." In: *Phys. Rev. Lett.* 80 (11 Mar. 1998), pp. 2261–2264. DOI: 10.1103/PhysRevLett.80.2261. URL: https://link.aps.org/doi/10.1103/PhysRevLett.80.2261.

[32] Owidiusz Makuta and Remigiusz Augusiak. "Self-testing maximally dimensional genuinely entangled subspaces within the stabilizer formalism." In: *New Journal of Physics* 23.4 (Apr. 2021), p. 043042. DOI: 10.1088/1367-2630/abee40. URL: https://doi.org/10.1088/1367-2630/abee40.

[33] D. Mayers and A. Yao. "Quantum cryptography with imperfect apparatus." In: *Proceedings 39th Annual Symposium on Foundations of Computer*

Science (Cat. No.98CB36280). 1998, pp. 503–509. DOI: 10.1109/SFCS.1998.743501.

[34] Dominic Mayers and Andrew Chi-Chih Yao. "Self testing quantum apparatus." In: *Quantum Inf. Comput.* 4 (2004), pp. 273–286.

[35] Matthew McKague. "Self-Testing Graph States." In: *Theory of Quantum Computation, Communication, and Cryptography*. Ed. by Dave Bacon, Miguel Martin-Delgado, and Martin Roetteler. Berlin, Heidelberg: Springer Berlin Heidelberg, 2014, pp. 104–120. ISBN: 978-3-642-54429-3.

[36] Michael A. Nielsen and Isaac L. Chuang. *Quantum Computation and Quantum Information: 10th Anniversary Edition*. Cambridge University Press, 2010. DOI: 10.1017/CBO9780511976667.

[37] Tobias J. Osborne and Frank Verstraete. "General Monogamy Inequality for Bipartite Qubit Entanglement." In: *Phys. Rev. Lett.* 96 (22 June 2006), p. 220503. DOI: 10.1103/PhysRevLett.96.220503. URL: https://link.aps.org/doi/10.1103/PhysRevLett.96.220503.

[38] Masanao Ozawa. "Entanglement measures and the Hilbert–Schmidt distance." In: *Physics Letters A* 268.3 (2000), pp. 158–160. ISSN: 0375-9601. DOI: https://doi.org/10.1016/S0375-9601(00)00171-7. URL: https://www.sciencedirect.com/science/article/pii/S0375960100001717.

[39] Palash Pandya, Omer Sakarya, and Marcin Wie śniak. "Hilbert-Schmidt distance and entanglement witnessing." In: *Phys. Rev. A* 102 (1 July 2020), p. 012409. DOI: 10.1103/PhysRevA.102.012409. URL: https://link.aps.org/doi/10.1103/PhysRevA.102.012409.

[40] Ekta Panwar, Palash Pandya, and Marcin Wieśniak. "An elegant proof of self-testing for multipartite Bell inequalities." In: (2022). arXiv: 2202.06908 [quant-ph].

[41] S. Pironio, M. Navascués, and A. Acín. "Convergent Relaxations of Polynomial Optimization Problems with Noncommuting Variables." In: *SIAM Journal on Optimization* 20.5 (2010), pp. 2157–2180. DOI: 10.1137/090760155. eprint: https://doi.org/10.1137/090760155. URL: https://doi.org/10.1137/090760155.

[42] Arthur O Pittenger and Morton H Rubin. "Convexity and the separability problem of quantum mechanical density matrices." In: *Linear Algebra and its Applications* 346.1-3 (2002), pp. 47–71.

[43] Ruben Quesada and Anna Sanpera. "Best separable approximation of multipartite diagonal symmetric states." In: *Phys. Rev. A* 89 (5 May 2014), p. 052319. DOI: 10.1103/PhysRevA.89.052319. URL: https://link.aps.org/doi/10.1103/PhysRevA.89.052319.

[44] Oliver Rudolph. "On the cross norm criterion for separability." In: *Journal of Physics A: Mathematical and General* 36.21 (May 2003), pp. 5825–5825. DOI: 10.1088/0305-4470/36/21/311. URL: https://doi.org/10.1088/0305-4470/36/21/311.

[45] Oliver Rudolph. "Some properties of the computable cross-norm criterion for separability." In: *Physical Review A* 67.3 (2003), p. 032312.

[46] Anna Sanpera, Rolf Tarrach, and Guifré Vidal. "Local description of quantum inseparability." In: *Phys. Rev. A* 58 (2 Aug. 1998), pp. 826–830. DOI: 10.1103/PhysRevA.58.826. URL: https://link.aps.org/doi/10.1103/PhysRevA.58.826.

[47] Pavel Sekatski et al. "Certifying the Building Blocks of Quantum Computers from Bell's Theorem." In: *Phys. Rev. Lett.* 121 (18 Nov. 2018), p. 180505. DOI: 10.1103/PhysRevLett.121.180505. URL: https://link.aps.org/doi/10.1103/PhysRevLett.121.180505.

[48] Michael Steiner. "Generalized robustness of entanglement." In: *Phys. Rev. A* 67 (5 May 2003), p. 054305. DOI: 10.1103/PhysRevA.67.054305. URL: https://link.aps.org/doi/10.1103/PhysRevA.67.054305.

[49] I Šupić et al. "Self-testing multipartite entangled states through projections onto two systems." In: *New Journal of Physics* 20.8 (Aug. 2018), p. 083041. DOI: 10.1088/1367-2630/aad89b. URL: https://doi.org/10.1088/1367-2630/aad89b.

[50] Ivan Šupić and Joseph Bowles. "Self-testing of quantum systems: a review." In: *Quantum* 4 (2020), p. 337.

[51] Barbara M. Terhal. "Bell inequalities and the separability criterion." In: *Physics Letters A* 271.5 (2000), pp. 319–326. ISSN: 0375-9601. DOI: https://doi.org/10.1016/S0375-9601(00)00401-1. URL: https://www.sciencedirect.com/science/article/pii/S0375960100004011.

[52] V. Vedral and M. B. Plenio. "Entanglement measures and purification procedures." In: *Phys. Rev. A* 57 (3 Mar. 1998), pp. 1619–1633. DOI: 10.1103/PhysRevA.57.1619. URL: https://link.aps.org/doi/10.1103/PhysRevA.57.1619.

[53] Frank Verstraete, Jeroen Dehaene, and Bart De Moor. "On the geometry of entangled states." In: *Journal of Modern Optics* 49.8 (2002), pp. 1277–1287. DOI: 10.1080/09500340110115488. eprint: https://doi.org/10.1080/09500340110115488. URL: https://doi.org/10.1080/09500340110115488.

[54] G. Vidal and R. F. Werner. "Computable measure of entanglement." In: *Phys. Rev. A* 65 (3 Feb. 2002), p. 032314. DOI: 10.1103/PhysRevA.65.032314. URL: https://link.aps.org/doi/10.1103/PhysRevA.65.032314.

[55] Guifré Vidal and Rolf Tarrach. "Robustness of entanglement." In: *Phys. Rev. A* 59 (1 Jan. 1999), pp. 141–155. DOI: 10.1103/PhysRevA.59.141. URL: https://link.aps.org/doi/10.1103/PhysRevA.59.141.

[56] Harald Weinfurter and Marek Żukowski. "Four-photon entanglement from down-conversion." In: *Phys. Rev. A* 64 (1 June 2001), p. 010102. DOI: 10.1103/PhysRevA.64.010102. URL: https://link.aps.org/doi/10.1103/PhysRevA.64.010102.

[57] R. F. Werner and M. M. Wolf. "All-multipartite Bell-correlation inequalities for two dichotomic observables per site." In: *Phys. Rev. A* 64 (3 Aug. 2001), p. 032112. DOI: 10.1103/PhysRevA.64.032112. URL: https://link.aps.org/doi/10.1103/PhysRevA.64.032112.

[58] Marcin Wieśniak et al. "Distance between Bound Entangled States from Unextendible Product Bases and Separable States." In: *Quantum Reports* 2.1 (2020), pp. 49–56. ISSN: 2624-960X. DOI: 10.3390/quantum2010004. URL: https://www.mdpi.com/2624-960X/2/1/4.

[59] Xingyao Wu et al. "Robust self-testing of the three-qubit *W* state." In: *Phys. Rev. A* 90 (4 Oct. 2014), p. 042339. DOI: 10.1103/PhysRevA.90.042339. URL: https://link.aps.org/doi/10.1103/PhysRevA.90.042339.

[60] Tzyh Haur Yang et al. "Robust and Versatile Black-Box Certification of Quantum Devices." In: *Phys. Rev. Lett.* 113 (4 July 2014), p. 040401. DOI: 10.1103/PhysRevLett.113.040401. URL: https://link.aps.org/doi/10.1103/PhysRevLett.113.040401.

[61] Marek Żukowski and Časlav Brukner. "Bell's Theorem for General N-Qubit States." In: *Phys. Rev. Lett.* 88 (21 May 2002), p. 210401. DOI: 10.1103/PhysRevLett.88.210401. URL: https://link.aps.org/doi/10.1103/PhysRevLett.88.210401.

[62] Karol Zyczkowski and Hans-J rgen Sommers. "Hilbert–Schmidt volume of the set of mixed quantum states." In: *Journal of Physics A: Mathematical and General* 36.39 (Sept. 2003), pp. 10115–10130. DOI: 10.1088/0305-4470/36/39/310. URL: https://doi.org/10.1088/0305-4470/36/39/310.