Witnessing entanglement in qudit systems

Vom Fachbereich Physik der Universität Hannover
zur Erlangung des Grades
Doktor der Naturwissenschaften
Dr. rer. nat.
genehmigte Dissertation
von

MSc Philipp Hyllus
geboren am 4. Mai 1976 in Wolfenbüttel in Niedersachsen

2005

Referent: Prof. Dr. M. Lewenstein
Korreferentin: Prof. Dr. D. Bruß

Tag der Promotion: 17. Januar 2005

Abstract

In this thesis, we deal with several aspects of the theory of entanglement, all of which are connected to the problem of finding ways to witness the presence of entanglement in a system of qudits, i.e., d-level quantum systems.

After having recalled some basic facts and definitions concerning the theory of entanglement, we concentrate on the local detection of entanglement via witness operators. A negative expectation value of these operators signals the presence of entanglement because their expectation value is positive with respect to all separable states. We discuss known ways and introduce a new simple method to construct witnesses. In this form, they are not easily applicable in an experiment, because they require measurements on the total system. Thus, we construct local decompositions such that the witnesses can be measured with local measurements only, and minimize the number of local measurements necessary. We concentrate on witnesses in high dimensional bipartite systems, and on witnesses for bound entanglement, a weak form of entanglement, also in multipartite systems. We show how to estimate the number of local measurements necessary for a witness and prove optimality in some cases. Finally, we briefly summarize results of the experimental implementation of witnesses, which has been performed in the group of H. Weinfurter in Munich. Using witnesses, it has been experimentally proven there that certain states of three and four photons are multipartite entangled in the polarization degrees of freedom.

Then we introduce simple networks for the experimental generation of bound entangled states of three 2-level systems. We show how the entanglement can be proven experimentally via locally decomposed witness operators and discuss ways to check the biseparability properties of the states which are responsible for the bondage of the entanglement.

Following this, we investigate optimization problems occuring in entanglement theory from the point of view of convex optimization. We show how the problems can be written such that recently obtained known results from the theory of semi-definite relaxations can be applied. This leads to a complete hierarchy of approximations to the optimal solutions. Applications include witnesses operators for bound entangled states, a known measure of entanglement for multipartite pure states, as well as a new entanglement criterion for multipartite systems of qudits.

Finally, we discuss the relationship between witness operators and Bell inequalities, which give bounds on the maximal correlations that can occur in any local and realistic theory. Formulated in the language of quantum mechanics, these inequalities can be written as witnesses. We investigate the relation in detail for a Bell inequality for two 2-level systems.

Keywords: Entanglement, Entanglement witnesses, Bound entanglement, Nonconvex optimization, Bell inequalities

Zusammenfassung

In dieser Arbeit behandeln wir mehrere Aspekte der Verschränkungstheorie, die alle einen Bezug haben zum Problem des Verschränkungsnachweises in Systemen bestehend aus mehreren „qudits“, d.h. quantenmechanischen d-Niveau Systemen.

Wir beginnen mit einer Einführung in die Grundbegriffe der Verschränkungstheorie, und wenden uns dann dem Verschränkungsnachweis mit Hilfe von lokal zerlegten Zeugenoperatoren zu. Ein negativer Erwartungswert dieser Operatoren weist Verschränkung nach, da die Operatoren einen positiven Erwartungswert bezüglich aller separierbaren Zustände haben. Wir erläutern bekannte Konstruktionen und führen eine neue einfache Methode zur Konstruktion von Verschränkungszeugen ein. In dieser Form sind die Zeugen experimentell nicht einfach anwendbar, da sie Messungen am Gesamtsystem erfordern. Deswegen zerlegen wir die Operatoren lokal, so daß sich der Erwartungswert des Zeugen mit mehreren lokalen Messungen messen lässt, und minimieren die Anzahl der nötigen lokalen Messungen. Wir betrachten Zeugen in Zweiparteiensystemen hoher Dimension sowie Zeugen zur Detektion von gebundener Verschränkung, einer schwer nachweisbaren Form der Verschränkung, auch für Mehrparteiensysteme. Wir gewinnen Abschätzungen für die minimale Anzahl der lokalen Messungen und beweisen in einigen Fällen, daß die Zerlegungen optimal sind. Schließlich berichten wir in Kürze von Experimenten, die in der Gruppe von Harald Weinfurter in München durchgeführt worden sind. Dabei wurde die Mehrparteienverschränkung von Zuständen von drei bzw. vier Photonen in den Polarisationsfreiheitsgeraden mit Hilfe von lokal zerlegten Zeugenoperatoren nachgewiesen.

Dann präsentieren wir einfache Netzwerke zur Erzeugung von gebunden verschränkten Zuständen von drei Zweiniveausystemen. Wir zeigen, wie die Verschränkung experimentell mit lokal zerlegten Zeugenoperatoren nachgewiesen werden kann, und vergleichen drei Methoden zum Test der Biseparabilitätseigenschaften, die mit der Gebundenheit der Verschränkung zusammenhängen.

Danach betrachten wir mehrere Optimierungsprobleme, die in der Verschränkungstheorie auftauchen vom Standpunkt der konvexen Optimierungstheorie aus. Wir zeigen, daß die Probleme so umgeschrieben werden können, daß kürzlich gewonnene Erkenntnisse der Theorie der semi-definiten Relaxationen angewandt werden können. Damit erzeugt man eine Hierarchie von Annäherungen an die optimale Lösung. Beispiele, bei denen solche Optimierungsprobleme auftreten, sind Zeugenoperatoren für gebunden verschränkte Zustände, ein Verschränkungsmaß für reine Mehrparteienzustände, sowie ein neues Verschränkungskriterium für Systeme beliebiger Dimension und Parteienzahl.

Am Ende der Arbeit wenden wir uns dem Zusammenhang von Zeugenoperatoren und Bell'schen Ungleichungen zu, die die maximalen Korrelationen begrenzen, die in einer lokalen und realistischen Theorie auftreten können. In der Sprache der Quantenmechanik entsprechen diese Ungleichungen Zeugenoperatoren. Wir erforschen diese Beziehung detailliert für eine Bell Ungleichung für zwei Zweiniveausysteme.

Schlagworte: Verschränkung, Verschränkungszeugen, gebundene Verschränkung, Nicht-konvexe Optimierung, Bell Ungleichungen.

Contents

Introduction

One of the most remarkable features that distinguishes quantum mechanics from classical mechanics is entanglement. Entanglement refers to quantum correlations between separated physical systems that can be stronger than correlations allowed by classical mechanics. The possibility of such correlations was observed already in the early days of quantum theory by Einstein, Podolsky and Rosen (EPR) [1], who used it to argue that quantum mechanics could not be regarded as a complete physical theory. The term entanglement itself was coined by Schrödinger, in a reaction to the EPR contribution [2]. Three decades later, Bell succeeded in constructing inequalities that *any* theory fulfilling the basic assumptions that EPR used in their argument has to obey [3]. Even more, he showed that entangled states can violate these inequalities, thereby ruling out the possibility of unifying EPRs beliefs about the way that physical theories have to be constructed and quantum mechanics.

The attitude towards entanglement changed in the end of the last century from being focused on the fundamental implications to questions of more practical nature, and it was realized that quantum systems might be used to perform tasks impossible or very hard for classical systems. Along these lines, Feynmann suggested to use quantum systems to simulate other, more complicated quantum systems [6], a very hard task for a classical computer.

Shortly after, algorithms based on the laws of quantum mechanics were found that could solve certain tasks faster than any classical computing device, founding the field of quantum computation. The first of these algorithms was due to Deutsch [7]. Further prominent examples are Shor's algorithm for factorizing prime numbers [8] and the search algorithm of Grover [9].

In parallel, other potential applications were developed. The first protocol for the secure transmission of a random secret key using nonorthogonal polarization states of photons was proposed by Bennett and Brassard [10], founding the field of quantum cryptography. The first protocol for secret key distribution using entangled states was proposed few years later by Ekert [11]. Other applications include the teleportation of quantum states [12], quantum dense coding, enabling the transmission of two classical bits by sending only one quantum bit or *qubit* if the two parties shared an entangled state before [13], and quantum communication complexity protocols, where several parties have to estimate a function separately with restricted communication only [14].

Along with the theoretical discoveries came significant progress on the experimental side with respect to the capabilities of controlling and manipulating elementary quantum systems in various experimental set ups. For example, Shor's algorithm has been implemented, factorizing 15, in a liquid state nuclear magnetic resonance (NMR) system [15]. Using NMR control techniques, Deutsch's algorithm has been performed recently in an ion trap [16]. An extensive list of achievements can be found in Refs. [17–19].

Up to now, it is not clear what the crucial ingredient for the success of all these applications is. However, there is strong evidence that entanglement plays a very important role. For instance, it was shown that it is necessary for quantum key distribution [20, 21]. Hence entanglement is interesting both from a fundamental as well as from a practical point of view. Further, entanglement is not only a theoretical construct, it has been realized in the laboratory. Even entangled states of more than two subsystems have been generated in several set ups, e.g., using the polarization degree of freedom of photons [22] or internal degrees of freedom of trapped ions [23]. Therefore, the characterization of entangled states is of great importance in many respects.

In this thesis, we deal with several aspects of the theory of entanglement, all of which are connected to the problem of finding ways to witness the presence of entanglement. In particular, we discuss the local detection of entanglement via witness operators, the generation and detection of so-called bound entangled states, which is a particular weak form of entanglement, complete hierarchies of efficient approximations to typical optimization problems in entanglement theory, as well as the relation between witness operators and Bell inequalities.

The thesis is organized as follows:

In chapter 1 we introduce the basic notions needed for the understanding of the rest of the thesis. We define entanglement of pure and mixed states and introduce ways to classify the state space for bipartite systems as well as for multipartite systems with respect to entanglement properties. Further, we introduce criteria for entanglement, in particular witness operators. Finally, we give a brief introduction to Bell's inequalities.

Then, we concentrate on the construction and local decomposition of witness operators in chapter 2. We focus on systems consisting of two parties, which can be of arbitrary dimension, and bound entangled states. Further, we construct new families of bound entangled states for multiqubit systems, and show how their entanglement can be detected in a local way with entanglement witnesses. We also present results of experiments performed by Mohamed Bourennane and coworkers in the group of Harald Weinfurter in Munich implementing witnesses in multiqubit systems. The results presented here are based on Refs. [II,V-VIII].[1]

In the following chapter 3, we turn our attention to the experimental generation of bound entanglement. We construct simple networks that generate two families of bound entangled states of three qubits. The motivation is that despite being

[1] References in roman numerals refer to the publication list on page 103.

interesting from a fundamental point of view such states have not been produced in the laboratory so far. We further provide ways to prove that the states are indeed bound entangled. This chapter is based on Ref. [IX].

In chapter 4, we discuss ways to solve typical optimization problems occuring in entanglement theory. For example, in the construction of witness operators for bound entangled states typically minimizations of the expectation value of an operator with respect to product states have to be performed. We show that such problems can be formulated as minimizations of a linear function subject to polynomial constraints of a degree of at most three, or subject to a semidefinite constraint and polynomial constraints of a degree no larger than two. We then apply recently obtained known results from the theory of semi-definite relaxations to the formulated optimization problems. These approximate the original computationally hard problems by a hierarchy of efficiently solvable semidefinite programms. In the formulation that involves only polynomial constraints, the solution of the original problem is obtained asymptotically.

The results, which are based on Ref. [X], are very useful from a practical point of view. Further, they are also interesting from a fundamental theoretical point of view, because it is possible to obtain a new criterion for entanglement of mixed quantum states, for any dimension and for any number of parties.

Finally, in chapter 5, we return to the origin of the introduction: the point raised by EPR. Formulated in the language of quantum mechanics, the Bell inequalities correspond to witness operators. In the last chapter, we investigate this relationship in detail for a prominent two qubit Bell inequality.

Chapter 1

Entanglement

In this chapter we give an introduction to the main concepts which are needed in the following chapters. The intention is to give an overview over the basic notions needed for the rest of the thesis, so we will in general omit proofs and refer to the literature, while discussing more technical details in later chapters in case they are needed.

We introduce the concept of entanglement explicitly for pure states of two physical systems only, and generalize this to mixed states where both quantum and classical correlations occur. We also mention criteria designed to distinguish classical from quantum correlations, with an emphasis on entanglement witnesses, which are the main objects of interest of this thesis. Finally, we introduce the concept of Bell inequalities.

1.1 Bipartite entanglement

1.1.1 Pure states

A pure quantum state of a single quantum system is described by a state vector $|\psi\rangle$ which is an element of a N-dimensional Hilbert space $\mathcal{H}$, where N corresponds to the number of degrees of freedom of that system. Here, as well as in the rest of the thesis, we consider only systems with a finite number of degrees of freedom. For such a space there exists an orthonormal basis $\{|i\rangle\}_{i=1}^{N}$. Every state can be written in this basis as

$$|\psi\rangle = \sum_{i=1}^{N} \psi_i |i\rangle, \quad \text{where} \quad \psi_i = \langle i|\psi\rangle. \tag{1.1}$$

In analogy, we can also describe two separate systems, with N and M degrees of freedom, say. Let the first be in a state $|\psi\rangle \in \mathcal{H}_A$ and the second in a state $|\phi\rangle \in \mathcal{H}_B$, where the dimensions of the Hilbert space are again given by the respective numbers of the degrees of freedom. The state of the composite system can then be described with the help of the tensor product, the composite state $|\psi\rangle \otimes |\phi\rangle$ lives in the composite Hilbert space $\mathcal{H}_A \otimes \mathcal{H}_B = \mathcal{H}_{\text{composite}}$ with the dimension $N \times M$.

The tensor product has the properties that for arbitrary states $|\psi_{1,2}\rangle \in \mathcal{H}_A$, $|\phi_{1,2}\rangle \in \mathcal{H}_B$, and for any $\alpha \in \mathbb{C}$ the identities

$$\begin{aligned} (|\psi_1\rangle + |\psi_2\rangle) \otimes |\phi_1\rangle &= |\psi_1\rangle \otimes |\phi_1\rangle + |\psi_2\rangle \otimes |\phi_1\rangle & (1.2)\\ |\psi_1\rangle \otimes (|\phi_1\rangle + |\phi_2\rangle) &= |\psi_1\rangle \otimes |\phi_1\rangle + |\psi_1\rangle \otimes |\phi_2\rangle & (1.3)\\ \alpha|\psi_1\rangle \otimes |\phi_1\rangle &= (\alpha|\psi_1\rangle) \otimes |\phi_1\rangle = |\psi_1\rangle \otimes (\alpha|\phi_1\rangle) & (1.4) \end{aligned}$$

hold.

A particular *product* basis for the composite system can be constructed from the tensor products of the bases of the individual systems, i.e., $\{|i\rangle \otimes |j\rangle\}_{(i,j)=(1,1)}^{(N,M)}$ is a basis of the composite system $\mathcal{H}_A \otimes \mathcal{H}_B$.

Using these definitions, we can state

Definition 1.1. A pure product state of the form $|\psi\rangle \otimes |\phi\rangle \in \mathcal{H}_A \otimes \mathcal{H}_B$ is called *separable*. A pure state $|\Psi\rangle \in \mathcal{H}_A \otimes \mathcal{H}_B$ which cannot be written as a product state is called *entangled*.

Separable states can be prepared by individual, independent local actions of each of the two parties alone. For a system in an entangled state, each subsystem is not in a definite state anymore, only the composite state is well defined.

If we want to write an entangled state of a bipartite state in terms of a product basis, we can use the freedom of choice of the local bases to obtain a very convenient form with the Schmidt decomposition. Before we state it, we introduce the singular value decomposition [24, 25]:

Theorem 1.2. Let C be a complex $N \times M$ matrix of rank k. Then there exist a unitary matrix U of dimension $N \times N$, a unitary matrix V of dimension $M \times M$, and k positive numbers λ_j, the *singular values* of C, such that

$$C = UDV^\dagger, \tag{1.5}$$

where D is a $N \times M$ matrix where the only nonvanishing entries are the k decreasingly ordered entries λ_j on the diagonal. Between the coefficients λ_j and the entries of the matrices U and V the following relations hold: The columns of U are given by the vectors $|u_j\rangle$ fulfilling $CC^\dagger|u_j\rangle = \lambda_j^2|u_j\rangle$, while the columns of V are given by the vectors $|v_j\rangle$ fulfilling $C^\dagger C|v_j\rangle = \lambda_j^2|v_j\rangle$.

Theorem 1.3 (Schmidt decomposition). Let $|\Psi\rangle$ be a state of a composite system in the Hilbert space $\mathcal{H}_A \otimes \mathcal{H}_B$ of dimension $N \times M$. Then there exist bases $\{|i\rangle\}_{i=1}^N$ of $\mathcal{H}_A$ and $\{|\tilde{i}\rangle\}_{i=1}^M$ of $\mathcal{H}_B$ such that

$$|\Psi\rangle = \sum_{i=1}^{r} a_i |i\rangle_A \otimes |\tilde{i}\rangle \tag{1.6}$$

and $a_i > 0$. The number r is called the *Schmidt rank* of $|\Psi\rangle$. It cannot exceed the minimum of N and M, i.e., $r \leq \min(N, M)$. For product states, $r = 1$.

Proof. Expanding the state in a product basis and applying the singular value decomposition to the coefficient matrix it follows that

$$\begin{aligned} |\psi\rangle &= \sum_{i,j} C_{ij}|i\rangle \otimes |j\rangle = \sum_{m,n,i,j} U_{im} D_{mn} V^\dagger_{nj} |i\rangle \otimes |j\rangle \\ &= \sum_m \lambda_m \big(\sum_i U_{im}|i\rangle\big) \otimes \big(\sum_j V^*_{jm}|j\rangle\big) = \sum_m \lambda_m |\tilde{m}\rangle \otimes |\bar{m}\rangle. \end{aligned}$$

The vectors $|\tilde{m}\rangle = \sum_i U_{im}|i\rangle$ are orthogonal because U is unitary, and the vectors $|\bar{m}\rangle = \sum_j V^*_{jm}|j\rangle$ are orthogonal because V is unitary. Further, the summation over m goes from 1 to k, and k cannot exceed the minimum of the number of columns N and the number of rows M of the matrix of coefficients C. □

Here and in the following, $*$ denotes complex conjugation.

The state

$$|\Psi_{\max}\rangle = \frac{1}{\sqrt{d}} \sum_{i=1}^{d} |ii\rangle \tag{1.7}$$

is usually called *maximally* entangled. A rather intuitive justification is that it has the maximal Schmidt rank and balanced Schmidt coefficients. However, there are more concrete reasons for this, some of which we will name below. For two qubits, the following maximally entangled states are denoted as *Bell* states

$$|\psi^\pm\rangle = \frac{1}{\sqrt{2}}(|01\rangle \pm |10\rangle) \tag{1.8}$$

$$|\phi^\pm\rangle = \frac{1}{\sqrt{2}}(|00\rangle \pm |11\rangle). \tag{1.9}$$

The state $|\psi^-\rangle$ is special because it is invariant under local unitary transformations of the form $U \otimes U$ and is referred to as the *singlet* state. Note that the Bell states can be transformed into each other just by changing the local bases.

1.1.2 Mixed states

Let us go back to a single system. We can imagine a source that produces distinct states $|\psi_i\rangle$ with probabilities p_i. In order to calculate expectation values of observables with respect to the average state produced by the source, the density matrix

$$\rho = \sum_i p_i |\psi_i\rangle\langle\psi_i| \tag{1.10}$$

can be used. From the definition it follows that ρ fulfills

$$(i)\ \rho \geq 0 \quad (ii)\ \rho = \rho^\dagger \quad \text{and } (iii)\ \mathrm{Tr}[\rho] = 1. \tag{1.11}$$

Here (i) means that all eigenvalues of ρ are positive semi-definite, while (iii) holds because $\mathrm{Tr}|\psi\rangle\langle\psi| = 1$ for any pure state $|\psi\rangle$, and $\sum_i p_i = 1$ by definition.

For systems a single qubit, the density operator can be conveniently written as

$$\rho = \frac{1}{2}\big(\mathbb{1} + \mathbf{s}\cdot\boldsymbol{\sigma}\big) \tag{1.12}$$

with the help of the *Pauli matrices*

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \text{and} \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{1.13}$$

Together with the identity $\mathbb{1}$, they form a basis of the space of 2×2 dimensional Hermitean matrices. In this form, the condition $\mathrm{Tr}[\rho] = 1$ is already ensured by the prefactor $1/2$. Further, the condition $\rho \geq 0$ requires that $s = |\mathbf{s}| \leq 1$. This can be easily seen by expanding ρ in the basis of eigenstates of $\mathbf{s}\cdot\boldsymbol{\sigma}$ as

$$\rho = \frac{1}{2}\big((1+s)|\mathbf{s}+\rangle\langle\mathbf{s}+| + (1-s)|\mathbf{s}-\rangle\langle\mathbf{s}-|\big). \tag{1.14}$$

Hence the pure states are at the border ($s = 1$) of the so-called *Bloch ball* of vectors $\mathbf{s}$.

If Alice possesses a source of qubits and wants to know the density matrix describing the source, then she can obtain it by measuring the expectation values of Stern-Gerlach type experiments oriented along the x, y, and z directions, corresponding to the σ_x, σ_y, and σ_z operators. From these expectation values she can infer the vector $\mathbf{s}$,

$$\mathrm{Tr}[\rho\sigma_i] = s_i, \tag{1.15}$$

for $i = x, y, z$.

Let us come back to bipartite systems again. We can also write any density matrix of a two qubit system with the help of the Pauli matrices as

$$\rho = \frac{1}{4}\sum_{i,j=0}^{3} \lambda_{ij}\sigma_i \otimes \sigma_j, \tag{1.16}$$

where $\sigma_0 = \mathbb{1}$ and $\sigma_{1,2,3} = \sigma_{x,y,z}$. In this case, the tensor products of Pauli matrices and the identity span the space of Hermitean 4×4 matrices. The $(0,0)$ element of the λ matrix has to be equal to 1 due to the normalization of ρ, and the other are calculated as follows:

$$\lambda_{kl} = \mathrm{Tr}[\rho(\sigma_k \otimes \sigma_l)]. \tag{1.17}$$

In fact, if Alice and Bob share a system in the state ρ, then they can obtain the elements of the λ matrix, and in effect the density matrix itself, just by local measurements due to the tensor structure. For instance, $\mathrm{Tr}[\rho(\sigma_x \otimes \sigma_y)]$ corresponds to the expectation value of Alice performing a Stern-Gerlach experiment directed along the x direction on her part of the system and Bob along the y direction on his part.

In fact, any density operator can be written in terms of a basis of tensor products of single particle operators, enabling the local measurement of the density matrix. This is referred to a quantum state tomography.

We can now extend the definition 1.1 to mixed states:

Definition 1.4 [26]. A density matrix $\rho \in \mathcal{H}_A \otimes \mathcal{H}_B$ is separable iff it can be written as a convex combination of pure product states, i.e.,

$$\rho = \sum_i p_i |\psi_i\rangle\langle\psi_i| \otimes |\phi_i\rangle\langle\phi_i| \tag{1.18}$$

with $\sum_i p_i = 1$. Otherwise, ρ is entangled.

The question arises: Given a state ρ, is it entangled or not? The opposite question regarding to whether a state is separable or not is the so-called *separability problem*. Before we state one of the most important separability criteria, the *positive partial transpose* (PPT) criterion, we define the operation of partial transposition.

Definition 1.5. If ρ is the state of a composite system living in a Hilbert space $\mathcal{H}_A \otimes \mathcal{H}_B$ of dimension $N \times M$, then the transposition with respect to subsystem A is defined as

$$\rho^{T_A} = \sum_{ijkl} \rho_{ik,jl} (|i\rangle\langle j|)^T \otimes |k\rangle\langle l| = \sum_{ijkl} \rho_{ik,jl} |j\rangle\langle i| \otimes |k\rangle\langle l|. \tag{1.19}$$

The partial transpose is basis dependent, but the eigenvalues of the partially transposed state are not. For bipartite systems, $(\rho^{T_B})^{T_A} = \rho^T = \rho^* \geq 0$, so that $\rho^{T_A} \geq 0$ implies $\rho^{T_B} \geq 0$ and vice versa.

If a density operator fulfils $\rho^{T_A} \geq 0$, it is common to say "ρ has a PPT", or even "ρ is PPT."

Now we can state the

Theorem 1.6 (Peres-Horodecki). A state ρ of a bipartite system of dimension 2×2 or 2×3 is separable iff $\rho^{T_A} \geq 0$ [27, 28].

For systems of higher dimensions, the positive partial transpose is only a necessary condition for separability. That it is necessary for any dimension is easy to see because for a separable state of the form of Eq. (1.18),

$$\sum_i p_i |\psi_i\rangle\langle\psi_i|^T \otimes |\phi_i\rangle\langle\phi_i| = \sum_i p_i |\psi_i^*\rangle\langle\psi_i^*| \otimes |\phi_i\rangle\langle\phi_i| \geq 0, \tag{1.20}$$

holds.

That the criterion is only necessary for systems of higher dimensions or those consisting of more than two parties implies that there exist states of those systems which are PPT but nevertheless entangled. A criterion which we will use later to detect the entanglement of PPT entangled states is

Theorem 1.7 (*range* criterion) [29]. Let ρ act on a Hilbert space $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B$, $\dim(\mathcal{H})=M$. If ρ is separable then there exists a set of product vectors $\{|\psi_i\rangle \otimes |\phi_j\rangle\}$ with $n \leq M^2$ pairs of indices (i, j) and probabilities p_{ij} such that

1. $\rho = \sum_{i,j} p_{ij} |\psi_i\rangle\langle\psi_i| \otimes |\phi_j\rangle\langle\phi_j|$ and $\rho^{T_A} = \sum_{i,j} p_{ij} |\psi_i^*\rangle\langle\psi_i^*| \otimes |\phi_j\rangle\langle\phi_j|$,

2. the vectors $\{|\psi_i\rangle \otimes |\phi_j\rangle\}$ ($\{|\psi_i^*\rangle \otimes |\phi_j\rangle\}$) span the range of ρ (ρ^{T_A}), in particular any of the vectors $\{|\psi_i\rangle \otimes |\phi_j\rangle\}$ ($\{|\psi_i^*\rangle \otimes |\phi_j\rangle\}$) belongs to the range of ρ (ρ^{T_A}).

A criterion related to the PPT criterion, which is also based on a reordering of the density matrix, is the *cross norm* [30] or *realignment* [31] criterion. Other operational criteria include the *reduction* criterion [32], the *majorization* criterion [33], criteria for low rank density matrices in $2 \times N$ [34] and $2 \times 2 \times N$ [35] dimensions, as well as recently established criteria based on semidefinite programming [36–39]. Such a criterion [X] is part of this thesis and will be discussed in detail in chapter 4.

All the criteria mentioned above are operational in that given a density operator ρ, the criteria can be evaluated directly. A further strong nonoperational criterion based on positive maps can be found in Ref. [28]. In addition, there are several criteria which do not require the complete knowledge of the density matrix. In the following subsection, we discuss a prominent example of these: the entanglement witnesses [28, 40]. In chapter 2, we discuss how they can be applied in experiments via local decompositions, where we will also compare this method to others which do not require the knowledge of the complete density matrix either.

1.1.3 Entanglement witnesses

From the definition of a density operator in Eq. (1.10) it follows that the set $\mathcal{M}$ of all physical states ρ fulfilling the conditions (1.11) is a convex set. It is a subset of the real vector space of Hermitean operators acting on a Hilbert space $\mathcal{H}$, which we denote as $\mathcal{HS}$. Furthermore, the set is bounded and closed. In addition, it follows

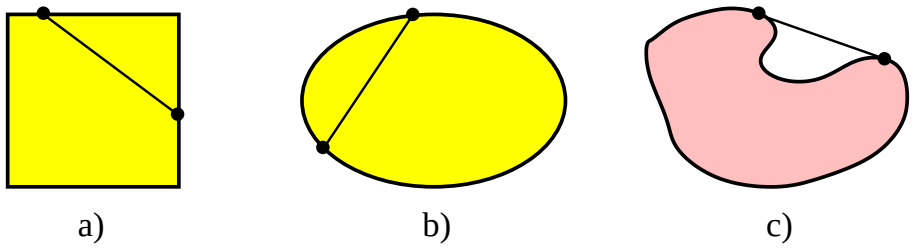

Figure 1.1. *The sets a) and b) are convex, because the line connecting any two points belonging to the set lies within the set. This is not the case for set c).*

from the definition (1.18) of separable states that the set S of separable states is a convex, bounded, and closed subset of $\mathcal{M}$, with the projectors onto product vectors as extremal points.

Basic separation theorems of convex analysis ensure [41] that for each state ρ which does *not* belong to S there exists a hyperplane separating ρ from S. This can also be extended to infinite dimensional systems by using a corollary of the Hahn-Banach theorem [28, 42]. The condition $\mathrm{Tr}[\rho] = 1$ restricts the set of density matrices to an

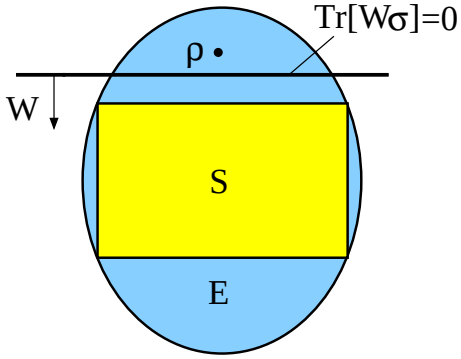

Figure 1.2. *Illustration: the set S is a convex subset of the set M of all states, and $E = \mathcal{M} \setminus S$ is the set of entangled states. The state ρ is separated from S by the hyperplane consisting of all states σ with* $\mathrm{Tr}[W\sigma] = 0$.

affine hyperplane in $\mathcal{HS}$, i.e., a hyperplane not containing the 0 element. Therefore, the hyperplane separating ρ from S can be chosen as a linear hyperplane $\mathcal{L}$, i.e., a hyperplane including the 0 element. By virtue of the Hilbert Schmidt scalar product $\langle A, B\rangle = \mathrm{Tr}[A^\dagger B] = \mathrm{Tr}[AB]$ it can be parametrized with an operator $W = W^\dagger$ as consisting of all the states σ for which the scalar product with W vanishes,

$$\mathcal{L} = \{\sigma \in \mathcal{M} \mid \mathrm{Tr}[W\sigma] = 0\}. \tag{1.21}$$

It is convenient to choose the direction of W such that $\mathrm{Tr}[W\sigma_s] \geq 0$ for all separable states σ_s. Now we can formulate [28, 40]

Theorem 1.8. For every entangled state ρ there exists an Hermitean *entanglement witness* W such that

$$\mathrm{Tr}[W\rho] < 0 \tag{1.22}$$

$$\mathrm{Tr}[W\sigma_s] \geq 0 \quad \text{for all} \quad \sigma_s \in S. \tag{1.23}$$

Further useful definitions are

Definition 1.9 [43]. A witness W_1 is *finer* than a witness W_2 iff $\mathrm{Tr}[W_2\rho] < 0 \Rightarrow \mathrm{Tr}[W_1\rho] < 0$, i.e., if it detects all states that W_2 detects – and more. A witness is *optimal* if there is no finer witness. Criteria for optimality can be found in [43].

In order to characterize the set of separable states, an infinite number of optimal witnesses is necessary. This is because there exist infinitely many extremal vectors $|\phi\rangle$ of S, and further there are vectors $|\psi\rangle \notin S$ arbitrarily close to each $|\phi\rangle$. Hence if nothing is known about a state, an infinite number of witnesses is necessary to detect its entanglement in the worst case. That is why this method of entanglement detecting is only useful from a practical point of view if some knowledge can be assumed about the state. We will discuss how to construct witnesses explicitly in chapter 2.

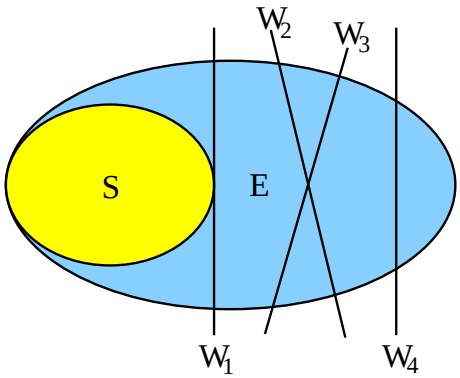

Figure 1.3. *Illustration: Both W_2, and W_3 are finer than W_4, but neither is finer than the other. W_1 is optimal and finer than all other witnesses. In this pictures not all extremal points of S are on the outside. Note that for 2 qubits already, the state space is 15-dimensional, turning the production of a proper illustration into a daunting task.*

1.1.4 Distillability, bound entanglement, and entanglement quantification

The distillability problem concerns the question: If we are given m copies of a bipartite state ρ, can we transform those copies to n copies of the singlet state $|\psi^-\rangle = (|01\rangle - |10\rangle)/\sqrt{2}$ by local operations and classical communications (LOCC) only [44]? It turns out that for bipartite pure states, $\rho = |\psi\rangle\langle\psi|$, it is always possible. In the asymptotic limit $n \to \infty$ the ratio n/m approaches the von Neumann entropy [18]

$$S(\rho_A) = -\mathrm{Tr}[\rho_A \log \rho_A] \tag{1.24}$$

of the reduces state

$$\rho_A = \mathrm{Tr}_B |\psi\rangle\langle\psi| \tag{1.25}$$

[45]. This rate is called the entanglement of distillation. It is now a natural to ask how many copies n of singlet states are needed to obtain m copies of a state ρ. The asymptotic ratio n/m is called entanglement of formation [44]. For bipartite pure states, it can be shown that both the entanglement of formation as well as the entanglement of distillation are given by the reduced von Neumann entropy [45].

For the singlet state, the reduced state ρ_A is given by $\mathbb{1}_2/2$, hence $S(\rho_A) = 1$. For all other two qubit states, $S \in [0, 1]$, which is one of the reasons why the state can be considered to be a maximally entangled state of two qubits. A motivation for choosing it as a reference state is that it is the resource necessary for quantum information tasks like teleportation [12]. Hence with a supply of m states $|\psi\rangle$ the same quantum information tasks can be performed as with n singlet states.

In conclusion, any bipartite pure state can be asymptotically reversibly transformed into any other, the achievable rates being given by the von Neumann entropy of entanglement. In this sense, any bi-partite entanglement of pure states is essentially

equivalent to that of the singlet state, which forms the so-called minimal reversible entanglement-generating set (MREGS) [46]. The situation is very different in the multi-partite case, where the MREGS have not even been identified for three-qubit systems, let alone for more general settings [47]. In the view of this fact, several more pragmatic (and inequivalent) measures of entanglement have been proposed, reasonably grasping the degree of multi-particle entanglement [48–50]. To evaluate these quantities typically amounts to solving a computationally hard problem. In chapter 4, we show how one of them, the geometric measure of entanglement [49], can be approximated efficiently.

Coming back to the distillability problem, it is surprising that the natural assumption that all entangled states are distillable to the singlet form holds for systems of two qubits or of one qubit and a three-level system (qutrit) only. This is related to the fact that a necessary condition for distillability of a state ρ is that it has a non-positive partial transpose (NPPT) [51]. However, as mentioned above, for bipartite systems of a dimensions higher than 2×2 or 2×3 there exist entangled states with a PPT [29]. Further, there is evidence that there exist even NPPT states which cannot be destilled [52, 53]. Undistillable states are called *bound* entangled [51].

Apart from being interesting from a fundamental point of view, bound entangled states are useful for certain quantum information processing tasks: they can activate the distillability of one copy of a bipartite state with non-positive partial transpose [54, 55]. It has also recently been shown that one can extract a secure key from bound entangled states [56]. In chapter 2, we construct and decompose witnesses for such states, and design networks for the generation of two families of three-qubit bound entangled states in chapter 3.

1.2 Multipartite entanglement

The basic definitions of separability and entanglement of the last section can easily be generalized for more than two parties.

Definition 1.10. A pure state of n parties $|\psi\rangle \in \mathcal{H}_1 \otimes \mathcal{H}_2 \otimes \ldots \otimes \mathcal{H}_n$ is called $k-$*separable* with respect to a specific partition into k parties iff

$$|\psi\rangle = |\psi^1\rangle \otimes |\psi^2\rangle \otimes \ldots \otimes |\psi^k\rangle. \tag{1.26}$$

For $k = 2$ the state $|\psi\rangle$ is called *biseparable*. A mixed state ρ of n parties acting on $\mathcal{H}_1 \otimes \mathcal{H}_2 \otimes \ldots \otimes \mathcal{H}_n$ is fully separable iff it can be written as

$$\rho = \sum_i p_i |\psi_i^1\rangle\langle\psi_i^1| \otimes |\psi_i^2\rangle\langle\psi_i^2| \otimes \ldots \otimes |\psi_i^n\rangle\langle\psi_i^n|, \tag{1.27}$$

where $\sum_i p_i = 1$ and $p_i \geq 0$. The state ρ is $k-$*separable* iff it can be written as

$$\rho = \sum_i p_i |\psi_i^1\rangle\langle\psi_i^1| \otimes |\psi_i^2\rangle\langle\psi_i^2| \otimes \ldots \otimes |\psi_i^k\rangle\langle\psi_i^k|. \tag{1.28}$$

A state carries true n-party entanglement iff it cannot be decomposed as a mixture of biseparable states.

The treatment of multiparty states becomes more difficult because there is no Schmidt decomposition in general [57]. However, it is still possible to eliminate some parameters of the states by using the choice of the local bases [58, 59]. Further, the possibilities of transforming states are more restricted. As mentioned above, there exist classes of pure state which cannot be transformed into each other asymptotically by LOCC only, in contrast to the bipartite pure state case.

One way of characterizing multiqubit states with respect to their separability and distillability properties was introduced by Dür, Cirac and Tarrach (DCT) [60–62]. They constructed multipartite qubit states depending on relatively few parameters, from which PPT properties of all possible cuts can be easily read off. Further, they showed that any n-qubit state can be transformed or *depolarized* to that form by local operations, preserving these parameters. So from entanglement and distillability properties of the depolarized states properties of the original states can be inferred. Here, distillability is defined as the possibility of distilling a singlet state between any two parties. This is why – contrary to the bipartite case – a state might be undistillable even though it does not have a PPT with respect to *all* bipartite splits. We will further discuss this property in chapter 3, where we construct explicit networks for the generation of three qubit states of the family introduced by DCT.

In the following section, we turn towards another way of classifying multiparty states.

1.2.1 Classification of mixed states via SLOCC

We consider here equivalence classes of states with respect to stochastic LOCC (SLOCC), i.e., LOCC without requiring that the operation succeeds with unit probability. We illustrate the idea by considering the simplest multipartite system, i.e., three qubits.

A three qubit state $|\phi\rangle$ is locally convertible to a state $|\psi\rangle$ by SLOCC iff invertible operators A, B, and C exist such that [63]

$$|\psi\rangle = A \otimes B \otimes C|\phi\rangle. \tag{1.29}$$

This leads to the following 6 equivalence classes under SLOCC

- The class of fully separable states $|\psi_A\rangle \otimes |\phi_B\rangle \otimes |\gamma_C\rangle$.
- 3 classes of biseparable states $|\psi_A\rangle \otimes |\delta_{BC}\rangle$, and in analogy for the biseparable splittings $B - AC$ and $C - AB$.
- The W-class represented by $|\mathrm{W}\rangle = \frac{1}{\sqrt{3}}(|001\rangle + |010\rangle + |100\rangle)$.
- The GHZ-class [64] represented by $|\mathrm{GHZ}\rangle = \frac{1}{\sqrt{2}}(|000\rangle + |111\rangle)$.

If non-invertible operators are allowed, then from both the W class and the GHZ class any of the bipartite classes can be reached, and from any of those the fully separable class.

As mentioned above, there does not exist a Schmidt decomposition for three qubits, but the freedom of choosing the local bases allows to bring any pure state to the form [59]

$$\lambda_0|000\rangle + \lambda_1 e^{i\theta}|100\rangle + \lambda_2|101\rangle + \lambda_3|110\rangle + \lambda_4|111\rangle, \tag{1.30}$$

where $\lambda_i \geq 0$ for $i = 0 \ldots 4$ and $\theta \in [0, \pi]$. Such a state is generically of the GHZ class, while a W vector can be written as

$$\lambda_0|000\rangle + \lambda_1|100\rangle + \lambda_2|101\rangle + \lambda_3|110\rangle. \tag{1.31}$$

It is not obvious at first sight that the W state is part of this family, but if we relabel $|0\rangle \leftrightarrow |1\rangle$ for the first party, then all the kets from the W state are present in this form. Hence even though states from the GHZ class and the W class cannot be converted into each other by SLOCC, there is a GHZ state arbitrarily close to any W state, and the set of pure W states is of measure zero among all pure states [63].

For mixed states, the following classes can be defined

- The class S of states which can be written as a convex combination of fully separable states.
- The class B of states which can be written as a convex combination of biseparable and fully separable states. The decomposition might contain states of different partitions.
- The class W which can be written as a convex combination of W states and of states of the classes B and S.
- Decompositions of states of the GHZ class contain at least one GHZ vector.

Only for the production of states belonging to the classes W and GHZ true tripartite entanglement is needed. All classes are compact, convex, and embedded into each other as $S \subset B \subset$W$\subset$GHZ. Hence we can again find hyperplanes separating states from a convex set in order to prove that they do not belong to the set. In chapter 2, we explain how this fact can be used to construct witnesses detecting true multipartite entanglement, and present results from an experiment where these witnesses were applied [VII].

For pure bipartite systems, the classification via SLOCC leads to equivalence classes of states with equal Schmidt rank, where the maximally entangled states with equal Schmidt coefficients can be taken as representatives [63]. When noninvertible local operations are allowed, it is possible to decrease the Schmidt rank of a vector, hence classes of higher Schmidt rank can be regarded as containing more powerfull entanglement. This also means that it is possible to reach any other state from a state with highest Schmidt rank in contrast to the case of three qubits, where W class states cannot be transformed to GHZ class states and vice versa.

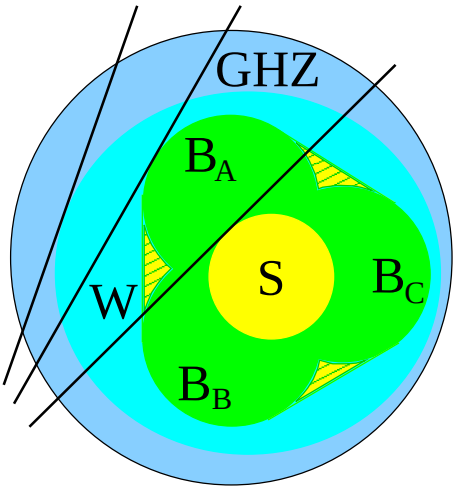

Figure 1.4. *The set of mixed three qubit states. The set S is the set of fully separable states, while there are three sets of biseparable states, labelled by B_A, B_B, and B_C, corresponding to the partitions $A-BC$, $B-AC$, and $C-AB$, respectively. The lines indicate hyperplanes or witnesses that separate GHZ class states from W class states, W class states from B class states, and B class states from S class states. To be precise, we should label the classes by GHZ$\setminus$W and so on in order to have the same notation as in the previous chapter.*

It is also possible to classify the set of mixed bipartite states in a similar way as the set of mixed three qubit states. The set of states of a given Schmidt rank r is defined as the set of all states that can be written as a convex combination of vectors with Schmidt rank smaller than or equal to r. For each r, the sets are convex and compact subsets of the sets with higher Schmidt rank [65].

There have been attempts to construct an analogous classification for systems of higher dimension or consisting of more parties. For systems of two qubits and one N level system, the number of equivalence classes is still finite [63, 66], the three qubit classes being among them. In contrast, for 4 qubits already, the number of equivalence classes under SLOCC in infinite [63, 67]. This is due to the fact that the number of parameters that can be changed locally by SLOCC grows slower than the number of parameters needed to parametrize a state if the dimension of the systems and the number of parties is increased [63]. Because of that, the relative size of the equivalence classes shrinks effectively when the size of the system is increased.

Nevertheless, for n parties, there is always a convex, compact class of biseparable states embedded into the class of truely n-party entangled states. Hence, hyperplanes or witness operators can always be used to distinguish true n-partite entanglement from biseparable states.

1.3 Bell inequalities

We mentioned already the criticism of quantum mechanics by EPR, and that Bell found a way to formulate the assumptions used in the form of inequalities. Bell's argument was based on the correlations of the singlet state. This was extended later by Clauser, Horne, Shimony, and Holt (CHSH) to an inequality for arbitrary two qubit states [68]. In the following, we will illustrate the idea by deriving the CHSH inequality in the spirit of [69].

Consider a source emitting two particles at a time to the receivers Alice and Bob. For each pair, Alice can choose to measure one of the observables $\hat{A}_1$ and $\hat{A}_2$, while Bob can choose between $\hat{B}_1$ and $\hat{B}_2$. Each experiment on either side has the possible outcomes ± 1.

The first EPR assumption underlying Bell's inequalities is that Alice and Bob have *free will* in choosing which measurement they want to perform in each run. Second, the *reality* assumption implies that to each possible measurement the outcome ± 1 can be assigned in advance. Finally, *locality* is assumed to hold, meaning that the outcome of a measurement of Alice should not depend on the choice of a spacelike separated measurement of Bob and vice versa.

By the locality and reality assumptions, we can assign independent values A_1^j, A_2^j, B_1^j, and B_2^j to the outcomes of all possible measurements on both sides for the $j-$th pair of particles. Then, the following relation holds

$$A_1^j(B_1^j + B_2^j) + A_2^j(B_1^j - B_2^j) = \pm 2, \tag{1.32}$$

because only one of the two terms can be nonvanishing. Averaging over many runs, this leads to

$$|E(A_1, B_1) + E(A_1, B_2) + E(A_2, B_1) - E(A_2, B_2)| \leq 2, \tag{1.33}$$

where $E(A_i, B_j)$ is the expectation value of the correlation experiment $\hat{A}_i\hat{B}_j$. This is the CHSH inequality [68] that gives a bound on *any* local realistic theory trying to explain the results. Each correlation term is measured with a subensemble only, because in a single run only one of the correlation terms is measured. Therefore, it has to be assumed that the measured values with respect to these subensembles does not deviate too much from the outcomes that would have been obtained if the whole ensemble would have been used.

In the derivations, we did not specify how the values are assigned to the outcomes of the measurements. This can be done by introducing so-called *hidden* variables. A model that can account for all correlations occuring in the measurements in this way is called a local hidden variable (LHV) model. The violation of a Bell inequality implies the non-existence of a LHV model for the correlations observed with respect to a certain state [70]. In the following, when we say that a state admits a LHV model, it is understood that this model is constructed with respect to a particular Bell inequality, with a fixed number of measurements.

In order to test quantum mechanics, the substitutions $\hat{A}_i \to \mathbf{a}_i \cdot \boldsymbol{\sigma} \equiv \sigma_{a_i}$ have to be made, and the correlations with respect to a state ρ are calculated as $E(\hat{A}_i, \hat{B}_j) = \mathrm{Tr}(\sigma_{a_i} \otimes \sigma_{b_j})\rho$. For the singlet state, $E(A_i, B_j) = -\mathbf{a}_i \cdot \mathbf{b}_j$. Choosing all measurement directions to lie in a plane parametrized by the angle θ with respect to the x axis, then for the choice $\theta_{a_1} = 0$, $\theta_{a_2} = \pi/2$, $\theta_{b_1} = \pi/4$, and $\theta_{b_2} = -\pi/2$, we obtain for the left hand side (lhs) in Eq. (1.33) the result $|3 \cdot (-1/\sqrt{2}) - 1/\sqrt{2}| = 2\sqrt{2}$, hence the CHSH inequality can be violated by quantum mechanical states. Note that for product states, the correlation terms factorize,

$$\mathrm{Tr}\Big[(\sigma_{a_i} \otimes \sigma_{b_j})|a,b\rangle\langle a,b|\Big] = \langle a|\sigma_{a_i}|a\rangle\langle b|\sigma_{b_j}|b\rangle. \tag{1.34}$$

As a consequence, the CHSH inequality (1.33), as well as all other Bell inequalities, is fulfilled for all separable states, may they be pure or mixed.

Several generalizations of the CHSH inequality have been derived in the following years. Inequalities for n-qubits with two dichotomic, i.e., two outcome, measurement settings per site were studied by Mermin [71], Ardehali [72], Belinskii and Klyshko [73]. The complete set of such inequalities was recently constructed [74, 75]. Further, generalizations for to more outcomes [76] and to several settings per site have been made, see, for instance, [77–80].

The first violation of Bell inequalities where all the EPR assumptions were met to a relatively high degree was observed in the experiment of Aspect *et al.* [81]. Since that time, several experiments have been performed improving the detection efficiency of the detectors [82], as well as the spacelike separation of the measurements [83].

Let us come back to the relation of Bell inequalities and quantum mechanics. We mentioned that all separable states fulfil all Bell inequalities. Remarkably enough, the natural assumption that all entangled states violate a Bell inequality is not necessarily true. For a one parameter family of $U \otimes U$ invariant states in $d \times d$ dimensions, Werner constructed a LHV model for a parameter range where the states are entangled [26].

Using the same family of states, Popescu showed [84] that it is possible to obtain violations of Bell inequalities by local operations from entangled states not violating a Bell inequality before. Gisin obtained the same effect even for two qubits [85]. A natural conjecture following from these observations is that all undistillable entangled states admit a LHV model. However, Dür showed that for a family of multipartite qubit states there exists a range of parameters where the states are undistillable while violating a Bell inequality with two dichotomic measurement settings per site [86].

This leaves open the question whether all PPT states admit a LHV model [87], because it was shown by Acín that the violation of the Bell inequalities considered by Dür always involves at least one bipartite split with a NPPT of the state [88]. This is related to the result of Werner and Wolf that no n-qubit inequality with two dichotomic measurement settings per site is violated by PPT entangled states [89].

We will come back to Bell inequalities in chapter 5, where we investigate the relation between the CHSH inequality and witness operators.

Chapter 2

Local detection of entanglement via entanglement witnesses

2.1 Overview

This chapter is devoted to providing an efficient and easy method for the experimental detection of entanglement, where detection refers to the experimental proof that the state of a system that two or several parties share is entangled. As an easy method we consider a method which requires each of the parties to perform measurements on its subsystem only, because global measurements on the whole system are in general technically demanding. The method should be efficient in the sense that the number of local measurements that are needed for the decision should be as small as possible. Further, the method should also be efficient in detecting as many states as possible, and the entanglement test should be decisive, i.e., it should not rely on any assumptions.

A method that is always applicable is the measurement of the density matrix with tomographic methods, followed by the application of one of the entanglement criteria mentioned in the introduction. The tomographic measurement is possible with local measurements only, as sketched in chapter 1. The drawback is that this method requires a large number of measurements in general. Further, even if the density matrix is known, the criteria might fail to detect the entanglement, or extensive computations might have to be performed.

There are methods which can be applied with less effort. For instance, if it can be assumed that the state in question is a pure state of two qubits, a method for measuring the amount of entanglement with local measurements only was introduced in Ref. [90]. Recently, another method for entanglement detection based on properties of the states which are invariant under local unitary operations has been introduced in Ref. [91]. For the entanglement test, the same local measurements as those needed for state tomography can be performed, but in this case the entanglement might be detected before the complete tomography has been done. So this criterion is rather easily implemented, but it is in general weaker than the PPT criterion [91].

Yet another method was developed in Ref. [92] which is based on the PPT criterion. Here, an approximation to the partial transpose operation is implemented on the

state in question, followed by an estimation of the eigenvalues of the resulting state. From the eigenvalues it is possible to tell whether the original state was PPT or not. The eigenvalues can be estimated by multiple applications of an interferometric network, acting on different numbers of copies in each run, which can be performed by LOCC [93]. Each party has to act on its part of the copies of the state then. This network can be modified such that it estimates the eigenvalues directly [94], without the need of the approximation of the partial transpose operation. With this interferometric network less parameters have to be estimated than with tomography to decide whether a state is PPT, namely the eigenvalues of the partially transposed state only, while the potential difficulties lie in the realization of the network.

The most frequently used method for proving the entanglement of a state up to now is the violation of a Bell inequality. Even though they are constructed to formulate the constraints every local and realistic theory has to obey, their violation is also a signature for entanglement, because every separable state can be described by a LHV model, as we saw in section 1.3. Further, they only require local measurements by construction, so that they can be readily implemented in a laboratory.

In a quantum mechanical world, Bell inequalities correspond to entanglement witnesses [40, 95]. This relation will be investigated in detail for the case of two qubits in chapter 5. Most likely, Bell inequalities correspond to non-optimal witnesses in general, because there exist entangled states with a LHV model even for systems of two qubits [26]. Further, no existing Bell inequality is violated by PPT entangled states [70]. This is a drawback of this method as far as the detection of entangled states is concerned. Another problem in this regard is that up to now it is not possible to construct a Bell inequality for a given state.

In conclusion, even though the presented methods can all be applied with local means, they either require many local measurements or cannot guarantee that any state is detected in general. However, in chapter 1 we already introduced operators capable of detecting *any* entangled state: the witness operators. If some knowledge about the state to be detected can be assumed, e.g., when an experiment is aimed at producing a particular state, then witnesses are well suited for delivering the entanglement proof. Even further, they are also capable of proving that a state is multipartite entangled.

So entanglement witnesses are well suited for detecting entanglement. However, it is essential to find ways to measure them locally, with as little effort as possible. This, as well as the construction of witnesses, is the aim of this chapter, with the focus on witnesses for finite-dimensional bipartite NPPT states and PPT entangled states. Notice that if we would not aim at reducing the experimental effort, then one could just measure the state ρ with a complete local state tomography and calculate the expectation value $\mathrm{Tr}[W\rho]$ directly.

The scheme of local detection via entanglement witnesses can be formulate explicitly as follows: Given a state ρ, we construct an entanglement witness W such that $\mathrm{Tr}[W\rho] < 0$. It is clear from section 1.1.3 that the witness should be as fine as possible, because finer witnesses detect a larger volume of the set of entangled states, so that the chances are better that a state is detected even in the presence of noise

in the laboratory. Then we decompose this operator into a sum of terms which can be measured locally. We further try to optimize, i.e., minimize, the number of local measurements that have to be performed in order to reduce the experimental effort. We distinguish two optimization strategies: i) optimization of the number of projectors onto product vectors that have to be measured, and ii) optimization of the number of locally correlated measurement settings where each party has to perform a von Neumann measurement on its system.

The chapter is organized as follows: In the first two sections 2.2 and 2.3 we discuss methods to construct witnesses, both for NPPT as well as for PPT entangled states, and introduce the different optimization strategies for local decompositions of witnesses explicitly.

In the following sections, we apply these methods to several examples, starting in section 2.4 with entanglement witnesses for bipartite NPPT entangled states of $N \times M$ systems. In order to get started, we first discuss a simple example in a two qubit system there.

Then we focus our attention on PPT entangled states in section 2.5. First, we find witnesses and local decompositions for three examples of bipartite PPT entangled states from the literature: UPB states [96] and chessboard states [97] in 3×3 systems [96], and Horodecki's states in a 2×4 system [29].

We also introduce two families of multiqubit PPT entangled states. The first is an extension of a family of three-qubit PPTES based on the GHZ state introduced in [98] to n qubits, while the second is a family of three-qubit PPTES based on the W state. For the first family, we construct witnesses and local decompositions for their local detection.

In section 2.6, we tell briefly about how witnesses were used to prove the multipartite entanglement of the W state in the group of H. Weinfurther [VII,VIII].

Finally, the conclusions and open questions can be found in section 2.7.

2.2 Constructing entanglement witnesses

Here, we introduce three methods of constructing entanglement witnesses: the first one is based on the NPPT property, while the second can distinguish entangled from biseparable states, which is useful for proving that a state is truely multiparty entangled. We show that witnesses constructed with the second method cannot detect PPT entangled states either. This is exactly what witnesses produced with the third method are capable of.

2.2.1 Witnesses for NPPT states

If ρ is NPPT, then there exist an entangled vector $|\phi\rangle$ and $\lambda < 0$ such that $\rho^{T_A}|\phi\rangle = \lambda|\phi\rangle$. Then

$$W = |\phi\rangle\langle\phi|^{T_A} \tag{2.1}$$

is a witness detecting ρ.

From section 1.1.3 we know that in order to be an entanglement witness, W has to be positive on all separable states, while having a negative expectation value for at least one entangled state. The witness from Eq. (2.1) is positive on all separable states because $\mathrm{Tr}[|a,b\rangle\langle a,b||\phi\rangle\langle\phi|^{T_A}] = \mathrm{Tr}[|a,b\rangle\langle a,b|^{T_A}|\phi\rangle\langle\phi|] = \mathrm{Tr}[|a^*,b\rangle\langle a^*,b||\phi\rangle\langle\phi|] \geq 0$. The state ρ is detected by W because $\mathrm{Tr}[\rho|\phi\rangle\langle\phi|^{T_A}] = \mathrm{Tr}[\rho^{T_A}|\phi\rangle\langle\phi|] = \lambda < 0$. This witness belongs to the class of decomposable witness, to be defined in

Definition 2.1. *Decomposable* witnesses can be written as $W = P + Q^{T_A}$, where P and Q are both positive semi-definite operators.

Decomposable witnesses cannot detect PPT entangled states because

$$\mathrm{Tr}[(P + Q^{T_A})\rho] = \mathrm{Tr}[P\rho] + \mathrm{Tr}[Q\rho^{T_A}] \geq 0, \tag{2.2}$$

since P, Q, and ρ^{T_A} are positive semi-definite operators. In the following, we will sometimes refer to positive semi-definite operators just as positive operators. In systems of one qubit and a qubit or a qutrit, all the entanglement witnesses are decomposable, because there the PPT criterion is necessary and sufficient.

Theorem 2.2 [43]. An *optimal* decomposable witness can be written as $W = Q^{T_A}$, where $Q \geq 0$ contains no product vector in its range.

Hence the construction of witnesses from Eq. (2.1) produces optimal decomposable witness operators. For 2×2 systems, there is always a product vector in a plane spanned by two entangled vectors [99]. Hence Q has to be of rank one in this case.

2.2.2 Witnesses excluding biseparability

A witness operator that detects the entanglement of a pure bipartite state $|\psi\rangle$ is given by

$$W = \alpha\mathbb{1} - |\psi\rangle\langle\psi| \ , \tag{2.3}$$

where $\mathbb{1}$ is the identity operator,

$$\alpha = \max_{|\phi\rangle\in S} |\langle\phi|\psi\rangle|^2 \ , \tag{2.4}$$

and S denotes the set of separable states. This construction guarantees that $\mathrm{Tr}[W\sigma_s] \geq 0$ for all separable states σ_s, and that $\mathrm{Tr}[W|\psi\rangle\langle\psi|] < 0$.

The overlap α can be calculated as follows. We choose an orthonormal product basis $|ij\rangle$ and expand $|\psi\rangle = \sum_{ij} c_{ij}|ij\rangle$ and $|\phi\rangle = |a\rangle|b\rangle = \sum_{ij} a_i b_j |ij\rangle$. The coefficient matrix is denoted by $C = (c_{ij})$ and the normalized coefficient vectors by $\vec{a} = (a_i)$ and $\vec{b} = (b_i)$. Then

$$\begin{aligned} \max_{|\phi\rangle\in B_1} |\langle\phi|\psi\rangle| &= \max_{a_i,b_j} |\sum_{ij} ({a_i}^* c_{ij} {b_j}^*)| \\ &= \max_{\vec{a},\vec{b}} |\langle\vec{a}|C|\vec{b}^*\rangle| = \max_k \{\lambda_k(C)\} \ , \end{aligned} \tag{2.5}$$

where $\lambda_k(C)$ denotes the singular values of C, *i.e.* the roots of the eigenvalues of $CC^\dagger$. In other words, $\lambda_k(C)$ are the Schmidt coefficients of $|\psi\rangle$, and α is therefore simply the square of the largest Schmidt coefficient of $|\psi\rangle$.

In the following, we will show that all witnesses of the form (2.3) are decomposable, i.e., they cannot detect PPT entangled states. For doing this, we need the following

Lemma 2.3 [43]. A witness B is finer than a witness A iff there exist a positive x and a positive operator P such that $A = xB + P$.

Note that in particular $A - xB \geq 0$, which we will use in the following Lemma.

Lemma 2.4. Let $|\phi_N\rangle = \sum_{i=0}^{N-1} a_i|ii\rangle$ be a normalized state of a $N \times N$ system written in the Schmidt decomposition with $a_0 \geq a_1 \geq \ldots a_{N-1}$. Then $W_{\phi_N} = a_0^2\mathbb{1} - |\phi_N\rangle\langle\phi_N|$ is a decomposable witness operator.
Proof. In [100] it was shown that the witnesses

$$W_{\phi_N^+} = \frac{1}{N}\mathbb{1} - |\phi_N^+\rangle\langle\phi_N^+| \tag{2.6}$$

are decomposable, where $|\phi_N^+\rangle = \sum_{i=0}^{N-1} |ii\rangle/\sqrt{N}$ is the maximally entangled state in $N \times N$ systems. The reason is that $|\phi_N^+\rangle\langle\phi_N^+| = \frac{1}{N}(\mathbb{1} - 2P_a^{T_A})$ holds, where P_a is the projector onto the antisymmetric subspace. Hence $W = 2P_a^{T_A}$.

The witness W_{ϕ_N} is of the form of Eq. (2.3). We will show that $W_{\phi_N^+}$ is finer than W_{ϕ_N}. From this it follows that W_{ϕ_N} is decomposable because $W_{\phi_N^+}$ is a decomposable witness. By virtue of Lemma 2.4, we can do this by showing that there exists a positive x such that $W_{\phi_N} - xW_{\phi_N^+} \geq 0$. The lhs is given by

$$W_{\phi_N} - xW_{\phi_N^+} = (a_0^2 - \frac{x}{N})\mathbb{1} + x|\phi_N^+\rangle\langle\phi_N^+| - |\phi_N\rangle\langle\phi_N| \equiv \rho_x - |\phi_N\rangle\langle\phi_N|. \tag{2.7}$$

In [101] it was shown that $\rho_x - \lambda|\phi_N\rangle\langle\phi_N| \geq 0$ for $\rho_x \geq 0$ if $\lambda \leq \langle\phi_N|\rho^{-1}|\phi_N\rangle^{-1}$, where the inverse of ρ_x is taken on its range. The first condition is $\rho_x \geq 0 \Leftrightarrow x \leq Na_0^2$. Then we have to show that there exists a x such that $\lambda_{\mathbf{max}} = \langle\phi_x|\rho^{-1}|\phi_x\rangle^{-1} \geq 1$. Let us see if $x = Na_0^2$ is a solution. In this case $\rho_x^{-1} = \frac{1}{Na_0^2}|\phi_N^+\rangle\langle\phi_N^+|$ and we obtain

$$\lambda_{\mathbf{max}} = \frac{Na_0^2}{|\langle\phi_N|\phi_N^+\rangle|^2} = \frac{N^2a_0^2}{(\sum_{i=0}^{N-1} a_i)^2}. \tag{2.8}$$

Using Lagrangian multipliers it can be shown that $a_0^2/(\sum_{i=0}^{N-1} a_i)^2 \in [1, 1/N^2]$, hence $\lambda_{\max}(x = Na_0^2) \geq 1$. □

Note that this result also holds for $N \times M$ dimensional systems, where $N \leq M$, because the maximal Schmidt rank of a state is also N in this case.

So far, we only used the construction of Eq. (2.3) for detecting bipartite states. However, the construction can be easily extended for multipartite systems. We just have to calculate the coefficient α for every bipartite splitting by choosing a basis containing only product vectors with respect to this partition, and take the maximal value from all partitions as coefficient for the witness. Then, a negative

expectation value of the observable W clearly signifies that the state $|\psi\rangle$ carries true multipartite entanglement, because the witness is positive on all biseparable states by construction. We can also extend Lemma 2.4 to the multipartite case.

Theorem 2.5. Every multipartite witness operator of the form $W = \alpha\mathbb{1} - |\phi\rangle\langle\phi|$, where α is the square of the largest Schmidt coefficient over all bipartite cuts of the multipartite state $|\phi\rangle$, is a decomposable witness operator.
Proof. For every bipartite cut, W is of the form of $W_{\phi_N} = a_0^2\mathbb{1} - |\phi_N\rangle\langle\phi_N|$, with a coeffiecient $a_0^2 \leq \alpha$, and decomposable due to Lemma 2.4. Hence it is decomposable with respect to any bipartite splitting and cannot detect PPT entangled states. □

2.2.3 Witnesses for PPT entangled states

In this subsection we present witnesses that are capable of detecting PPT entangled states. In particular, we present here a method for construction witnesses for the so-called *edge* states [43].

A state δ is called an edge state iff it cannot be represented as $\delta = q\delta' + (1 - q)\sigma_s$, where $0 \leq q < 1$, σ_s is a separable state and δ' is a state with a positive partial transpose. In other words, for all product vectors $|e, f\rangle$ and $\epsilon > 0$, $\delta - \epsilon|e, f\rangle\langle e, f|$ is not a state anymore. This implies that the edge states lie on the boundary between the bound entangled states and the entangled states with non-positive partial transpose. They violate the range criterion introduced in section 1.1.2 in an extremal sense, i.e., δ is an entangled edge state with a positive partial transpose iff for all product vectors $|e, f\rangle \in R(\delta)$, $|e^*, f\rangle \notin R(\delta^{T_A})$, where $R(\delta)$ denotes the range of δ.

The generic form of an entanglement witness for such a state δ is [43]

$$W = \bar{W} - \epsilon\mathbb{1}, \tag{2.9}$$

where

$$\bar{W} = (P + Q^{T_A}) \tag{2.10}$$

$$\epsilon = \inf_{|e,f\rangle} \langle e, f|\bar{W}|e, f\rangle, \tag{2.11}$$

and P and Q denote the projectors onto the kernel of δ and δ^{T_A}, respectively, and we sometimes call $\bar{W}$ the *prewitness*. Because of the edge state properties, the coefficient ϵ is positive:

$$\langle e, f|P + Q^{T_A}|e, f\rangle = \langle e, f|P|e, f\rangle + \langle e^*, f|Q|e^*, f\rangle > 0. \tag{2.12}$$

This construction can be easily generalized to more than two parties, when more partitions play a role, as we will see in section 2.5.4.

It is also possible to substract an operator I which is positive on the range of δ [43]. The coefficient ϵ has to be adopted to

$$\epsilon' = \inf_{|e,f\rangle} \frac{\langle e, f|\bar{W}|e, f\rangle}{\langle e, f|I|e, f\rangle}. \tag{2.13}$$

However, from a practical point of view, it is very convenient to substract the identity, because this term does not require a measurement at all since $\mathrm{Tr}\delta = 1$, so that the prewitness $\bar{W}$ only has to be decomposed.

The parameter ϵ can be determined by the use of multiparameter minimization routines. However, these cannot in general guarantee that a global minimum is reached. This is a severe problem because W from Eq. (2.9) is not a witness if ϵ larger than the prescribed value of Eq. (2.11). In chapter 4, we will see that it is possible to obtain lower bounds for $\inf_{|e,f\rangle}\langle e, f|\bar{W}|e, f\rangle$, which makes it possible to ensure that W is a proper witness.

2.3 Local decompositions of entanglement witnesses

Having constructed an entanglement witness, it is necessary to find a decomposition into operators which can be measured locally. Such a decomposition is of the general form

$$W = \sum_{i=1}^{k} c_i |e_i\rangle\langle e_i| \otimes |f_i\rangle\langle f_i|. \tag{2.14}$$

Such a decomposition can be measured locally: If a source distributes bipartite states ρ to Alice and Bob, they have to measure the expectation value of the projectors $|e_i\rangle\langle e_i| \otimes |f_i\rangle\langle f_i|$ with respect to this state and add their results with the weights c_i. One can construct such a decomposition in many ways, but it is reasonable to do it in a way which corresponds to expenses of Alice and Bob which are as small as possible. There are several possibilities to define an optimal decomposition:

One possibility is to look for the optimal number of product vectors (ONP), i.e., one can try to minimize k in (2.14). This optimization strategy looks very natural and has already been considered in the literature. It was proven in [101] that in general the ONP k_- for any two-qubit state is equal to 5, while separable two-qubit state need no more than 4 product vectors. Also a constructive way for computing this optimal decomposition was given.

What is the "cost" Alice and Bob have to pay when measuring W via such a decomposition? It is the number of measurements they have to perform. When we talk about measurements here, we consider only von Neumann measurements. We do not consider more general positive operator valued measurements (POVMs) here, because their implementation would require additional ancilla systems. One measurement on Alice's side in the sense above consists of a choice of one orthonormal basis for Alice's Hilbert space. For an particle with spin s one may interpret this as the choice of a direction for a Stern-Gerlach-like apparatus. Alice sets up her device in the desired direction and is able to distinguish between $2s + 1$ different states.

For one local measurement Bob also has to choose an orthonormal basis in his Hilbert space; all together this yields one orthonormal product basis for both. Thus, if we

are in a $N \times N$ system a term of the form

$$\sum_{k,l=1}^{N} c_{kl} |A_k\rangle\langle A_k| \otimes |B_l\rangle\langle B_l| \tag{2.15}$$

with $\langle A_s|A_t\rangle = \langle B_s|B_t\rangle = \delta_{st}$ can be measured with one collective setting of measurement devices of Alice and Bob. Alice and Bob can discriminate between the states $|A_k, B_l\rangle$, measure the probabilities of these states and add their results with the weights c_{kl} using one collective setting and some classical communication. We call such a collective setting of measurement devices a local von Neumann measurement (LvNM).

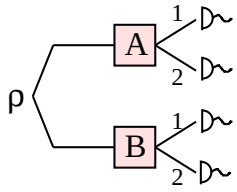

Figure 2.1. *The state ρ is distributed to Alice and Bob. They perform local measurements A and B, respectively, from which they can obtain the probabilities* $\mathrm{Tr}[|A_k, B_l\rangle\langle A_k, B_l|\rho]$ *for all the terms occuring in Eq. (2.15), here for two-valued measurements on both sides.*

It is therefore reasonable to find a decomposition of the form

$$W = \sum_{i=1}^{m} \sum_{k,l=1}^{N} c_{kl}^{i} |A_k^i\rangle\langle A_k^i| \otimes |B_l^i\rangle\langle B_l^i| \tag{2.16}$$

with $\langle A_s^i|A_t^i\rangle = \langle B_s^i|B_t^i\rangle = \delta_{st}$ and an optimal number of devices' settings (ONS), i.e., a minimal m. In this sense m is the minimal number of measurements Alice and Bob have to perform.

Please note that a decomposition like (2.14) with the minimal k_- (ONP) will probably require k_- LvNMs because the vectors $|e_i, f_i\rangle$ are not orthogonal to each other in general.

We also would like to emphasize that a decomposition of the form (2.16) is more general than a decomposition into a sum of tensor products of operators:

$$W = \sum_{i=1}^{m} \gamma_i A_i \otimes B_i. \tag{2.17}$$

The decomposition (2.17) has the advantage that Alice and Bob do not have to distinguish between some states, they only have to measure locally some expectation values of Hermitean operators. A decomposition like in Eq. (2.16) can be written in the form of (2.17) if for all i the matrices (c_{kl}^i) are of rank one. In the following

we will see that for qubit systems there is not a big difference between (2.16) and (2.17). From the optimal decomposition in the sense of (2.16) we can derive a decomposition of the form (2.17) where some of the operators are the identity ($\mathbb{1}$), so they do not require new measurement settings. For $N \times N$ systems we will see that it is straightforward to derive the optimal decomposition in the sense of (2.17).

2.4 Local detection of bipartite NPPT entanglement

In this section, we construct for the first time explicit examples of witness and local decompositions. The constructions are based on section 2.2.1. We start looking at witnesses for two-qubit systems. The results are then used to obtain more general result for such witnesses in $N \times M$ systems.

2.4.1 Witnesses for two-qubit systems

Let us assume that the state $|\psi\rangle = a|01\rangle + b|10\rangle$ written in the Schmidt decomposition is the output state of a device in a laboratory. In order to being able to proof its entanglement experimentally, we construct a witnesses, and find different local decompositions: those with optimal number of projectors (ONP) and with optimal number of settings (ONS).

The projector onto $|\psi\rangle$ partially transposed has the negative eigenvalue $-ab$, and the corresponding eigenvector is $|\phi^-\rangle = (|00\rangle - |11\rangle)/\sqrt{2}$. Hence the witness is given by

$$W = |\phi^-\rangle\langle\phi^-|^{T_A} = \frac{1}{2}\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \tag{2.18}$$

independent of a. Further, the witness is robust to noise. In the case where the source emits the state mixed with white noise, so that it has to be described by the mixture

$$\rho_p = p|\psi\rangle\langle\psi| + \frac{1-p}{4}\mathbb{1}, \tag{2.19}$$

then the state is still entangled and detected by the witness for $p > (1+4ab)^{-1}$. This is not surprising, because W is an optimal witness, as explained in section 2.2.1. The case where noise is admixed which is different from white noise, but not too much, has been discussed in Ref. [V]. This witness has been implemented experimentally, the results being described in Ref. [102].

For general states, there may be other states corresponding to the negative eigenvalue of the partially transposed state. It is therefore natural to decompose a more general state written in the Schmidt decomposition as $|\psi\rangle = \alpha|00\rangle + \beta|11\rangle$. For our W we have the special case $\alpha = 1/\sqrt{2} = -\beta$, but we want to deal with the most general $|\psi\rangle$.

First, we compute the ONP-decomposition with the minimal k_- according to [99], arriving at

$$|\psi\rangle\langle\psi|^{T_A} = \frac{(\alpha+\beta)^2}{3}\sum_{i=1}^{3}|A_i'B_i'\rangle\langle A_i'B_i'| - \alpha\beta(|01\rangle\langle 01| + |10\rangle\langle 10|), \tag{2.20}$$

where we have used the definitions

$$\begin{aligned} |A_1'\rangle &= e^{i\frac{\pi}{3}}\cos(\theta)|0\rangle + e^{-i\frac{\pi}{3}}\sin(\theta)|1\rangle = |B_1'\rangle \\ |A_2'\rangle &= e^{-i\frac{\pi}{3}}\cos(\theta)|0\rangle + e^{i\frac{\pi}{3}}\sin(\theta)|1\rangle = |B_2'\rangle \\ |A_3'\rangle &= \cos(\theta)|0\rangle + \sin(\theta)|1\rangle = |B_3'\rangle \\ \cos(\theta) &= \sqrt{\alpha/(\alpha+\beta)} \\ \sin(\theta) &= \sqrt{\beta/(\alpha+\beta)}. \end{aligned} \tag{2.21}$$

This decomposition into five product vectors requires four correlated measurement settings for Alice and Bob.

But we can measure W with less settings. We define the spin directions by $|z^+\rangle = |0\rangle, |z^-\rangle = |1\rangle, |x^\pm\rangle = \frac{1}{\sqrt{2}}(|0\rangle \pm |1\rangle), |y^\pm\rangle = \frac{1}{\sqrt{2}}(|0\rangle \pm i|1\rangle)$. The projector to be decomposed is given by

$$|\psi\rangle\langle\psi| = \alpha^2|00\rangle\langle 00| + \beta^2|11\rangle\langle 11| + \alpha\beta\big(|00\rangle\langle 11| + |11\rangle\langle 00|\big). \tag{2.22}$$

Only the last term is not in the form of a local decomposition yet. We can use the first of the following simple algebraic identities

$$ab + cd = \frac{1}{2}\Big((a+c)(b+d) + (a-c)(b-d)\Big) \tag{2.23}$$

$$ab - cd = \frac{1}{2}\Big((a-c)(b+d) + (a+c)(b-d)\Big), \tag{2.24}$$

the *golden equations*, to decompose

$$\begin{aligned} |0\rangle\langle 1| \otimes |0\rangle\langle 1| + |1\rangle\langle 0| \otimes |1\rangle\langle 0| &= \frac{1}{2}\big(\sigma_x \otimes \sigma_x - \sigma_y \otimes \sigma_y\big) \\ &= |x+, x+\rangle\langle x+, x+| + |x-, x-\rangle\langle x-, x-| \\ &\quad -|y+, y+\rangle\langle y+, y+| - |y-, y-\rangle\langle y-, y-|, \end{aligned} \tag{2.25}$$

where we used the completeness relation twice to obtain the second equality. We arrive at the decomposition

$$\begin{aligned} |\psi\rangle\langle\psi|^{T_A} &= \alpha^2|z^+z^+\rangle\langle z^+z^+| + \beta^2|z^-z^-\rangle\langle z^-z^-| + \alpha\beta\Big(|x^+x^+\rangle\langle x^+x^+| + \\ &\quad +|x^-x^-\rangle\langle x^-x^-| - |y^+y^-\rangle\langle y^+y^-| - |y^-y^+\rangle\langle y^-y^+|\Big) \\ &= \frac{1}{4}\Big(\mathbb{1}\otimes\mathbb{1} + \sigma_z\otimes\sigma_z + (\alpha^2-\beta^2)(\sigma_z\otimes\mathbb{1} + \mathbb{1}\otimes\sigma_z) \\ &\quad +2\alpha\beta(\sigma_x\otimes\sigma_x + \sigma_y\otimes\sigma_y)\Big). \end{aligned} \tag{2.26}$$

This decomposition into six product vectors requires only the measurement of three settings: Alice and Bob only have to set up their Stern-Gerlach devices in the x-, y- and z-direction to measure $|\psi\rangle\langle\psi|^{T_A}$.

Now we want to prove that three LvNM are really necessary, which means that the decomposition of Eq. (2.26) has an optimal number of settings (ONS). Our proof is a special case of a theorem about $N \times N$ systems we will show later. But in the two-qubit case the proof is particularly simple, and therefore we present it here separately.

Proposition 2.6. In a two-qubit system a decomposition of $|\psi\rangle\langle\psi|^{T_A}$ of the form (2.16) requires at least three measurements.
Proof. Consider a decomposition requiring two measurements:

$$|\psi\rangle\langle\psi|^{T_A} = \sum_{i,j=1}^{2} c_{ij}^1 |A_i^1\rangle\langle A_i^1| \otimes |B_j^1\rangle\langle B_j^1| + \sum_{i,j=1}^{2} c_{ij}^2 |A_i^2\rangle\langle A_i^2| \otimes |B_j^2\rangle\langle B_j^2|. \tag{2.27}$$

With the help of a Schmidt decomposition as above we can write $|\psi\rangle\langle\psi|^{T_A} = \sum_{i,j=0}^{3} \lambda_{ij}\, \sigma_i \otimes \sigma_j$ with the matrix

$$(\lambda_{ij}) = \begin{pmatrix} \frac{1}{4} & 0 & 0 & \frac{\alpha^2-\beta^2}{4} \\ 0 & \frac{\alpha\beta}{2} & 0 & 0 \\ 0 & 0 & \frac{\alpha\beta}{2} & 0 \\ \frac{\alpha^2-\beta^2}{4} & 0 & 0 & \frac{1}{4} \end{pmatrix}. \tag{2.28}$$

Note that the 4x3 submatrix containing the right 3 columns is of rank 3.

Now we write any projector on the rhs of (2.27) as a vector in the Bloch sphere: $|A_1^1\rangle\langle A_1^1| = \sum_{i=0}^{3} s_i^A \sigma_i$ is represented by the vector $\vec{s}_{A_1^1} = (1/2, s_1^A, s_2^A, s_3^A)$ and $|A_2^1\rangle\langle A_2^1|$ by $\vec{s}_{A_2^1} = (1/2, -s_1^A, -s_2^A, -s_3^A)$; $|B_1^1\rangle\langle B_1^1|$ can be written similarly. If we expand the first sum on the rhs of (2.27) in the $(\sigma_i \otimes \sigma_j)$ basis, the 4x3 submatrix containing the right 3 columns is given by $(c_{11}^1 \vec{s}_{A_1^1} - c_{12}^1 \vec{s}_{A_1^1} + c_{21}^1 \vec{s}_{A_2^1} - c_{22}^1 \vec{s}_{A_2^1})^T (s_1^B, s_2^B, s_3^B)$. Hence this submatrix is of the form $a_i b_j$, which implies that it is of rank one. The corresponding submatrix from the second sum on the rhs of (2.27) is also of rank one and we arrive at a contradiction: No matrix of rank 3 can be written as a sum of two matrices of rank one. □

2.4.2 Witnesses for $N \times M$ systems

Now we want to generalize our results to higher dimensions. First we consider $N \times N$ systems, and at the end of this section we will see that all the results obtained remain valid also for $N \times M$ systems.

Again, the witness can be constructed from the projector onto a negative eigenvector of the partially transposed density operator, because we are only looking at NPPT states in this section. In the following, we decompose the projector, the partial transposition is easily performed later.

Our discussion proceeds as follows: After explaining our notation we construct a decomposition of a projector onto a state with Schmidt rank l using about $2l$ measurements. This decomposition is a generalization of the decomposition for the two-qubit case. It is not clear whether this decomposition is optimal. Then we derive a lower bound for the number of measurements needed if the Schmidt rank l is maximal. We show that if $l = N$ at least $l + 1$ measurements are necessary.

We first explain some notational and technical details, partly taken from Ref. [103]. As in the introduction, we denote the real vector space of all Hermitean operators on $\mathcal{H}_A$ by $\mathcal{HS}_A$. In this space one can use the orthogonal basis $\{\mathbb{1}, G_i^A, i = 1...N^2 - 1\}$, where the G_i^A are the traceless generators of the SU(N), normalized to $\mathrm{Tr}(G_i^2) = 1$. For $N = 2$ they are the Pauli matrices, for $N = 3$ the Gell-Mann matrices, and so on. We can define $G_0^A := \mathbb{1}$ and expand every projector (and any other element of $\mathcal{HS}_A$) in this basis:

$$|\phi\rangle\langle\phi| = \sum_{i=0}^{N^2-1} f_i G_i^A, \tag{2.29}$$

where the entries of the Bloch vector f_i are real, $f_0 = 1/N$, and from the fact that $|\phi\rangle\langle\phi|$ is a pure state it follows that

$$\sum_{i=1}^{N^2-1} f_i^2 = 1 - \frac{1}{N}. \tag{2.30}$$

We sometimes write $(f_0, ..., f_{N^2-1}) = (f_0, \vec{f}) = \hat{f}$. It is easy to see that an operator described by $\hat{g}$ is a projector onto a vector orthogonal to $|\phi\rangle$ if and only if $\vec{g}$ fulfills (2.30) and

$$< \vec{f}, \vec{g} >:= \sum_{i=1}^{N^2-1} f_i g_i = -\frac{1}{N}. \tag{2.31}$$

One can also expand any projector (as every operator) on $\mathcal{H}_A \otimes \mathcal{H}_B$ as

$$|\psi\rangle\langle\psi| = \sum_{i,j=0}^{N^2-1} \lambda_{ij} G_i^A \otimes G_j^B \tag{2.32}$$

since the $G_i^A \otimes G_j^B$ form a product basis of the space $\mathcal{HS} = \mathcal{HS}_A \otimes \mathcal{HS}_B$.

Before we show our decomposition please note that is easy to decompose any operator $A \in \mathcal{HS}$ into N^2 local measurements. One can always write

$$A = \sum_{i,j=0}^{N^2-1} \mu_{ij} G_i^A \otimes G_j^B = \sum_{i=0}^{N^2-1} G_i^A \otimes \left(\sum_{j=0}^{N^2-1} \mu_{ij} G_j^B \right) \tag{2.33}$$

to obtain such a decomposition. This is also a decomposition of the form (2.17) from section 2.3. In principle, one could also do a singular value decomposition of the coefficient matrix (λ_{ij}) to arrive at a Schmidt decomposition for operators.

Theorem 2.7. Let $|\psi\rangle\langle\psi|$ be a projector onto a state with Schmidt rank l. If l is even, $|\psi\rangle\langle\psi|$ can be decomposed into $2l-1$ local measurements. If l is odd, $|\psi\rangle\langle\psi|$ can be decomposed into $2l$ local measurements.
Proof. If we have $|\psi\rangle = \sum_{i=1}^{l} s_i|ii\rangle$ we can write

$$|\psi\rangle\langle\psi| = \sum_{i=1}^{l} s_i^2 |ii\rangle\langle ii| + \sum_{i,j=1, i<j}^{l} s_i s_j K(i,j) \tag{2.34}$$

with $K(i,j) = |ii\rangle\langle jj| + |jj\rangle\langle ii|$. The first sum corresponds to one measurement, and each of the $l(l-1)/2$ terms of the second sum can be decomposed by defining for every $K(i,j)$ the states $|X_{i,j}^{\pm}\rangle = \frac{1}{\sqrt{2}}(|i\rangle \pm |j\rangle), |Y_{i,j}^{\pm}\rangle = \frac{1}{\sqrt{2}}(|i\rangle \pm i|j\rangle)$

$$\begin{aligned} K(i,j) &= |X_{i,j}^{+}X_{i,j}^{+}\rangle\langle X_{i,j}^{+}X_{i,j}^{+}| + |X_{i,j}^{-}X_{i,j}^{-}\rangle\langle X_{i,j}^{-}X_{i,j}^{-}| \\ &\quad -|Y_{i,j}^{+}Y_{i,j}^{+}\rangle\langle Y_{i,j}^{+}Y_{i,j}^{+}| - |Y_{i,j}^{-}Y_{i,j}^{-}\rangle\langle Y_{i,j}^{-}Y_{i,j}^{-}|, \end{aligned} \tag{2.35}$$

as we have done before for 2×2 systems. This corresponds to 2 measurements for each $K(i,j)$.

In order to reduce the number of necessary measurements we sum up the terms from (2.35) for different $K(i,j)$ and $K(m,n)$ in a way that the terms from different $K(i,j)$ and $K(m,n)$ can be measured with one measurement.

Let us first consider the case that l is even. We have $l(l-1)/2$ index pairs (i,j). These pairs can be grouped into $l-1$ sets of $l/2$ pairs in a way that in every set every index $1 \leq i \leq N$ appears exactly in one pair. For $l=4$, for instance, the 3 sets may be defined as $\{(1,2),(3,4)\}, \{(1,3),(2,4)\}, \{(1,4),(2,3)\}$. If we look at the $l/2$ $K(i,j)$ belonging to one set, the l vectors $|X_{i,j}^{\pm}\rangle$ are mutually orthogonal, they form an orthogonal basis of $\mathcal{H}_A$, as well as of $\mathcal{H}_B$. The vectors $|X_{i,j}^{\pm}\rangle|X_{i,j}^{\pm}\rangle$ occuring in the $K(i,j)$ of one set can then be viewed as eigenvectors of some Hermitean operator on $\mathcal{H}_A \otimes \mathcal{H}_B$ which can be measured with one locally correlated measurement. The vectors $|Y_{i,j}^{\pm}\rangle|Y_{i,j}^{\pm}\rangle$ can also be measured with one measurement. So we need 2 measurements for one set of $l/2$ $K(i,j)$ and $2(l-1)$ measurements for all $K(i,j)$. Finally, we need one measurement for the first sum on the rhs of (2.34) and this completes the proof for even l.

If l is odd, we can similarly group the $l(l-1)/2$ index pairs into l sets of $(l-1)/2$ pairs. Again, every index appears at most one time in every set, but now one index is missing in every set, and every index is missing in exactly one set. As before, we need 2 measurements for one set and therefore $2l$ measurements for all $K(i,j)$. For the first sum on the rhs of (2.34) we do not need another measurement since we can put the vector $|i,i\rangle$ to the set of index pairs where i is missing. □

As mentioned in the beginning, it is not clear whether the number of measurements needed with our decomposition is the optimal one. In the remaining part of this section, we want to give a lower bound for the number of required measurements. The derivation is based on the same idea as the proof of Proposition 2.6 and needs three lemmata. In Lemma 2.8 we give a lower bound for the rank of some matrix of the form (2.32) in $N \times N$-systems. In the Lemmata 2.11 and 2.12 we show that

the matrix coming from one measurement has a low rank. Together this proves our bound.

Lemma 2.8. If $|\psi\rangle \in \mathcal{H}_A \otimes \mathcal{H}_B$ has the full Schmidt rank N then the matrix (λ_{ij}) in (2.32) has the full rank N^2.
Proof. First, notice that the rank of (λ_{ij}) is independent of the choice of the basis $G_i^A \otimes G_j^B$. This is in analogy to the fixed Schmidt rank of a state vector and follows from the singular value decomposition of the matrix of coefficients (λ_{ij}) [24].

Now we simply construct an orthonormal product basis of $\mathcal{HS}$ where (λ'_{ij}) is diagonal and the diagonal elements do not vanish. Starting from the Schmidt-decomposition $|\psi\rangle = \sum_{i=1}^N s_i |ii\rangle$ we define on $\mathcal{H}_A$, as well as on $\mathcal{H}_B$:

$$P_k = |k\rangle\langle k|, \qquad 1 \leq k \leq N \tag{2.36}$$

$$Q_{jk} = \frac{1}{\sqrt{2}}(|j\rangle\langle k| + |k\rangle\langle j|), \quad 1 \leq j < k \leq N \tag{2.37}$$

$$R_{jk} = \frac{i}{\sqrt{2}}(|j\rangle\langle k| - |k\rangle\langle j|), \quad 1 \leq j < k \leq N. \tag{2.38}$$

These N^2 operators form an orthonormal basis of $\mathcal{HS}_A$ (resp. $\mathcal{HS}_B$), denoted by H_i^A (resp. H_i^B), and if one computes

$$\lambda'_{rs} = \sum_{\alpha,\beta=1}^{N} s_\alpha s_\beta \langle\alpha|H_r^A|\beta\rangle\langle\alpha|H_s^B|\beta\rangle \tag{2.39}$$

one can directly verify that (λ'_{rs}) is in the basis $H_r^A \otimes H_s^B$ diagonal with entries $(s_1^2, \ldots, s_N^2, s_1 s_2, \ldots, s_{N-1} s_N, -s_1 s_2, \ldots, -s_{N-1} s_N)$, and has the full rank. □

Corollary 2.9. If $|\psi\rangle \in \mathcal{H}_A \otimes \mathcal{H}_B$ has the Schmidt rank l then the matrix (λ_{ij}) in (2.32) has the rank l^2.
Proof. The proof is essentially the same as the proof of Lemma 2.8, the matrix with entries given by Eq. (2.39) is diagonal with only l^2 nonvanishing entries if $|\psi\rangle$ has Schmidt rank l. □

Corollary 2.10. If $|\psi\rangle \in \mathcal{H}_A \otimes \mathcal{H}_B$ has the Schmidt rank l then a decomposition in the sense of (2.32) requires l^2 Hermitean operators for every party.
Proof. If one would need less, this would be a direct contradiction to Lemma 2.8 and Corollary 2.9. Please note that we have already computed this decomposition – see (2.33). □

Lemma 2.11. Let $\vec{v}_1, ..., \vec{v}_r \in \mathbb{R}^n$ be some vectors obeying the equations

$$< \vec{v}_i, \vec{v}_j >= C \neq 0 \qquad \forall\, i \neq j. \tag{2.40}$$

$\vec{v}_r$ should be uniquely defined by $\vec{v}_1, ..., \vec{v}_{r-1}$ and the equations (2.40) while $\vec{v}_{r-1}$ should not be uniquely defined by $\vec{v}_1, ..., \vec{v}_{r-2}$ and the equations (2.40). Then we have

$$\vec{v}_r \in \mathrm{Lin}(\vec{v}_1, ..., \vec{v}_{r-1}) \tag{2.41}$$

and

$$\dim(\mathrm{Lin}(\vec{v}_1, ..., \vec{v}_r)) = r - 1 \tag{2.42}$$

where $\mathrm{Lin}(\vec{v}_1, ..., \vec{v}_r)$ denotes the linear subspace spanned by $\vec{v}_1, ..., \vec{v}_r$.
Proof. We can split $\vec{v}_r$ in two parts:

$$\vec{v}_r = \vec{v}_{r\|} + \vec{v}_{r_\perp}, \tag{2.43}$$

where $\vec{v}_{r\|} \in \mathrm{Lin}(\vec{v}_1, ..., \vec{v}_{r-1})$ and $\vec{v}_{r\perp} \perp \mathrm{Lin}(\vec{v}_1, ..., \vec{v}_{r-1})$. Since $\vec{v}_r$ is uniquely determined, it follows that $\vec{v}_{r\perp} = 0$ (otherwise $\vec{v}_r = \vec{v}_{r\|} - \vec{v}_{r\perp}$ would be a different solution) and the first part of the statement is proven. The equality in (2.42) comes from the fact that $\vec{v}_{r-1}$ is not unique. □

Lemma 2.12. Let M be one LvNM in the sense of (2.15) expanded in the $G_i^A \otimes G_j^B$ basis:

$$M = \sum_{i,j=1}^{N} c_{ij} |A_i\rangle\langle A_i| \otimes |B_j\rangle\langle B_j| = \sum_{i,j=0}^{N^2-1} \mu_{ij} G_i^A \otimes G_j^B \tag{2.44}$$

Then the $(N^2-1) \times N^2$ submatrix consisting of the lower N^2-1 rows of the $N^2 \times N^2$ matrix (μ_{ij}) (called $(\mu_{ij})^{\mathrm{red}} = (\mu_{ij})_{i>0}$) has the rank $N-1$.
Proof. We can write any of the projectors $|A_i\rangle\langle A_i|$ and $|B_j\rangle\langle B_j|$ as Bloch vectors $\hat{A}_i$ and $\hat{B}_j$ (resp. $\vec{A}_i$ and $\vec{B}_j$) with the help of (2.29). Then we have

$$(\mu_{ij})^{\mathrm{red}} = \sum_{i,j=1}^{N} c_{ij} (\vec{A}_i)^T (\hat{B}_j). \tag{2.45}$$

This is a $(N^2-1) \times N^2$ matrix, since every $(\vec{A}_i)^T \hat{B}_j$ is a $(N^2-1) \times N^2$ matrix. The range of this matrix is spanned by the vectors $(\vec{A}_i)^T$. The vectors $(\vec{A}_i)^T$ correspond to the vectors $|A_i\rangle$, and they obey relations of the form (2.31). Furthermore, $\vec{A}_N$ is uniquely determined by $\vec{A}_1, ..., \vec{A}_{N-1}$, since $|A_N\rangle$ is uniquely determined by $|A_1\rangle, ..., |A_{N-1}\rangle$. Thus, we can apply our Lemma 2.11, and the rank of $(\mu_{ij})^{\mathrm{red}}$ is $N-1$. □

Theorem 2.13. Let $|\psi\rangle \in \mathcal{H}_A \otimes \mathcal{H}_B$ have full Schmidt rank $N > 1$. Then a local measurement of the projector $|\psi\rangle\langle\psi|$ requires at least $N+1$ measurements.
Proof. If we look at $|\psi\rangle\langle\psi|$ in the form (2.32) the matrix λ_{ij} has, according to Lemma 2.8, the full rank N^2, and the reduced matrix $(\lambda_{ij})^{\mathrm{red}} = (\lambda_{ij})_{i>0}$ has a rank of N^2-1.

Since the matrix $(\mu_{ij})^{\mathrm{red}}$ corresponding to a single LvNM has, according to Lemma 2.12, the rank $N-1$ we need at least $(N^2-1)/(N-1) = N+1$ measurements. This proves the statement. □

Using a particular basis for the elements of $\mathcal{HS}$, it was shown recently that, if N is prime, the projector onto a pure entangled state with equal Schmidt coefficients and full Schmidt rank can be measured with $N+1$ LvNMs, saturating the bound from Theorem 2.13 [104].

The question remains which of these results remain valid for $N \times M$-systems with $M > N$. The answer is simple: All results remain valid. Since the maximal Schmidt rank in a $N \times M$-system is N, Theorem 2.7 can be proven in just the same way. Also the arguments which led to Theorem 2.13 can be applied.

2.5 Local detection of PPT entanglement

So far we were concerned with NPPT entangled states which are easily detected by entanglement witnesses, and obtained local decompositions requiring only few local measurements. Now we turn to the detection of a particular kind of entangled states: those having a positive partial transpose, which are undistillable because of this property. First, we construct and decompose witnesses for bipartite PPT entangled states. Then, we introduce two new families of entangled multiqubit states which have a PPT with respect to every bipartite splitting, and construct and decompose witnesses for the first family.

2.5.1 UPB states for two-qutrit systems

In Ref. [96], an easy method for constructing PPT entangled states was introduced with the help of unextendible product bases (UPBs). A UPB is a set of orthogonal product vectors spanning a subspace of the total space such that there is no other product vector orthogonal to all of them. For systems of two qutrits, the states

$$\begin{aligned} |\psi_0\rangle &= \frac{1}{\sqrt{2}}|0\rangle(|0\rangle - |1\rangle), \quad |\psi_2\rangle = \frac{1}{\sqrt{2}}|2\rangle(|1\rangle - |2\rangle), \\ |\psi_1\rangle &= \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)|2\rangle, \quad |\psi_3\rangle = \frac{1}{\sqrt{2}}(|1\rangle - |2\rangle)|0\rangle, \\ |\psi_4\rangle &= \frac{1}{3}(|0\rangle + |1\rangle + |2\rangle)(|0\rangle + |1\rangle + |2\rangle) \end{aligned} \tag{2.46}$$

form a UPB. By construction, the state

$$\rho_{\mathrm{UPB}} = \frac{1}{4}\big(\mathbb{1} - \sum_{i=0}^{4} |\psi_i\rangle\langle\psi_i|\big), \tag{2.47}$$

which is the projection on the space orthogonal to that spanned by the UPB, does not contain any product state in its range. Furthermore, since all the vectors forming the UPB are real, the partial transposition leaves the state invariant. Hence it is PPT. In conclusion, ρ_{UPB} is an entangled edge state with a positive partial transpose.

Witnesses for these states can be constructed following the method of section 2.2.3. The projectors P and Q are related to each other by

$$P = Q^{T_A} = \sum_{i=0}^{4} |\psi_i\rangle\langle\psi_i|, \tag{2.48}$$

therefore we skip Q^{T_A} and write the witness as

$$W_{\text{UPB}} = \sum_{i=0}^{4} |\psi_i\rangle\langle\psi_i| - \epsilon\mathbb{1}. \tag{2.49}$$

The witness is already in a local form, and five measurements settings are necessary to measure this witness, one for each of the five projectors, since the UPB is constructed in such a way that no two projectors can be evaluated in the same basis.

The main problem of this construction is to find ϵ. An analytical bound obtained with the method of Terhal [105] gives

$$\epsilon \geq \frac{1}{9}\frac{(6-\sqrt{30})}{6}\frac{(2-\sqrt{3})}{2} \approx 0.001297. \tag{2.50}$$

Numerical analysis using a multivariable minimization routine [106], however, leads to the much bigger value $\epsilon \approx 0.02842$. As mentioned in section 2.9, we will present an efficient method to obtain numerical lower bounds for ϵ in chapter 4, ensuring the positivity on product states of the witness.

Note that when the bound entangled state is affected by white noise, namely $\rho_p = p \cdot \rho_{\text{UPB}} + (1-p)\mathbb{1}/9$, the witness given above is still suitable for the detection of entanglement. We find that $\text{Tr}[W_{\text{UPB}}\rho_p] < 0$ when $p > (1 - 9\epsilon/5)$, leading to $p > 0.949$ for the numerical bound on ϵ.

2.5.2 Chessboard states for two qutrits

In Ref. [97], another method for constructing bound entangled states of two qutrits was introduced. The states are constructed from 4 entangled vectors as follows

$$\rho_{\text{cb}} = N\sum_{i=1}^{4} |V_i\rangle\langle V_i|, \tag{2.51}$$

$$\begin{aligned} |V_1\rangle &= (m,0,s;0,n,0;0,0,0) \\ |V_2\rangle &= (0,a,0;b,0,c;0,0,0) \\ |V_3\rangle &= (n^*,0,0;0,-m^*,0;t,0,0) \\ |V_4\rangle &= (0,b^*,0;0,-a^*,0;0,d,0), \end{aligned} \tag{2.52}$$

where the normalization is ensured by the factor $N = 1/\sum_j \langle V_j|V_j\rangle$. By choosing the phases of the $|V_i\rangle$ and the basis vectors, 6 of the parameters can be made real, so that we can assume without loss of generality that only t and s remain complex.

In matrix form, $\rho_{\rm cb}$ can then be written as

$$\rho_{\rm cb} = N \begin{pmatrix} m^2+n^2 & 0 & ms^* & 0 & 0 & 0 & nt^* & 0 & 0 \\ 0 & a^2+b^2 & 0 & 0 & 0 & ac & 0 & bd & 0 \\ sm & 0 & |s|^2 & 0 & sn & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & a^2+b^2 & 0 & bc & 0 & -ad & 0 \\ 0 & 0 & ns^* & 0 & m^2+n^2 & 0 & -mt^* & 0 & 0 \\ 0 & ac & 0 & cb & 0 & c^2 & 0 & 0 & 0 \\ tn & 0 & 0 & 0 & -tm & 0 & |t|^2 & 0 & 0 \\ 0 & bd & 0 & -da & 0 & 0 & 0 & d^2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}. \tag{2.53}$$

The states are sometimes called chessboard states because the only nonvanishing matrix elements are distributed as the fields of the same color on a chessboard.

In Ref. [97], two methods were provided of ensuring that $\rho_{\rm cb}$ is bound entangled. We will employ the first one, i.e., we demand that $\rho_{\rm cb} = \rho_{\rm cb}^{T_A}$, which is fulfilled for $t = ad/m$ and $s = ac/n$ real, and implies that $P = Q$. The kernel of $\rho_{\rm cb}$ (and of $\rho_{\rm cb}^{T_A}$) is spanned by the (non-normalized) vectors

$$|k_1\rangle = |22\rangle \tag{2.54}$$

$$|k_2\rangle = (\frac{m}{n}, 0, -\frac{m^2+n^2}{ac}; 0, 1, 0; 0, 0, 0) \tag{2.55}$$

$$|k_3\rangle = (0, -\frac{ac}{a^2+b^2}, 0; -\frac{bc}{a^2+b^2}, 0, 1; 0, 0, 0) \tag{2.56}$$

$$|k_4\rangle = (-\frac{ad}{mn}, 0, \frac{d}{c}; 0, 0, 0; 1, 0, 0) \tag{2.57}$$

$$|k_5\rangle = (0, -\frac{bd}{a^2+b^2}, 0; \frac{ad}{a^2+b^2}, 0, 0; 0, 1, 0). \tag{2.58}$$

In order to construct and decompose the witness $W_{\rm cb}$ in terms of projectors onto product states with the method of section 2.9, we first examine whether there are more product vectors in the kernel of $\rho_{\rm cb}$, so we try to solve

$$(a_1|0\rangle + a_2|1\rangle + a_3|2\rangle) \otimes (b_1|0\rangle + b_2|1\rangle + b_3|2\rangle) = \sum_{i=1}^{5} x_i |k_i\rangle. \tag{2.59}$$

When writing down the equations one can see that the x_i can be substituted by products of one a_j and one b_k. Then it is possible to solve the set of equations which is then linear in the parameters b_k. This in turn gives two equations for the a_j. The first solution is given by

$$|k_4'\rangle = |k_4\rangle - \frac{mn}{ac}|k_1\rangle = (-\frac{ad}{mn}, 0, 1) \otimes (1, 0, -\frac{mn}{ac}), \tag{2.60}$$

and with

$$\begin{aligned}
\alpha_1 &\equiv (m^2+n^2)bmn-(a^2+b^2)am^2 \\
\alpha_3 &\equiv ad^2n^2 \\
\alpha_{13} &\equiv (m^2+n^2)(mn+ab)d-2abdm^2 \qquad (2.61)\\
\gamma_1^{0,1} &\equiv \Big(-\alpha_{13}\pm\sqrt{\alpha_{13}^2-4\alpha_1\alpha_3}\Big)/2\alpha_3 \\
\gamma_2^{0,1} &\equiv \Big[\frac{1}{am^2}\Big(bmn+d(mn+ab)\gamma_1^{0,1}+ad^2(\gamma_1^{0,1})^2\Big)\Big]^{\frac{1}{2}} \qquad (2.62)
\end{aligned}$$

the other solutions can be written as

$$|e,f\rangle = a_1\Big(1,\pm\gamma_2^{0,1},\gamma_1^{0,1}\Big)\otimes b_2\Big(\pm\frac{m^2\gamma_2^{0,1}}{mn+ad\gamma_1^{0,1}},1,\mp\frac{a^2+b^2+bd\gamma_1^{0,1}}{ac\gamma_2^{0,1}}\Big). \qquad (2.63)$$

The parameters a_1 and b_2 can be used to normalize the vectors. We found 6 product vectors in the kernels. Of those vectors, 5 will be linearly independent in general. Since they do not form an orthonormal set in general, we cannot construct P and Q from them. However, the witness can also be constructed by using instead of the projector onto the kernel of $\rho_{\rm cb}$ ($\rho_{\rm cb}^{T_A}$) an operator $\tilde{P}$ ($\tilde{Q}$) which is strictly positive on the range of the kernel of $\rho_{\rm cb}$ ($\rho_{\rm cb}^{T_A}$). This will only affect the value of ϵ. Here $\tilde{P}=\tilde{Q}$ can be constructed by summing the projectors onto five linearly independent product vectors from the kernel of $\rho_{\rm cb}$. Since the vectors are real, we have in addition that $\tilde{P}=\tilde{Q}^{T_A}$. Hence in general the pre-witness $\bar{W}_{\rm cb}$ can be decomposed into 5 projectors onto product vectors requiring 5 settings to measure the witness $W_{\rm cb}$, one for each of the projectors.

Again, an upper bound on the coefficient ϵ can be calculated numerically with the use of a multivariable minimization routine [106], while we refer to chapter 4 for a method capable of generating a lower bound.

2.5.3 Horodecki states for 2×4 systems

The first example of a PPT entangled state in the literature was introduced by P. Horodecki in Ref. [29]. It is a state of a 2×4 system which can be written in matrix as

$$\rho_b=\frac{1}{7b+1}\begin{pmatrix}
b & 0 & 0 & 0 & 0 & b & 0 & 0\\
0 & b & 0 & 0 & 0 & 0 & b & 0\\
0 & 0 & b & 0 & 0 & 0 & 0 & b\\
0 & 0 & 0 & b & 0 & 0 & 0 & 0\\
0 & 0 & 0 & 0 & \frac{1}{2}(1+b) & 0 & 0 & \frac{1}{2}\sqrt{1-b^2}\\
b & 0 & 0 & 0 & 0 & b & 0 & 0\\
0 & b & 0 & 0 & 0 & 0 & b & 0\\
0 & 0 & b & 0 & \frac{1}{2}\sqrt{1-b^2} & 0 & 0 & \frac{1}{2}(1+b)
\end{pmatrix}, \qquad (2.64)$$

where $b\in[0,1]$. For $b=0,1$ the matrix ρ_b is separable, and bound entangled for all other values of b. In the following we assume $b\neq 0,1$. The kernel of ρ_b is spanned

by the entangled vectors

$$|k_1\rangle = \frac{1}{\sqrt{2}}(1,0,0,0;0,-1,0,0) \tag{2.65}$$

$$|k_2\rangle = \frac{1}{\sqrt{2}}(0,1,0,0;0,0,-1,0) \tag{2.66}$$

$$|k_3\rangle = \frac{1}{\sqrt{2+y^2}}(0,0,1,0;y,0,0,-1), \tag{2.67}$$

where $y = \sqrt{(1-b)/(1+b)}$, so any vector in the kernel can be represented as

$$|k\rangle = (A,B,C,0;yC,-A,-B,-C) \tag{2.68}$$

where A, B, C, D are complex parameters. For $|k\rangle$ to be a product vector, it must be of the form

$$|e,f\rangle = (r,s)\otimes(A',B',C',D') \equiv (r(A',B',C',D');s(A',B',C',D')), \tag{2.69}$$

where r, s, A', B', C', and D' are complex parameters. It can be readily checked that there is no possibility to write $|k\rangle$ in this form, therefore there is no product vector in the kernel of ρ_b. On the other hand, the kernel of $\rho_b^{T_B}$ is spanned by the entangled vectors

$$|k_1\rangle = \frac{1}{\sqrt{2}}(0,0,1,0;0,-1,0,0) \tag{2.70}$$

$$|k_2\rangle = \frac{1}{\sqrt{2}}(0,0,0,1;0,0,-1,0) \tag{2.71}$$

$$|k_3\rangle = \frac{1}{\sqrt{2+y^2}}(0,1,0,0;1,0,0,y), \tag{2.72}$$

and does not contain any product vector, either. Therefore, the decomposition of the witness of the from of Eq. (2.9) in projectors onto product vectors is a rather tedious task. On the other hand, we can write down the witness W as in Eq. (2.9) and then decompose it as in Eq. (2.17) and (2.33)

$$W = P + Q^{T_A} - \epsilon\mathbb{1} = \sum_{i=0}^{3}\sum_{j=0}^{15} w_{ij}\sigma_i\otimes\tau_j \equiv \sum_{i=0}^{3}\sigma_i\otimes\tilde{\tau}_i \tag{2.73}$$

in a straightforward manner. Here σ_i and τ_i, for $i > 0$, are the generators of the SU(2) and SU(4), respectively, while $\sigma_0 = \mathbb{1}$ and $\tau_0 = \mathbb{1}_4/4$. The matrices $\tilde{\tau}$ can be calculated from

$$2\tilde{\tau}_i = \mathrm{Tr}_A\big(W(\sigma_i\otimes\mathbb{1}_4)\big), \tag{2.74}$$

they turn out to be

$$\tilde{\tau}_0 = \frac{c}{2}\begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} + (3-4\epsilon)\tau_0, \tag{2.75}$$

$$\tilde{\tau}_1 = \frac{1}{4}\begin{pmatrix} 0 & -cy^2 & 2cy & 0 \\ -cy^2 & 0 & -2 & 2cy \\ 2cy & -2 & 0 & -c(4+y^2) \\ 0 & 2cy & -c(4+y^2) & 0 \end{pmatrix}, \tag{2.76}$$

$$\tilde{\tau}_2 = \frac{i}{4}\begin{pmatrix} 0 & -cy^2 & -2cy & 0 \\ cy2 & 0 & -2 & -2cy \\ 2cy & 2 & 0 & -c(4+y^2) \\ 0 & 2cy & c(4+y^2) & 0 \end{pmatrix}, \tag{2.77}$$

and

$$\tilde{\tau}_3 = -cy^2\tau_0, \tag{2.78}$$

where $c = 1/(2+y^2)$. The number of correlated local measurement settings is 4, but Alice and Bob need only 3 different settings each to measure the witness.

As before, upper bound on the coefficient ϵ can be obtained by minimization routines [106], while lower bounds can be produced with the method to be introduced in chapter 4.

After having discussed three examples of bipartite PPT entangled states, we turn now to the multiqubit setting.

2.5.4 A family of n-qubit PPTES from a GHZ state

In Ref. [98], the following family of states of three qubits depending on three positive parameters a, b, and c has been introduced

$$\rho_3(a,b,c) = \frac{1}{N}\begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & a & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & b & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{c} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & c & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \frac{1}{b} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{a} & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}, \tag{2.79}$$

where the normalization is ensured by $N = 2 + \sum_x (x + 1/x)$, $x = a, b, c$, and we exchanged $c \leftrightarrow 1/c$ as compared to the original notation. It can be shown [98] that if $abc \neq 1$, then the states have the following very peculiar properties

1. The states have a positive partial transpose with respect to every bipartite split.

2. The states are separable with respect to every biparte split.

3. The states are entangled.

Here, we generalize this construction to states of n qubits, for arbitrary n, and construct entanglement witnesses detecting them for every n. We further present local decompositions for small n and prove the optimality for $n = 3$.

Let us first introduce some notation. In the standard computational basis for n qubits, the basis states can be labeled by an index k running from $000\ldots0 \equiv 0$ up to $111\ldots1 \equiv 2^n - 1$. By $\bar{k}$ we denote the number which is reached by inverting all digits in the binary representation, e.g., $\bar{0} = 1$.

The n-qubit family can then be written as

$$\rho_n(\mathbf{a}) = \frac{1}{N}\Big[2|\mathrm{GHZ}_n\rangle\langle\mathrm{GHZ}_n| + \sum_{k=1}^{2^{n-1}-1} a_k|k\rangle\langle k| + \frac{1}{a_k}|\bar{k}\rangle\langle\bar{k}|\Big], \tag{2.80}$$

where $N = n + \sum_{k=1}^{2^{n-1}-1}(a_k + \frac{1}{a_k})$, and $\mathbf{a}$ is the vector of positive coefficients a_k. These coefficients have to fulfil $a_1 a_2 a_4 \cdots a_{2^{n-1}} = \Pi_{k=0}^{n-1} a_{2^k} \neq 1$ in order to ensure that the states are entangled. Now we will prove the named properties of the states.

Property 1 (positive partial transpose). When the states (2.80) are partially transposed, only the projector onto the GHZ state is affected, while all the product vectors remain unchanged. To be precise, only the off-diagonal term $|0\rangle\langle 1|^{\otimes n} + |1\rangle\langle 0|^{\otimes n}$ is changed. If I is a list of all the indices of those systems with respect to which the state is transposed, then this term is transformed to $|k_I\rangle\langle\bar{k}_I| + |\bar{k}_I\rangle\langle k_I|$, where k_I has a 0 on all binary positions except for those which have been partially transposed. Hence, the partially transposed states is a mixture of a state with only positve entries on the diagonal and the following state in the subspace $\{|k_I\rangle, |\bar{k}_I\rangle\}$

$$\begin{pmatrix} a_{k_I} & 1 \\ 1 & \frac{1}{a_{k_I}} \end{pmatrix}, \tag{2.81}$$

which is equal to the projector onto the nonnormalized state $\sqrt{a_{k_I}}|k_I\rangle + (1/\sqrt{a_{k_I}})|\bar{k}_I\rangle$. Therefore $\rho_n(\mathbf{a})$ is PPT with respect to every bipartite split. □

Property 2 (biseparability). If we fix a certain bipartite splitting labelled by I, then we saw in the proof of property 1 that the partial transpose affects only a small subspace of the whole space, so that we can write

$$\rho_n(\mathbf{a}) = \rho_{s,I} + \tilde{\rho}_I. \tag{2.82}$$

Here $\rho_{s,I}$ is the unaffected part which contains only projectors onto product states, and

$$\tilde{\rho}_I = \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & a_{k_I} & 0 & 0 \\ 0 & 0 & \frac{1}{a_{k_I}} & 0 \\ 1 & 0 & 0 & 1 \end{pmatrix}, \tag{2.83}$$

living in the subspace $\{|0\rangle, |k_I\rangle, |\bar{k}_I\rangle, |2^n - 1\rangle\}$. Since $\rho_{s,I}$ is separable, $\rho_n(\mathbf{a})$ is biseparable if $\tilde{\rho}$ is separable. But $\tilde{\rho}$ lives in a 2×2 dimensional subspace, and it is positive with respect to partial transposition with respect to one part of this subspace. In

these dimensions, the PPT criterion is a sufficient criterion for separability. Hence $\rho_n(\mathbf{a})$ is separable with respect to every bipartite split. □

Property 3 (entanglement). We prove that the states are entangled when the parameters $\mathbf{a}$ fulfil the conditions named above by showing that the states violate the range criterion of section 1.1.2 extremely in this case, i.e., that there exists no product vector $|\phi\rangle = \otimes_{m=1}^{n}|\phi_m\rangle$ such that i) $|\phi\rangle \in R(\rho_n(\mathbf{a}))$ and ii) $|\phi\rangle^{*I} \in R(\rho_n(\mathbf{a})^{T_I})$ for all bipartite splittings labelled by I. Here $*I$ denotes complex conjugation of all $|\phi_l\rangle$ with $l \in I$ and T_I denotes partial transposition with respect to all systems listed in I. With this notation, $|k_I\rangle = |0\rangle^{T_I}$.

This can be proven by looking at the respective kernels. Let $|\phi_m\rangle = \cos\theta_m|0\rangle + \exp(i\varphi_m)\sin\theta_m|1\rangle \equiv c_m|0\rangle + e_m s_m|1\rangle$, and $\theta_m \in [0, \pi/4]$. The kernel of $\rho_n(\mathbf{a})$ is spanned by $|\mathrm{GHZ}_n^-\rangle \equiv (|0\rangle^{\otimes n} - |1\rangle^{\otimes n})/\sqrt{2}$. From Eq. (2.81) it is easy to see that the kernel of $\rho_n(\mathbf{a})^{T_I}$ is spanned by the (nonnormalized) vector $|K_I\rangle \equiv (|k_I\rangle - a_{k_I}|\bar{k}_I\rangle)$, for all possible sets I. A product vector fulfilling the range criterion then has to obey

$$\langle \mathrm{GHZ}_n^-|\phi\rangle = 0 \tag{2.84}$$

$$\langle K_I|\phi\rangle^{*I} = 0, \tag{2.85}$$

for all sets I. These conditions are equivalent to

$$\Pi_{j=1}^{n}\tan(\theta_j) = \exp(-i\sum_{l=1}^{n}\varphi_l) = 1 \tag{2.86}$$

$$\frac{\Pi_{j\notin I}\tan(\theta_j)}{\Pi_{l\in I}\tan(\theta_l)} = \frac{1}{a_{k_I}}\exp(-i\sum_{m=1}^{n}\varphi_m) = \frac{1}{a_{k_I}}, \tag{2.87}$$

where the last equalities on the right hand sides are due to the restriction of the angles θ_k to the interval $[0, \pi/4]$. Therefore, we can without loss of generality assume that $\varphi_k = 0$, for all k. The second condition can be rewritten using the first as

$$\left(\Pi_{l\in I}\tan(\theta_l)\right)^2 = a_{k_I}. \tag{2.88}$$

A contradiction can now be reached by multiplying the conditions (2.88) for several splittings such that the dependence on the angles can be taken out by virtue of (2.86). For instance, let us consider the sets $I_j = \{j\}, j = 1, 2, \ldots, n$, i.e., all the sets where only one party is separated from all the others. If we multiply the left hand sides and right hand sides of Eq. (2.88) for all I_j, we obtain

$$\left(\Pi_{j=1}^{n}\tan(\theta_j)\right)^2 = \Pi_{j=1}^{n}a_{k_{I_j}} = 1, \tag{2.89}$$

where the last equality follows again from Eq. (2.86). If we demand that $\Pi_{j=1}^{n}a_{k_{I_j}} \neq 1$, then $\rho_n(\mathbf{a})$ violates the range criterion in an extremal sense, and is therefore entangled. Further, $k_{I_j} = 2^{n-j}$, so that we arrive at the condition below Eq. (2.80).

□

Note that the minimal number of conditions one has to combine here is 3, because two sets I together can only contain each party once if I_1 is the complement of I_2, so that $a_{k_{I_1}} \cdot a_{k_{I_2}} = 1$ by construction, from which no contradiction can be obtained.

Finally, we would like to point out that it is also easy to generate states that are NPPT with respect to a splitting I by choosing $a_{k_I} \neq (a_{\bar{k}_I})^{-1}$. If I contains more than one party, then a violation of condition (2.89) still ensures entanglement of the whole state, not only with respect to I. On the other hand, if $I = \{1\}$, say, then by multiplying the conditions (2.88) for the sets $I = \{1, 2\}$ and $I_j, j = 3, 4, \ldots n$, we obtain

$$a_{k_{\{1,2\}}} \Pi_{j=3}^{n} a_{k_{I_j}} \neq 1 \tag{2.90}$$

as a condition for entanglement of the whole state.

2.5.5 Local detection of the family

We adopt the method from section 2.2.3 to the multiqubit setting to

$$W = P + \sum_I Q_I^{T_I} - \epsilon \mathbb{1}, \tag{2.91}$$

$$\epsilon = \inf_{|\phi\rangle} \langle\phi| \big(P + \sum_I Q_I^{T_I}\big) |\phi\rangle \tag{2.92}$$

where $Q_I = K(\rho_n(\mathbf{a})^{T_I})$, and $|\phi\rangle = \otimes_{i=1}^n |\phi_i\rangle$. The sum can be performed over all I, but it suffices to take into account those splittings that are involved in the contradiction from the proof of entanglement of the states, which is reflected in the condition on the parameters $\mathbf{a}$. This ensures already that $\epsilon > 0$. The witness is rather easy to construct because

$$\begin{aligned} Q_I^{T_I} &= \frac{1}{1 + a_{k_I}^2} |K_I\rangle\langle K_I|^{T_I} \\ &= \frac{1}{1 + a_{k_I}^2} \Big(|k_I\rangle\langle k_I| + a_{k_I}^2 |\bar{k}_I\rangle\langle \bar{k}_I| - a_{k_I} \big(|0\rangle\langle 1|^{\otimes n} + |1\rangle\langle 0|^{\otimes n} \big) \Big). \end{aligned} \tag{2.93}$$

It takes the form

$$\begin{aligned} W &= \frac{1}{2} \Big(|0\rangle\langle 0|^{\otimes n} + |1\rangle\langle 1|^{\otimes n} \Big) + \sum_I \frac{1}{1 + a_{k_I}^2} \Big(|k_I\rangle\langle k_I| + a_{k_I}^2 |\bar{k}_I\rangle\langle \bar{k}_I| \Big) \\ &\quad - \Big(\frac{1}{2} + \sum_I \frac{a_{k_I}}{1 + a_{k_I}^2} \Big) \Big(|0\rangle\langle 1|^{\otimes n} + |1\rangle\langle 0|^{\otimes n} \Big) - \epsilon \mathbb{1}. \end{aligned} \tag{2.94}$$

This is already a local decomposition except for the term $|0\rangle\langle 1|^{\otimes n} + |1\rangle\langle 0|^{\otimes n}$, which is also the problematic term when the projector onto the GHZ state has to be decomposed locally. It can be found by iterating the trick from section 2.4.1. Starting from

$$|00\rangle\langle 11| + |11\rangle\langle 00| = \frac{1}{2} \big(\sigma_x \sigma_x - \sigma_y \sigma_y \big) \tag{2.95}$$

$$|00\rangle\langle 11| - |11\rangle\langle 00| = \frac{i}{2} \big(\sigma_x \sigma_y + \sigma_y \sigma_x \big) \tag{2.96}$$

it is possible to obtain

$$\begin{aligned} |000\rangle\langle 111| + |111\rangle\langle 000| &= \frac{1}{4}\big(\sigma_x\sigma_x\sigma_x - \sigma_x\sigma_y\sigma_y - \sigma_y\sigma_x\sigma_y - \sigma_y\sigma_y\sigma_x\big) \\ &= \frac{1}{2}\sigma_x\sigma_x\sigma_x - \frac{1}{8}(\sigma_x+\sigma_y)^{\otimes 3} - \frac{1}{8}(\sigma_x-\sigma_y)^{\otimes 3}, \end{aligned} \tag{2.97}$$

where we left out the tensor product signs. The number of measurements necessary is reduced from four to three in the last decomposition, where only the measurements $\sigma_x\sigma_x\sigma_x$, and $\big((\sigma_x \pm \sigma_y)/\sqrt{2}\big)^{\otimes 3}$ have to be performed. For the latter measurements, each party has to perform spin measurements in directions in the $x-y$ plane at an angle of $\pm 45^\circ$ with respect to the x axis.

For four qubits,

$$\begin{aligned} |0\rangle\langle 1|^{\otimes 4} + |1\rangle\langle 0|^{\otimes 4} &= \frac{1}{8}\Big(\sigma_x^{\otimes 4} + \sigma_y^{\otimes 4} - \sigma_x\sigma_x\sigma_y\sigma_y - \sigma_x\sigma_y\sigma_x\sigma_y - \sigma_x\sigma_y\sigma_y\sigma_x \\ &\quad -\sigma_y\sigma_x\sigma_y\sigma_x - \sigma_y\sigma_y\sigma_x\sigma_x - \sigma_y\sigma_x\sigma_x\sigma_y\Big) \\ &= \frac{1}{4}\Big(\sigma_x^{\otimes 4} + \sigma_y^{\otimes 4} - \frac{1}{4}(\sigma_x+\sigma_y)^{\otimes 4} - \frac{1}{4}(\sigma_x-\sigma_y)^{\otimes 4}\Big). \end{aligned} \tag{2.98}$$

Here, eight measurements are necessary for the first decomposition and only four for the second. For five qubits, the number of locally correlated measurement settings necessary is eight already for a decomposition similar to the decompositions in the lower lines of the Eqs. (2.97) and (2.98), and the number seems to increase exponentially with the number of qubits. However, it is by no means clear that these are the optimal decompositions.

Optimality of the decomposition for three qubits

We consider again the decomposition of the witness for three qubits and prove that the decomposition is optimal with similar methods as applied in section 2.4. The witness is later used in chapter 3, where a network is constructed for the generation of $\rho_3(\mathbf{a})$, given in Eq. (2.79). The witness for $\rho_3(\mathbf{a})$ is given by

$$\begin{aligned} \rho_3 &= \frac{1}{N}\Big(2|\mathrm{GHZ}\rangle\langle\mathrm{GHZ}| + a|001\rangle\langle 001| + \frac{1}{a}|110\rangle\langle 110| + b|010\rangle\langle 010| + \frac{1}{b}|101\rangle\langle 101| \\ &\quad + c|100\rangle\langle 100| + \frac{1}{c}|011\rangle\langle 011|\Big). \end{aligned} \tag{2.99}$$

From Eq. (2.89) we see that the condition for entanglement is $abc \neq 1$. The witness is given by

$$\begin{aligned} W &= \frac{1}{2}\Big(|000\rangle\langle 000| + |111\rangle\langle 111|\Big) + \frac{1}{1+a^2}\Big(|001\rangle\langle 001| + a^2|110\rangle\langle 110|\Big) \\ &\quad + \frac{1}{1+b^2}\Big(|010\rangle\langle 010| + b^2|101\rangle\langle 101|\Big) + \frac{1}{1+c^2}\Big(|100\rangle\langle 100| + c^2|011\rangle\langle 011|\Big) \\ &\quad - \Big(\frac{1}{2} + \frac{a}{1+a^2} + \frac{b}{1+b^2} + \frac{c}{1+c^2}\Big)\Big(|000\rangle\langle 111| + |111\rangle\langle 000|\Big) - \epsilon\mathbb{1}. \end{aligned} \tag{2.100}$$

Proposition 2.14. The witness (2.100) cannot be measured with three LvNMs, i.e., the second decomposition of Eq. (2.97) is optimal.

Proof. The proof is an extension of the two-qubit case. First, we write the witness in the $\sigma_i \otimes \sigma_j \otimes \sigma_k$ basis: $W = 1/8 \sum_{i,j,k=0}^{3} \lambda_{ijk} \sigma_i \otimes \sigma_j \otimes \sigma_k$. From this we define the reduced $4 \times 3 \times 3$ tensor $(\lambda^{\mathrm{red}}_{ijk})_{i,j,k} := (\lambda_{ijk})_{j,k>0}$ which is given by

$$\begin{aligned}
\lambda^{\mathrm{red}}_{0jk} &= \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} =: A^{(0)} \quad \lambda^{\mathrm{red}}_{1jk} = \begin{pmatrix} -s_1 & 0 & 0 \\ 0 & s_1 & 0 \\ 0 & 0 & 0 \end{pmatrix} =: A^{(1)} \\
\lambda^{\mathrm{red}}_{2jk} &= \begin{pmatrix} 0 & s_1 & 0 \\ s_1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} =: A^{(2)} \quad \lambda^{\mathrm{red}}_{3jk} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -s_2 \end{pmatrix} =: A^{(3)}.
\end{aligned}$$

Here we used the abbreviations $s_1 = 1 + 2\sum_x \frac{x}{1+x^2}$ and $s_2 = \sum_x \frac{1-x^2}{1+x^2}$, where x is summed over a, b, c.

Let us now see what can be achieved with one measurement setting. One measurement setting is of the form

$$\begin{aligned}
M &= \sum_{r,s,t=0}^{1} c_{rst} |A_r\rangle\langle A_r| \otimes |B_s\rangle\langle B_s| \otimes |C_t\rangle\langle C_t| \\
&= \sum_{i,j,k=0}^{3} \mu_{ijk} \sigma_i \otimes \sigma_j \otimes \sigma_k, \qquad (2.101)
\end{aligned}$$

Defining s^A as the Bloch vector of $|A_0\rangle\langle A_0|$ (and similarly s^B and s^C for $|B_0\rangle\langle B_0|$ and $|C_0\rangle\langle C_0|$) and using the same argumentation as in the two-qubit case, it is easy to see that the reduced $4 \times 3 \times 3$ tensor μ^{red}_{ijk} is given by

$$\mu^{\mathrm{red}}_{ijk} = \sum_{r,s,t=0}^{1} c_{rst} (-1)^{s+t} (\vec{s}^A_r)_i s^B_j s^C_k, \qquad (2.102)$$

where $\vec{s}^A_r = (1/2; (-1)^r (s^A_1, s^A_2, s^A_3))$. Hence μ^{red}_{ijk} is of the form $a_i b_j c_k$, which can be defined as a tensor of rank one in analogy to rank one matrices [107], c.f. section 2.4.1.

If the witness W could be measured with three local measurements, then it would have to be possible to write its reduced tensor as the sum of three rank one tensors, i.e.,

$$\begin{aligned}
\lambda^{\mathrm{red}}_{ijk} &= \sum_{r=1}^{3} a^{(r)}_i b^{(r)}_j c^{(r)}_k \\
&= a^{(1)}_i B^{(1)}_{jk} + a^{(2)}_i B^{(2)}_{jk} + a^{(3)}_i B^{(3)}_{jk}. \qquad (2.103)
\end{aligned}$$

The $B^{(r)}$ are matrices of rank one, and their linear combination has to span the same three-dimensional subspace as the $A^{(i)}$ in the space of 3×3 matrices. In this

case it would be possible to write every single $B^{(r)}$ as a linear combination of the $A^{(i)}$. However, a general linear combination of the $A^{(i)}$ is of the form:

$$\mathcal{A} = \begin{pmatrix} -\alpha & \beta & 0 \\ \beta & \alpha & 0 \\ 0 & 0 & \gamma \end{pmatrix} \tag{2.104}$$

This is of rank one if and only if $\alpha = \beta = 0$. Thus, we arrive at a contradiction, the B_i cannot be of rank one and linear independent. □

The proof shows a link between the minimal number of measurements needed to the rank of the reduced λ tensor, a connection which is investigated further in Ref. [107].

2.5.6 A family of three qubit PPTES from a W state

In this section, we show that it is also possible to construct a family of three qubit states with the same properties as the states constructed from a GHZ state in the preceding section. It is given by

$$\rho_{\mathrm{W}}(a) = \frac{1}{N}\Big(3|\mathrm{W}\rangle\langle\mathrm{W}| + 2\mathrm{a}|000\rangle\langle 000| + \frac{1}{\mathrm{a}}\big(|011\rangle\langle 011| + |101\rangle\langle 101| + |110\rangle\langle 110|\big)\Big), \tag{2.105}$$

where $N = 3 + 2a + 3/a$ and $a > 0$. This state is invariant with respect to the exchange of two parties, which simplifies the following proofs.

Property 1 (positive partial transpose). Partially transposing the state with respect to subsystem A leads to

$$\rho^{T_A} = \frac{1}{N}\begin{pmatrix} 2a & 0 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{a} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & \frac{1}{a} & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & \frac{1}{a} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \tag{2.106}$$

which is a convex combination of projectors onto the product states $|011\rangle$ and $|100\rangle$ and on the (partly unnormalized) states $|0\rangle_A|\psi^+\rangle_{BC}$, $(\sqrt{a}|000\rangle + \frac{1}{\sqrt{a}}|1\rangle_A|01\rangle_{BC})$, and $(\sqrt{a}|000\rangle + \frac{1}{\sqrt{a}}|1\rangle_A|10\rangle_{BC})$, hence it is positive. The states partially transposed with respect to B and C are of a similar form, so that $\rho_{\mathrm{W}}(a)$ is PPT with respect to every partition. □

Property 2 (biseparability). We just consider the partition $A - BC$, and because the state is invariant under the exchange of two particles, separability of the other splits follows. Identifying $|00\rangle_{BC} \equiv |0\rangle, |01\rangle_{BC} \equiv |1\rangle$, and so on, we can write the state as

$$\rho_{\mathrm{W}}(a) = \tilde{\rho} + \frac{1}{Na}|03\rangle\langle 03|, \tag{2.107}$$

where $\tilde{\rho}$ has support on $\mathbb{C}^2 \otimes \mathbb{C}^3$ only, because the projector onto $|03\rangle$ and $\tilde{\rho}$ remain independent under partial transposition, c.f. Eq. (2.106). Further, $\tilde{\rho} \geq 0$ by construction of the state. Therefore it is separable, because the PPT criterion is also sufficient for 2×3 systems, so that $\rho_{\mathrm{W}}(a)$ is biseparable with respect to this splitting. □

Property 3 (entanglement). In order to show that the states are entangled despite the first two properties, we use the range criterion from section 1.1.2 as in the preceding section, i.e. we show that there is no product vector $|a,b,c\rangle \in R(\rho)$ fulfilling $|a,b,c\rangle^{*X} \in R(\rho^{T_X})$ for $X = A, B, C$ by looking at the respective kernels.

The kernel of ρ_{W} is spanned by $|\mathrm{W}_{1,2}\rangle = (|001\rangle + \exp(i\alpha_{1,2})|010\rangle + \exp(i2\alpha_{1,2})|100\rangle)/\sqrt{3}$, where $\alpha_1 = 2\pi/3 = \alpha_2/2$, and $|111\rangle$. The kernel of ρ^{T_A} is spanned by $|0\rangle_A|\psi^-\rangle_{BC}$, $(\frac{1}{\sqrt{a}}|000\rangle - \sqrt{a}|1\rangle_A|\psi^+\rangle_{BC})$, and $|111\rangle$. The kernels of ρ^{T_B} and ρ^{T_C} are of a similar form again, with (A, BC) replaced by (B, AC) for ρ^{T_B} and replaced by (C, AB) for ρ^{T_C}. The condition $K(\rho)|a,b,c\rangle = 0$ is fulfilled by $|a,b,c\rangle \in \{|000\rangle, |011\rangle, |101\rangle, |110\rangle\}$ only. On the other hand, $K(\rho^{T_X})|a,b,c\rangle^{*X} = 0$ for $|a,b,c\rangle = |0\rangle_X|11\rangle$ only, which shows that there exists no product vector with the desired properties. Hence $\rho_{\mathrm{W}}(a)$ are PPTES for all $a > 0$. □

2.6 Experimental implementation

In this section, we briefly discuss part of the experiments performed by M. Bourennane and coworkers in the group of H. Weinfurter in Munich [VII]. They produced two multiparticle states of photons entangled in the polarization degrees of freedom by parametric down-conversion, and proved that the produced states are truely multipartite entangled with entanglement witnesses. The first state is the three-qubit W state [63, 108]

$$|\mathrm{W}\rangle = \sqrt{\frac{1}{3}}\Big(|001\rangle + |010\rangle + |100\rangle\Big) \tag{2.108}$$

and the second is a four photon entangled state of the form

$$|\psi_4\rangle = \frac{1}{\sqrt{3}}\Big(|0011\rangle + |1100\rangle - \frac{1}{2}\big(|0110\rangle + |1001\rangle + |0101\rangle + |1010\rangle\big)\Big). \tag{2.109}$$

Here, we will concentrate on the first one and refer to Ref. [VII] for the treatment of the second one. For the detection of its entanglement the following two witnesses $\mathcal{W}_{\mathrm{W}}^{(1)}$ and $\mathcal{W}_{\mathrm{W}}^{(2)}$ [98] can be used. The first witness is constructed according to section 2.2.2 and can be decomposed with methods similar to those applied before in this chapter, resulting in

$$\begin{aligned}\mathcal{W}_{\mathrm{W}}^{(1)} &= \frac{2}{3}\mathbb{1} - |\mathrm{W}\rangle\langle\mathrm{W}| = \frac{1}{24}\Big[17\cdot\mathbb{1}^{\otimes 3} + 7\cdot\sigma_{\mathrm{z}}^{\otimes 3} + 3\cdot\big(\sigma_{\mathrm{z}}\mathbb{1}\mathbb{1} + \mathbb{1}\sigma_{\mathrm{z}}\mathbb{1} + \mathbb{1}\mathbb{1}\sigma_{\mathrm{z}}\big) \\ &\quad +5\cdot\big(\sigma_z\sigma_z\mathbb{1} + \sigma_z\mathbb{1}\sigma_z + \mathbb{1}\sigma_z\sigma_z\big) - (\mathbb{1} + \sigma_z + \sigma_x)^{\otimes 3} - (\mathbb{1} + \sigma_z - \sigma_x)^{\otimes 3} \\ &\quad -(\mathbb{1} + \sigma_z + \sigma_y)^{\otimes 3} - (\mathbb{1} + \sigma_z - \sigma_y)^{\otimes 3}\Big]. \end{aligned} \tag{2.110}$$

Its expectation value is positive on biseparable and fully separable states. It thus detects all states belonging to the two classes of states with genuine tripartite entanglement, the W class and the GHZ class, but without distinguishing between them. The factor 2/3 corresponds to the maximal squared overlap between the W state and biseparable states. From this we also see that a mixture of $|\mathrm{W}\rangle$ and white noise, $\rho = p|\mathrm{W}\rangle\langle\mathrm{W}| + (1-\mathrm{p})\mathbb{1}/8$, exhibits tripartite entanglement for a noise contribution of up to $p > 13/21$. This decomposition requires five measurement settings, namely $\sigma_z^{\otimes 3}$ and $((\sigma_z \pm \sigma_i)/\sqrt{2})^{\otimes 3}; i = x, y$, see also Fig. 2.3. The number of correlated local measurement settings can be shown to be optimal with similar methods as those used before in this section [VI].

A witness that detects genuine tripartite entanglement and, with just one extra local measurement, allows to distinguish between the W and GHZ states, in its original and decomposed form is [VI]

$$\begin{aligned}\mathcal{W}_{\mathrm{W}}^{(2)} &= \frac{1}{2}\mathbb{1} - |\overline{\mathrm{GHZ}}\rangle\langle\overline{\mathrm{GHZ}}| \\ &= \frac{1}{16}\Big[6\cdot\mathbb{1}^{\otimes 3} + 4\cdot\sigma_z^{\otimes 3} - 2\cdot\big(\sigma_y\sigma_y\mathbb{1} + \sigma_y\mathbb{1}\sigma_y + \mathbb{1}\sigma_y\sigma_y\big) \\ &\quad -(\sigma_z+\sigma_x)^{\otimes 3} - (\sigma_z-\sigma_x)^{\otimes 3}\Big], \end{aligned} \tag{2.111}$$

where $|\overline{\mathrm{GHZ}}\rangle = (|\bar{0}\bar{0}\bar{0}\rangle + |\bar{1}\bar{1}\bar{1}\rangle)/\sqrt{2} = (|000\rangle + |001\rangle + |010\rangle + |100\rangle)$ with $|\bar{0}\rangle = (|0\rangle + i|1\rangle)/\sqrt{2}$ and $|\bar{1}\rangle = -(|0\rangle - i|1\rangle)/\sqrt{2}$. This witness is constructed slightly differently from above, namely here 1/2 is the maximal squared overlap between $|\overline{\mathrm{GHZ}}\rangle$ and any biseparable state. Furthermore, since the maximum overlap between $|\overline{\mathrm{GHZ}}\rangle$ and any W state is 3/4 [98], the operator $\mathcal{W}_{\mathrm{GHZ}} = 3/4\cdot\mathbb{1} - |\overline{\mathrm{GHZ}}\rangle\langle\overline{\mathrm{GHZ}}|$ is a GHZ witness, i.e., it has a negative expectation value for GHZ states, but is positive for states belonging to the class of W states. Therefore we can prove with the witness (2.111) that a state ρ is fully tripartite entangled if $\mathrm{Tr}[\mathcal{W}_{\mathrm{W}}^{(2)}\rho] < 0$. If $\mathrm{Tr}[\mathcal{W}_{\mathrm{W}}^{(2)}\rho] < -1/4$ then the state ρ does not belong to the W state class. Note, however, that the witness cannot prove that a state belongs to the W class. Theoretically, one expects $\mathrm{Tr}[\mathcal{W}_{\mathrm{W}}^{(1)}|\mathrm{W}\rangle\langle\mathrm{W}|] = -1/3$ and $\mathrm{Tr}[\mathcal{W}_{\mathrm{W}}^{(2)}|\mathrm{W}\rangle\langle\mathrm{W}|] = -1/4$.

Let us now proceed with the experimental demonstration. For the experiments the qubits are represented by the polarization of photons, with "0" $\equiv$ horizontal (H) and "1" $\equiv$ vertical (V) linear polarization. The process of spontaneous parametric down-conversion (SPDC) is used to generate the polarization-entangled four photon state of Eq. (2.109) in the arms a_0 and b_0 [109, 110], c.f. Fig. 2.2. In order to transform the initial state into the W state two photon interference at a beam splitter (BS) is employed when the photons are distributed into arms a, b and c [108, 111]. Provided each of the three observers receives one photon they obtain the three photon state $|\mathrm{W}\rangle$. The general principle and the experimental techniques to observe multi photon entangled states are described in detail in [111, 112]; let us here focus on detecting their entanglement.

For implementing the witness observable polarization analyzers (PA) are used. A quarter- (QWP) and a half-wave-plate (HWP) together with a polarizing beam splitter (PBS) allow to set and to analyze any arbitrary polarization direction of

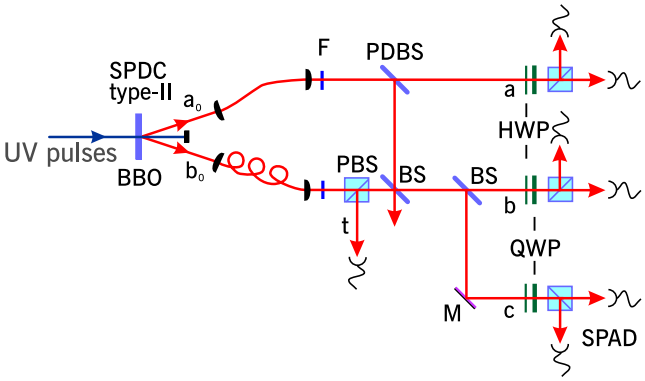

Figure 2.2. *Experimental setup to demonstrate the three-photon entanglement of the W state.*

each of the photons. As the computational basis of the qubit "0"/"1" and thus the spin observable σ_z corresponds to a measurement of the H/V linear polarization, σ_x (σ_y) corresponds to the analysis of $\pm 45°$ linear polarization (left/right circular polarization). Registration of a photon in one of the two detectors of a PA signals the observation of the corresponding eigenstate of the spin operator. For instance, when all three output ports are measured in the H/V linear polarization basis, then the probabilities for the eight possible eigenstates $|z\pm\rangle \otimes |z\pm\rangle \otimes |z\pm\rangle$ are obtained from the events where all three observers register a photon. From these probabilities, the terms like $\mathrm{Tr}[\rho\sigma_z^{\otimes 3}]$ of the expectation value of the entanglement witness can be calculated.

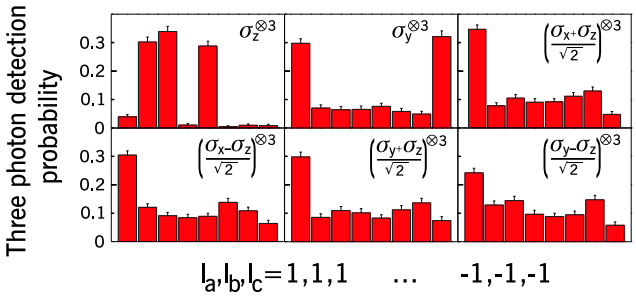

Figure 2.3. *Three photon detection probabilities for six settings of the polarization analyzers as required for the detection of three-photon entanglement using the witness operators* $\mathcal{W}_{\mathrm{W}}^{(1)}$ *and* $\mathcal{W}_{\mathrm{W}}^{(2)}$.

The multi photon detection probabilities for the three-qubit state $|W\rangle$ are shown in Fig. 2.3. From the experimental results the following expectation values can be calculated

$$\mathrm{Tr}[\mathcal{W}_W^{(1)}\rho_W]_{\mathrm{exp}} = -0.197 \pm 0.018, \tag{2.112}$$

$$\mathrm{Tr}[\mathcal{W}_W^{(2)}\rho_W]_{\mathrm{exp}} = -0.139 \pm 0.030. \tag{2.113}$$

This clearly proves with high statistical significance that the observed state is truly tripartite entangled, and the result of the second witness can be interpreted as an indication that the state belongs to the W class. In contrast, the evaluation of a three-photon Bell inequality failed to signify tripartite entanglement for the same experimental settings and noise [111], indicating the superiority of the witnesses as far as entanglement detection is concerned.

2.7 Conclusions

To summarize, in this chapter, we presented the scheme for local detection of entanglement via witness operators, concentrating on bipartite NPPT entangled states as well as PPT entangled states. We presented three different methods for the construction of witnesses, in particular witnesses detecting NPPT states, witnesses excluding biseparability, as well as witnesses detecting PPT entangled edge states. We constructed and decomposed witness for bipartite NPPT entangled states of arbitrary finite dimension, and for bipartite PPT entangled states from the literature. Further, we introduced a new and generalized an existing family of multiqubit PPT entangled states and constructed and decomposed witnesses detecting the latter. Finally, we discussed the experimental implementation of witnesses in the group of H. Weinfurter in Munich.

Further details concerning witnesses detecting multiparticle entanglement, as well as a more general discussion about ways to prove the optimality of a decomposition based on the tensor rank of an operator can be found in Ref. [107]. The results seem to imply that the number of measurement settings of witnesses detecting n-qubit entangled states such as the GHZ state increases exponentially with the number of qubits, as also indicated in section 2.5.4. However, it has been realized that witnesses requiring two locally correlated measurement settings only can be constructed for the multiparty entangled cluster and GHZ states of n qubits [113]. This reduction comes at a price of a slightly reduced robustness to noise.

Such witnesses can in turn be related to Hamiltonians of spin systems, enabling to connect the entanglement properties of states of such systems with macroscopic properties such as energy and temperature [114]. Witnesses have also been applied in the context of quantum cryptography, where it has been shown that the provable presence of entanglement is a necessary condition for the exchange of a secret key [20]. Another interesting relation is the one between Bell inequalities and witness operators. In chapter 5, we will investigate this relation for two Bell inequalities for two-qubit systems.

Finally, we would like to mention two other directions of research that might be interesting. The first concerns the question whether it might be of advantage to use POVMs instead of LvNMs in certain situations. The second is related to the question of how many runs of an experiment have to be performed for a secure decision whether the expectation value of the witness is positive or negative [115], see, for instance, Ref. [116, 117], where this question has been addressed in the context of Bell inequalities.

Chapter 3

Generation and detection of bound entanglement

3.1 Overview

In this chapter, we provide short networks for the experimental generation of two three-qubit families of bound entangled states. The motivation is that even though this special form of entanglement is very interesting from a theoretical point of view, it has not been produced in the laboratory so far.

How does one generate a certain bound entangled state experimentally? A solution that is straightforward from a theoretical point of view is to consider the spectral decomposition of the state and to compose a mixed density matrix by creating the eigenvectors with probabilities that are specified by the according eigenvalues. However, this is, in general, a demanding experimental task, as one would need a source that can emit various types of product vectors and entangled vectors with high fidelities and well-specified probabilities. A more satisfactory approach is to deterministically generate a state that is the purification of the wanted bound entangled state in some higher-dimensional Hilbert space. The additional dimensions are provided by ancilla systems. Then, by tracing out the ancilla (i.e. experimentally simply ignoring the ancilla part), one arrives at the desired bound entangled state.

Here we develop the latter method. Namely, we explicitly construct quantum networks that generate the two families of bound entangled states of three qubits. The first is PPT entangled and has been introduced in Ref. [98], it is the three-qubit family that we generalized to an arbitrary number qubits in section 2.5.4. The second family [60] has a parameter range in which it is NPPT only with respect to one subsystem, which is not sufficient for distillation of a singlet between any two of the parties [60–62]. The properties of the latter states have been used recently in the context of quantum cryptography to show that so-called bound information exists [118]. The networks in both cases act on a six-qubit register that is initially in state $|000000\rangle$, and from which they generate a six-qubit pure state, such that the reduced density operator ρ_{bound} of the first three qubits is the desired bound entangled state.

The network for the family [98] requires only eight two-qubit gates and one Toffoli gate with three control qubits, while the network for the family [60] requires six CNOT gates, one control-U with two control qubits and one Toffoli gate with three control qubits. The number of qubits and number of gates is in foreseeable reach of quantum information technology: at present, with NMR techniques an order-finding algorithm has been performed with 5 qubits [119], as well as the Shor algorithm with 7 qubits [15]. In ion traps, 6 qubits could be provided, and control gates [120] and simple algorithms have been demonstrated [16].

The second step for the experimental generation of bound entangled states is to show that the generated states indeed carry bound entanglement. For the family of bound entangled states in [98] we discuss this issue explicitly. We constructed an appropriate entanglement witness in section 2.5.4 already, where we also found an optimal local decomposition which requires only four measurements settings. Here, we optimize its robustness to noise by varying the families' parameters. Furthermore, this family of states is PPT with respect to any subsystem. For the experimental proof of this fact we briefly discuss three methods: we consider the full state estimation of the produced state ρ_{bound}, the more direct spectrum estimation of $\rho'_{\text{bound}} = \mathcal{A}(\rho_{\text{bound}})$, where $\mathcal{A}$ is the LOCC version of the structural physical approximation to the partial transpose [92, 93], and finally the spectrum estimation of the partial transpose of ρ_{bound} via the LOCC version of the network introduced in [94].

The chapter is organized as follows: In section 3.2 we will introduce the network that generates the class of bound entangled states described in [98]. In section 3.3 we construct a network that generates the family of bound entangled states of Ref. [60]. In section 3.4 we look for family parameters optimizing the robustness to noise of the entanglement witness. Finally, in section 3.5 we briefly discuss the three different approaches to check the positivity of the partial transpositions of the density matrix with respect to any of the three subsystems.

3.2 Generation of PPT entangled states

In this section we explicitly construct the quantum network that generates the following class of bound entangled states [98]:

$$\begin{aligned}\rho_{\text{bound}} &= \frac{1}{N}\Big(2|\text{GHZ}\rangle\langle\text{GHZ}| + \text{a}|001\rangle\langle 001| + \text{b}|010\rangle\langle 010| + \text{c}|011\rangle\langle 011| \\ &\quad + \frac{1}{c}|100\rangle\langle 100| + \frac{1}{b}|101\rangle\langle 101| + \frac{1}{a}|110\rangle\langle 110|\Big), \end{aligned} \tag{3.1}$$

where the coefficients fulfill $a, b, c > 0$ and $ab \neq c$, while the normalization reads $N = 2 + a + b + c + 1/a + 1/b + 1/c$. This state appeared in section 2.5.4 already, where we also constructed an optimal entanglement witness for it.

This mixed state can be generated deterministically by a quantum network that uses a register with three qubits plus three auxiliary qubits, all initialized at $|0\rangle$, and generates a pure states of 6 qubits, such that the reduced density operator of the three qubits of interest is ρ_{bound}.

The procedure to generate the bound entangled state consists of two parts: a preparation stage for the first three qubits, and a purification stage where from the prepared state and an ancilla state a purification of $\rho_{\rm bound}$ is generated. In the preparation stage one starts with the three-qubit state $|000\rangle$, and prepares a three-qubit pure state of the form

$$\begin{aligned} |\psi_{\rm bound}\rangle &= \frac{1}{\sqrt{N}}(|000\rangle + \sqrt{a}|001\rangle + \sqrt{b}|010\rangle + \sqrt{c}|011\rangle \\ &\quad + \frac{1}{\sqrt{c}}|100\rangle + \frac{1}{\sqrt{b}}|101\rangle + \frac{1}{\sqrt{a}}|110\rangle + |111\rangle). \end{aligned} \tag{3.2}$$

This is achieved by applying certain local rotations (LU) on the three qubits, a control-U gate $\mathrm{CU}_{(3,1)}$ between qubit 3 and qubit 1 (qubit 3 acts as the control qubit), a control-U gate $\mathrm{CU}_{(3,2)}$ between qubit 3 and qubit 2 (qubit 3 acts as the control qubit) and a CNOT gate between qubit 1 and qubit 3 (qubit 1 acts as the control qubit). This sequence of gates is illustrated in the left part of Fig. 3.1.

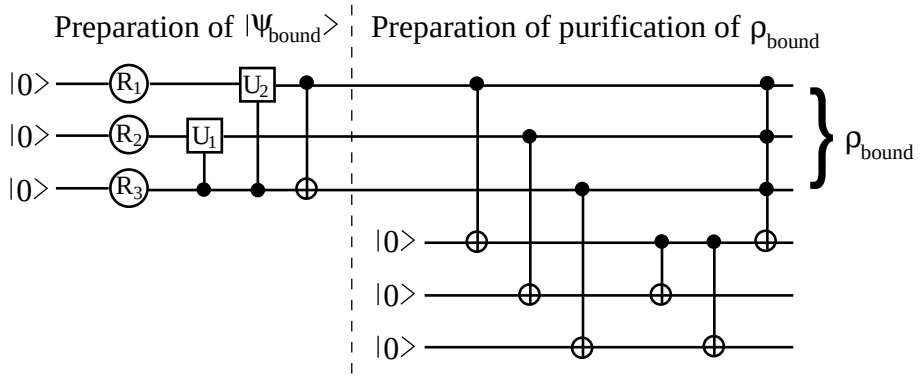

Figure 3.1. *The network for creating the bound entangled state given in Eq. (3.1).*

The specific form of these gates is given by

$$\mathrm{LU} = N_1 \begin{pmatrix} 1 & 1/\sqrt{b} \\ 1/\sqrt{b} & -1 \end{pmatrix}_1 \otimes N_2 \begin{pmatrix} 1 & \sqrt{b} \\ \sqrt{b} & -1 \end{pmatrix}_2 \otimes \begin{pmatrix} \alpha & \beta \\ \beta & -\alpha \end{pmatrix}_3, \tag{3.3}$$

$$\begin{aligned} \mathrm{CU}_{(3,1)} &= \mathbb{1}_{(1,2)} \otimes |0\rangle\langle 0|_3 \\ &\quad + N_1 N_3 \begin{pmatrix} (\sqrt{a} - \sqrt{1/bc}) & (\sqrt{a/b} + \sqrt{1/c}) \\ (\sqrt{a/b} + \sqrt{1/c}) & (-\sqrt{a} + \sqrt{1/bc}) \end{pmatrix}_1 \otimes \mathbb{1}_2 \otimes |1\rangle\langle 1|_3, \end{aligned} \tag{3.4}$$

$$\begin{aligned} \mathrm{CU}_{(3,2)} &= \mathbb{1}_{(1,2)} \otimes |0\rangle\langle 0|_3 \\ &\quad + \mathbb{1}_1 \otimes N_2 N_4 \begin{pmatrix} (1 - \sqrt{bc/a}) & (\sqrt{b} + \sqrt{c/a}) \\ (\sqrt{b} + \sqrt{c/a}) & (-1 + \sqrt{bc/a}) \end{pmatrix}_2 \otimes |1\rangle\langle 1|_3. \end{aligned} \tag{3.5}$$

where $N_1 = \sqrt{b/(1+b)}$, $N_2 = 1/\sqrt{1+b}$, $N_3 = \sqrt{c/(1+ac)}$, and $N_4 = \sqrt{a/(a+c)}$. The coefficients α and β depend on a, b, c and must be chosen such that $\alpha N_1 N_2 = \beta N_3 N_4$ and $\alpha^2 + \beta^2 = 1$.

It is straightforward to confirm that this set of gates is constructed such that it performs the following sequence of transformations:

$$\begin{aligned}
|000\rangle \quad &\stackrel{\mathrm{LU}}{\longrightarrow} \quad N_1\Big(|0\rangle+\frac{1}{\sqrt{b}}|1\rangle\Big)N_2\Big(|0\rangle+\sqrt{b}|1\rangle\Big)\Big(\alpha|0\rangle+\beta|1\rangle\Big)\\
&\stackrel{\mathrm{CU}_{(3,1)}\cdot\mathrm{CU}_{(3,2)}}{\longrightarrow} \quad \frac{1}{\sqrt{N}}\Big[\Big(|0\rangle+\frac{1}{\sqrt{b}}|1\rangle\Big)\Big(|0\rangle+\sqrt{b}|1\rangle\Big)|0\rangle\\
&\qquad\qquad +\Big(\sqrt{a}|0\rangle+\frac{1}{\sqrt{c}}|1\rangle\Big)\Big(|0\rangle+\sqrt{\frac{c}{a}}|1\rangle\Big)|1\rangle\Big]\\
&\stackrel{\mathrm{CNOT}_{(1,3)}}{\longrightarrow} \quad |\psi_{\mathrm{bound}}\rangle.
\end{aligned}$$

In the second part of the network one first applies a sequence of three CNOT gates between the main and the auxiliary qubits: in this way each term of $|\psi_{\mathrm{bound}}\rangle$ is copied to the ancilla system. Here, the first, second, and third qubits of the main system act as control qubits, and the first, second, and third ancilla qubits act as target qubits, respectively:

$$\begin{aligned}
|\psi_{\mathrm{bound}}\rangle|000\rangle \stackrel{3\ \mathrm{CNOTs}}{\longrightarrow} \quad &\frac{1}{\sqrt{N}}(|000\rangle|000\rangle+\sqrt{a}|001\rangle|001\rangle+\sqrt{b}|010\rangle|010\rangle\\
&+\sqrt{c}|011\rangle|011\rangle+\frac{1}{\sqrt{c}}|100\rangle|100\rangle+\frac{1}{\sqrt{c}}|101\rangle|101\rangle\\
&+\frac{1}{\sqrt{a}}|110\rangle|110\rangle+|111\rangle|111\rangle). \qquad (3.6)
\end{aligned}$$

Applying $\mathrm{CNOT}_{(4,5)}$ and $\mathrm{CNOT}_{(4,6)}$ then leads to

$$\begin{aligned}
\longrightarrow \quad &\frac{1}{\sqrt{N}}\Big(|000\rangle|000\rangle+\sqrt{a}|001\rangle|001\rangle+\sqrt{b}|010\rangle|010\rangle+\sqrt{c}|011\rangle|011\rangle\\
&+\frac{1}{\sqrt{c}}|100\rangle|111\rangle+\frac{1}{\sqrt{c}}|101\rangle|110\rangle+\frac{1}{\sqrt{a}}|110\rangle|101\rangle+|111\rangle|100\rangle\Big). \qquad (3.7)
\end{aligned}$$

Finally, one applies a 3-Toffoli gate, where the three system qubits are the control qubits and the first auxiliary qubit is the target. Its action is defined as [18]

$$|a,b,c\rangle|f\rangle \rightarrow |a,b,c\rangle|a\cdot b\cdot c\oplus f\rangle\ . \qquad (3.8)$$

The resulting state of the total system is then

$$\begin{aligned}
|\Psi_{\mathrm{bound}}\rangle \;=\; &\frac{1}{\sqrt{N}}\Big(\big(|000\rangle+|111\rangle\big)|000\rangle+\sqrt{a}|001\rangle|001\rangle+\sqrt{b}|010\rangle|010\rangle \qquad (3.9)\\
&+\sqrt{c}|011\rangle|011\rangle+\frac{1}{\sqrt{c}}|100\rangle|111\rangle+\frac{1}{\sqrt{b}}|101\rangle|110\rangle+\frac{1}{\sqrt{a}}|110\rangle|101\rangle\Big).
\end{aligned}$$

Tracing over the three auxiliary qubits, one obtains that the remaining state of the three system qubits is of the desired form of Eq. (3.1):

$$\begin{aligned}
\mathrm{Tr}_{\mathrm{aux}}\big(|\Psi_{\mathrm{bound}}\rangle\langle\Psi_{\mathrm{bound}}|\big) \;=\; &\frac{1}{N}\Big(2|\mathrm{GHZ}\rangle\langle\mathrm{GHZ}|+\mathrm{a}|001\rangle\langle 001|+\mathrm{b}|010\rangle\langle 010|\\
&+c|011\rangle\langle 011|+\frac{1}{c}|100\rangle\langle 100|+\frac{1}{b}|101\rangle\langle 101|\\
&+\frac{1}{a}|110\rangle\langle 110|\Big)=\rho_{\mathrm{bound}}. \qquad (3.10)
\end{aligned}$$

The total quantum network that generates the bound entangled state ρ_{bound} is shown in Fig. 3.1.

Note that for the generation of this bound entangled state a more general version of the Toffoli gate can also be applied, namely $|a,b,c\rangle|f\rangle \rightarrow \exp[i\theta(a,b,c)]|a,b,c\rangle|a \cdot b \cdot c \oplus f\rangle$, because the extra phases cancel when one traces over the ancilla qubits after the Toffoli gate. This requires less elementary operations than the Toffoli gate [18]. The Toffoli gate with three controls can be decomposed into 13 two-qubit gates [121]. We point out that here we are mainly interested in providing a network for the generation of bound entanglement with a small number of gates, rather than in the optimization of this network, or the decomposition of the necessary gates into elementary single and two-qubit gates. The latter issue is discussed elsewhere in the literature, see, for instance, Ref. [122].

3.3 Generation of the Dür-Cirac-Tarrach states

In this section we will show how to produce the second family experimentally, with a method similar to the one described above. Using the notation from [60], this family is given by:

$$\rho_{\text{DCT}} = \sum_{\sigma=\pm} \lambda_0^\sigma |\Psi_0^\sigma\rangle\langle\Psi_0^\sigma| + \sum_{k=01,10,11} \lambda_k (|\Psi_k^+\rangle\langle\Psi_k^+| + |\Psi_k^-\rangle\langle\Psi_k^-|). \tag{3.11}$$

Here $|\Psi_k^\pm\rangle = \frac{1}{\sqrt{2}}(|k_1 k_2 0\rangle \pm |\bar{k}_1 \bar{k}_2 1\rangle)$, where k_1 and k_2 are the binary digits of k, and $\bar{k}_i$ denotes the flipped k_i (Note that the state $|\Psi_0^+\rangle$ in this notation corresponds to $|\text{GHZ}\rangle$ from above.). The normalization condition reads $\lambda_0^+ + \lambda_0^- + 2(\lambda_{01} + \lambda_{10} + \lambda_{11}) = 1$. With the definitions $\Delta \equiv \lambda_0^+ - \lambda_0^- \geq 0$ and

$$s_k \equiv \begin{cases} 1 \text{ if } \lambda_k < \Delta/2 \\ 0 \text{ if } \lambda_k \geq \Delta/2 \end{cases} \tag{3.12}$$

the following properties of the partial transposes hold [60]:

$$s_{01} = 0 \Leftrightarrow \rho^{T_B} \geq 0, \quad s_{10} = 0 \Leftrightarrow \rho^{T_A} \geq 0, \quad s_{11} = 0 \Leftrightarrow \rho^{T_C} \geq 0. \tag{3.13}$$

A singlet state between two of the parties can be distilled iff the partial transposes with respect to the two parties are negative. For the following choice of the parameters

$$\lambda_0^+ = \frac{1}{3}; \quad \lambda_0^- = \lambda_{10} = 0; \quad \lambda_{01} = \lambda_{11} = \frac{1}{6} \tag{3.14}$$

the corresponding state is inseparable with respect to the splitting $A - (BC)$ but separable with respect to the other two splittings. In particular, the state is separable with respect to the splitting $B - (AC)$, hence no singlet can be destilled between B and A or B and C, and because it has a PPT also with respect to the splitting $C - (AB)$, no singlet can be destilled between C and A. Hence when all parties are separated no singlet can be destilled between any two of them, so the state is bound entangled. However, when it is mixed with two states that are obtained by

cyclic permutation of the parties it turns out that the mixture is inseparable with respect to any partition [62]. These properties were used recently to show that bound information exists and can be activated for distillation [118].

Let us sketch how the states of Eq. (3.11) could be prepared with our scheme. The density matrix is given by

$$\rho_{\mathrm{DCT}} = \begin{pmatrix} \frac{\lambda_0^+ + \lambda_0^-}{2} & 0 & 0 & 0 & 0 & 0 & 0 & \frac{\lambda_0^+ - \lambda_0^-}{2} \\ 0 & \lambda_{11} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \lambda_{01} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \lambda_{10} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \lambda_{10} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \lambda_{01} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \lambda_{11} & 0 \\ \frac{\lambda_0^+ - \lambda_0^-}{2} & 0 & 0 & 0 & 0 & 0 & 0 & \frac{\lambda_0^+ + \lambda_0^-}{2} \end{pmatrix}. \tag{3.15}$$

We start again with the state $|000\rangle$ and produce the pure state

$$\begin{aligned} |\psi_{\mathrm{DCT}}\rangle &= \frac{\gamma}{\sqrt{2}}\big(|000\rangle + |100\rangle\big) + \sqrt{\lambda_{01}}\big(|010\rangle + |110\rangle\big) \\ &\quad + \sqrt{\lambda_{10}}\big(|011\rangle + |111\rangle\big) + \sqrt{\lambda_{11}}\big(|001\rangle + |101\rangle\big), \end{aligned} \tag{3.16}$$

where $\gamma = \sqrt{\lambda_0^+ + \lambda_0^-}$. The state in Eq. (3.16) is reached as follows: Start by a local rotation and a CNOT gate

$$\begin{aligned} |000\rangle \quad \stackrel{\mathrm{LU}_1}{\longrightarrow} \quad |0\rangle\big(\alpha_+|0\rangle + \alpha_-|1\rangle\big)|0\rangle \quad \stackrel{\mathrm{CNOT}_{(2,3)}}{\longrightarrow} \quad |0\rangle\big(\alpha_+|00\rangle + \alpha_-|11\rangle\big) \\ \text{where} \quad \mathrm{LU}_1 = \mathbb{1} \otimes \begin{pmatrix} \alpha_+ & \alpha_- \\ \alpha_- & -\alpha_+ \end{pmatrix} \otimes \mathbb{1}. \end{aligned} \tag{3.17}$$

By proper choice of the coefficients $\alpha_\pm$ we can then reach $|\psi_{\mathrm{DCT}}\rangle$ with 3 local unitaries described below as follows

$$\stackrel{\mathrm{LU}_2}{\longrightarrow} |0\rangle\big(\gamma|00\rangle + \sqrt{2\lambda_{01}}|10\rangle + \sqrt{2\lambda_{10}}|11\rangle + \sqrt{2\lambda_{11}}|01\rangle\big) \stackrel{\mathrm{LU}_3}{\longrightarrow} |\psi_{\mathrm{DCT}}\rangle. \tag{3.18}$$

Hence we have to choose the coefficients and the local unitaries LU_2 such that

$$\begin{aligned} \alpha_+|00\rangle + \alpha_-|11\rangle \quad &\stackrel{\mathrm{LU}_2}{\longrightarrow} \quad \alpha_+|\phi\rangle|\psi\rangle + \alpha_-|\phi^\perp\rangle|\psi^\perp\rangle \\ &= \quad \gamma|00\rangle + \sqrt{2\lambda_{01}}|10\rangle + \sqrt{2\lambda_{10}}|11\rangle + \sqrt{2\lambda_{11}}|01\rangle, \end{aligned} \tag{3.19}$$

i.e. we have to find the Schmidt decomposition of the state on the right hand side (rhs) of the last equation. This state has the decomposition

$$|\varphi\rangle = \sum_{ij} C_{ij}|ij\rangle \quad \text{with} \quad C = \begin{pmatrix} \gamma & \sqrt{2\lambda_{11}} \\ \sqrt{2\lambda_{01}} & \sqrt{2\lambda_{10}} \end{pmatrix}. \tag{3.20}$$

The Schmidt coefficients are the positive square roots of the eigenvalues of $C^T C$, namely

$$\alpha_\pm^2 = \frac{1}{2}\Big(1 \pm \sqrt{1 - 4[\big(\gamma^2 + 2\lambda_{01}\big)\big(2\lambda_{10} + 2\lambda_{11}\big) - \big(\gamma\sqrt{2\lambda_{11}} + 2\sqrt{\lambda_{01}\lambda_{10}}\big)^2]}\Big). \tag{3.21}$$

Then the rotation is given by $\mathrm{LU}_2 = \mathbb{1} \otimes V_2 \otimes U_2$, $U_2 = (|u+\rangle, |u-\rangle)$ and $V_2 = (|v+\rangle, |v-\rangle)$. The vectors $|u\pm\rangle$ can be obtained from $(C^T C - \alpha_\pm \mathbb{1})|u\pm\rangle = 0$ and the vectors $|v\pm\rangle$ from $(CC^T - \alpha_\pm \mathbb{1})|v\pm\rangle = 0$. The last local unitary is given by $\mathrm{LU}_3 = H \otimes \mathbb{1} \otimes \mathbb{1}$, where

$$H = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}, \tag{3.22}$$

the Hadamard gate.

Now we add again three ancilla qubits in the state $|000\rangle$, and by using three CNOT gates the states of the first three qubits are copied to the ancilla qubits as in the previous section. This yields the state

$$\begin{aligned} |\psi_{\mathrm{DCT}}\rangle|000\rangle \quad \overset{3\ \mathrm{CNOT's}}{\longrightarrow} \quad & \Big(\frac{\gamma}{\sqrt{2}}\big(|000\rangle^{\otimes 2} + |100\rangle^{\otimes 2}\big) + \sqrt{\lambda_{01}}\big(|010\rangle^{\otimes 2} + |110\rangle^{\otimes 2}\big) \\ & + \sqrt{\lambda_{10}}\big(|011\rangle^{\otimes 2} + |111\rangle^{\otimes 2}\big) \\ & + \sqrt{\lambda_{11}}\big(|001\rangle^{\otimes 2} + |101\rangle^{\otimes 2}\big)\Big). \end{aligned} \tag{3.23}$$

Then we apply the unitary

$$U = \frac{1}{\gamma} \begin{pmatrix} \sqrt{\lambda_0^-} & \sqrt{\lambda_0^+} \\ \sqrt{\lambda_0^+} & -\sqrt{\lambda_0^-} \end{pmatrix} \tag{3.24}$$

on qubit 4 iff the qubits 2 and 3 are in the state $|00\rangle$. A 2-controlled operation usually acts when both control qubits are in the state $|1\rangle$, but this can be changed by flipping the control qubits before and after the gate. This operations leads to the state

$$\begin{aligned} & \frac{1}{\sqrt{2}}|000\rangle\big(\sqrt{\lambda_0^-}|0\rangle + \sqrt{\lambda_0^+}|1\rangle\big)|00\rangle + \frac{1}{\sqrt{2}}|100\rangle\big(\sqrt{\lambda_0^+}|0\rangle - \sqrt{\lambda_0^-}|1\rangle\big)|00\rangle \\ & + \sqrt{\lambda_{01}}\big(|010\rangle^{\otimes 2} + |110\rangle^{\otimes 2}\big) + \sqrt{\lambda_{10}}\big(|011\rangle^{\otimes 2} + |111\rangle^{\otimes 2}\big) \\ & + \sqrt{\lambda_{11}}\big(|001\rangle^{\otimes 2} + |101\rangle^{\otimes 2}\big) \end{aligned} \tag{3.25}$$

Then a 3-Toffoli gate flips qubit 4 iff the first three qubits are in the state $|100\rangle$. Finally two CNOT gates flip qubits 2 and 3 iff the first qubits' state is $|1\rangle$. Tracing out the ancilla particles then yields ρ_{DCT}. Summarizing, the procedure is

$$\begin{aligned} |\psi_{\mathrm{DCT}}\rangle|000\rangle \quad \overset{3\ \mathrm{CNOT's},^2\mathrm{CU},^3\mathrm{Toffoli}}{\longrightarrow} \quad & \sqrt{\frac{\lambda_0^-}{2}}\big(|000\rangle - |100\rangle\big)|000\rangle \\ & + \sqrt{\frac{\lambda_0^+}{2}}\big(|000\rangle + |100\rangle\big)|100\rangle + \dots \\ \overset{\mathrm{CNOT}_{1,2},\ \mathrm{CNOT}_{1,3}}{\longrightarrow} \quad & \Big(\sqrt{\lambda_0^-}|\mathrm{GHZ}^-\rangle|000\rangle + \sqrt{\lambda_0^+}|\mathrm{GHZ}\rangle|100\rangle \\ & + \sqrt{\lambda_{01}}(|010\rangle^{\otimes 2} + |101\rangle|110\rangle) \\ & + \sqrt{\lambda_{10}}(|011\rangle^{\otimes 2} + |100\rangle|111\rangle) \\ & + \sqrt{\lambda_{11}}(|001\rangle^{\otimes 2} + |110\rangle|101\rangle)\Big) \\ & \equiv |\Psi_{\mathrm{DCT}}\rangle \end{aligned}$$

which leads to

$$\mathrm{Tr}_{4,5,6}|\Psi_{\mathrm{DCT}}\rangle\langle\Psi_{\mathrm{DCT}}| = \rho_{\mathrm{DCT}}. \tag{3.26}$$

The complete network is shown in Fig. 3.2. The existence of bound entanglement for the choice of parameters in Eq. (3.14) can be proved by showing that the state has a PPT with respect to two subsystems, but not with respect to the third. This can be proved experimentally by applying the methods of section 3.5 below.

Note that the method works for a general choice of the parameters for which the rank of the density matrix is full.

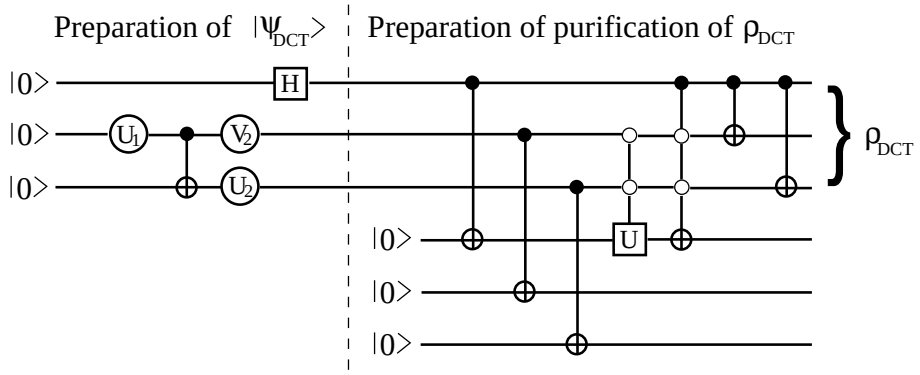

Figure 3.2. *The network for creating the bound entangled state given in Eq. (3.11). Open circles for the control bits indicate that the corresponding gate acts non-trivially on the target if the control is 0, rather than 1 as usually (filled circles).*

3.4 Preparation of the entanglement witness

An entanglement witness and its optimal local decomposition have been obtained in section 2.5.4 already. The witness with the decomposition of Eq. (2.97) requires only 4 local measurements. On the other hand, if state tomography is applied to confirm the positivity of the partial transposes (cf. section 3.5), then all measurements necessary for the witness with the first decomposition of Eq. (2.97) are already performed there.

The last step on the construction of our witness is the computation of the value of ϵ. We use the parametrization $|e\rangle = \cos\theta_e|0\rangle + \exp i\phi_e \sin\theta_e|1\rangle$ and accordingly for $|f\rangle$ and $|g\rangle$. This leads to

$$\begin{aligned}\epsilon \;=\; & \inf_{|e,f,g\rangle}\Big[\frac{1}{2}\Big((c_e c_f c_g)^2 + (s_e s_f s_g)^2\Big) + \frac{1}{1+c^2}\Big(c^2(s_e c_f c_g)^2 + (c_e s_f s_g)^2\Big) \\ & + \frac{1}{1+b^2}\Big((c_e s_f c_g)^2 + b^2(s_e c_f s_g)^2\Big) + \frac{1}{1+a^2}\Big((c_e c_f s_g)^2 + a^2(s_e s_f c_g)^2\Big) \\ & - \Big[\frac{1}{2} + \frac{c}{1+c^2} + \frac{b}{1+b^2} + \frac{a}{1+a^2}\Big]\Big(2\cos(\phi_e+\phi_f+\phi_g)c_e c_f c_g s_e s_f s_g\Big)\Big],\end{aligned} \tag{3.27}$$

where $c_{e,f,g} \equiv \cos\theta_{e,f,g}$ and $s_{e,f,g} \equiv \sin\theta_{e,f,g}$. In this equation the phases ϕ_e, ϕ_f, ϕ_g appear only in the term $\cos(\phi_e+\phi_f+\phi_g)$. Therefore the phases can be chosen to be equal to zero, using the following argument: the term $2\cos(\phi_e+\phi_f+\phi_g)c_e c_f c_g s_e s_f s_g$ in the above equation has to have a positive sign in order to minimize ϵ. As the coefficients $c_{e,f,g}$ and $s_{e,f,g}$ occur only quadratically in all other terms, all of them can be chosen to be positive. Then ϵ is minimized for $\phi_e = \phi_f = \phi_g = 0$. We are thus left with 6 real parameters. If the parameters a, b, c are determined by the experimental set-up, then the corresponding value of ϵ can be obtained numerically by use of a multivariable minimization routine [106].

If the parameters a, b, c can be chosen freely, then it is advantageous to maximize ϵ with respect to a, b, c. Making the natural assumption that white noise is introduced in the preparation procedure of the state, *i.e.* $\rho_p = p\rho_{\rm B} + \frac{1-p}{8}\mathbb{1}$, the witness will detect entanglement in the state for $p > 1-2\epsilon$. Hence the tolerance of the witness to the presence of noise is enlarged by maximizing ϵ. We searched for the maximum [106] in the parameter range $a = b = 1/c \in]0,1[$ (Remember from the definition of $\rho_{\rm bound}$ in Eq. (3.1) that one has to use the open interval here.). We obtain numerically that for $a < a_{\rm th} = 0.3460$ the minimum is reached at $\epsilon = a^2/(1+a^2)$, i.e., when the product state is one of the three possibilities $|e,f,g\rangle = |011\rangle, |101\rangle, |110\rangle$. For $a > a_{\rm th}$ the minimum of ϵ is obtained when $\theta_e = \theta_f = \theta_g$. These results are shown in Fig. 3.3. We find $\epsilon^{\max}_{a=b=1/c} \approx 0.1069$ which is reached for $a_{\rm th} \approx 0.3460$. This is also the highest value obtained numerically when $a, b, 1/c \in]0,1[$ without the restriction $a = b = 1/c$. For this choice of parameters the state mixed with white noise as described above is still detected for $p > 0.786$, i.e., more than 20% of white noise can be tolerated.

As mentioned already in section 2.2.3, the problem with the minimization routines is that it cannot be guaranteed that the minimum ϵ is indeed the global minimum for fixed a, b, and c. In chapter 4, we discuss ways to obtain numerical lower bounds for this situation.

3.5 Testing the positivity of the partial transpose

In this section we briefly discuss three different methods to check the positivity of the partially transposed density operator $\rho^{T_X}_{\rm bound}$ with respect to subsystem $X = A, B, C$. One possible option is to perform the full state estimation of $\rho_{\rm bound}$, and then to check whether all the eigenvalues of $\rho^{T_X}_{\rm bound}$ for $X = A, B, C$ are positive. This method requires the estimation of $(2\times2\times2)^2-1 = 63$ independent parameters of the density operator. This can be achieved by performing $3\times3\times3 = 27$ measurements on single copies of $\rho_{\rm bound}$, since one can write any three qubit state as

$$\rho \;=\; \frac{1}{8}\sum_{i,j,k=0,x,y,z} \lambda_{i,j,k}\sigma_i\otimes\sigma_j\otimes\sigma_k, \tag{3.28}$$

where $\lambda_{i,j,k} = tr(\rho\sigma_i\otimes\sigma_j\otimes\sigma_k)$, and $\sigma_0 = \mathbb{1}$. As also discussed in section 2.3, the data from estimating $\lambda_{l,m,n}$ with $l, m, n = x, y, z$ can also be used to estimate

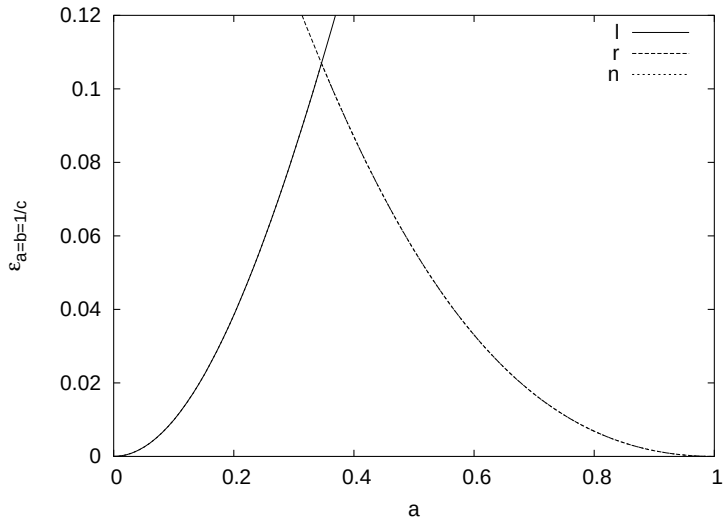

Figure 3.3. *Bounds for $\epsilon_{a=b=1/c}$ as a function of a: The left curve (l) is given by $a^2/(1+a^2)$, while the right curve (r) is the analytic minimum of ϵ for $\theta_e = \theta_f = \theta_g$. The maximal value $\epsilon^{\max} \approx 0.1069$ is obtained for $a \approx 0.3460$, where the two curves meet. The result of the numerical minimization (n) is plotted on top of the two analytical curves, and equals the lower branch of them for all a.*

$\lambda_{0,m,n}$, $\lambda_{l,0,n}$ and $\lambda_{l,m,0}$. Hence, only local measurements in the x, y, z directions have to be performed. One disadvantage of this option is the superfluous estimation of parameters of the density operator, since we are only interested in learning about the lowest eigenvalue of the partially transposed density operator.

Another method for finding out whether $\rho_{\text{bound}}^{T_X} > 0$ for $X = A, B, C$ is to start by applying the structural physical approximation (SPA) [92] to the partial transpose of ρ_{bound}, and then to estimate the lowest eigenvalue of the resulting density operator. From this one can infer whether or not the original state, partially transposed, is positive. This procedure has to be performed for the three possible partitions.

A structural physical approximation is a completely positive (CP) map, constructed from a positive, but not CP map, by adding white noise. The aim in constructing these approximations is to allow the physical implementation of maps which are useful in entanglement detection, but are non-physical. In this way one is able to bypass full state estimation when trying to detect the existence of entanglement in a given system, since one can estimate directly the relevant parameters, which is the lowest eigenvalue in this case.

The SPA to the partial transposition with respect to party C can be easily obtained from the SPA for two particles [IX], and reads

$$[\mathbb{1} \otimes \widetilde{\mathbb{1} \otimes T}](\rho) \quad = \quad \frac{1}{3}\mathbb{1} \otimes \Lambda_1 \otimes \Lambda_2 + \frac{2}{3}\mathbb{1} \otimes \mathbb{1} \otimes \sigma_x\sigma_z\Lambda_1\sigma_z\sigma_x, \tag{3.29}$$

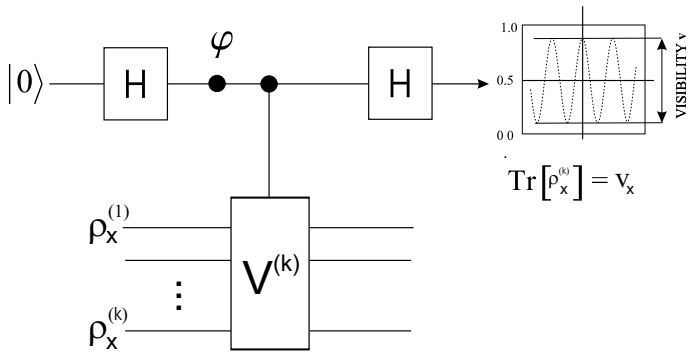

Figure 3.4. *Quantum network that each of the parties $X = A, B, C$ has to perform for the estimation of the non-linear functionals by LOCC.*

where $\Lambda_1(\rho) = 1/3 \sum_{i=x,y,z} \sigma_i \rho \sigma_i$, and $\Lambda_2(\rho) = 1/4 \sum_{i=0,x,y,z} \sigma_i \rho \sigma_i$. Note that the map $[\widetilde{\mathbb{1} \otimes \mathbb{1} \otimes T}]$ can be implemented using only LOCC. Now $\rho^{T_C} \geq 0$ if $[\widetilde{\mathbb{1} \otimes \mathbb{1} \otimes T}](\rho) \geq 2/9$ [92]. The SPA to the partial transposition with respect to A and B can be obtained from Eq. (3.29) by a permutation of the parties.

In order to estimate the eigenspectrum by LOCC of a state $\tilde{\rho}$ obtained from ρ by an application of a SPA, the parties first prepare k copies of $\tilde{\rho}$, $k = 2, \ldots 8$. Then each party performs the interferometric network of Fig. 3.4 that works on a control qubit coupled to the respective qubits of k copies of the state. If the state of the control qubit is $|1\rangle$, then $V^{(k)}$ performs the shift operation on the state qubits,

$$V^{(k)}|\phi_1\rangle|\phi_2\rangle...|\phi_k\rangle = |\phi_k\rangle|\phi_1\rangle...|\phi_{k-1}\rangle, \tag{3.30}$$

$\forall|\phi_i\rangle$, $i = 1, ..., k$, which can be implemented by a concatenation of swap gates between two qubits. By varying φ, an interference pattern can be observed in the measurement on the control qubit in the $\{|0\rangle, |1\rangle\}$ basis after the completion of the network. The visibility is given by $v_X = \mathrm{Tr}[\rho_X^k]$, where ρ_X is the reduced state of party $X = A, B, C$. From these visibilities the parties can infer $\mathrm{Tr}[\rho^k]$ [IX]. If they perform the network for $k = 2, \ldots 8$, they can estimate the spectrum of ρ [92].

Finally, there is the option of directly estimating the non-linear functionals $\mathrm{Tr}[(\rho^{T_X})^k]$ with $k = 1, 2, 3, ...$ and $X = A, B, C$, following [94]. This scheme is a modification of the scheme presented in [92], and can be also implemented using only LOCC [93]. The main difference between the quantum network of [94], when compared with [92], is that the $C - V^{(k)}$ gates acting on the different subsystems do not all shift in the same direction. The one with respect to which the partial transposition is supposed to be done has to shift in the opposite direction.

In conclusion, for the estimation of the density matrix via tomography, 27 locally correlated measurements are necessary on single copies of the state. On the other hand, in both of the other procedures, $3 \times 7 = 21$ measurements are necessary, 7 for each partition. Each of these measurement requires a network acting on up to 25

qubits. In the last proposal, the network gives an estimate directly, whereas in the former, a SPA has to be performed before.

3.6 Conclusions

To summarize, we have presented a quantum network that generates bound entangled states of three qubits. Explicitly, we have studied the production of the two families of bound entangled states that were introduced in [98] and [60]. Note that our method could be adapted in a straightforward way to the generation of other types of bound entangled states. As our networks consists of six qubits and several two-qubit gates, they go beyond present quantum information processing technology – however, it seems feasible to realize them in the not too distant future.

We also discussed different methods of testing whether the produced states generated by the network are indeed bound entangled. Namely, we suggested to detect the entanglement via a suitable witness operator, and to confirm positivity of the partial transposes by either full state estimation, or spectrum estimation of the structural physical approximation of the partial transpose, or direct estimation of some non-linear functionals.

Chapter 4

Non-convex optimization problems

4.1 Overview

In this chapter, we show that many problems occuring in entanglement theory can be formulated in such a way that recent known results from non-convex optimization theory can be applied to efficiently approximate the solutions.

Many problems occuring in the field are in fact convex problems. To state whether a state is separable or not is equivalent to stating whether a state is in the convex hull of product states. Also, the evaluation of many measures of entanglement essentially require the solution of a convex problem. This is why it has been increasingly realized in recent years [36–38, 123–128] that the solution of many problems of quantum information theory can be found or approximated with the help of the field of research that is primarily concerned with questions of this type: the theory of convex optimization [129]. Many problems from quantum information theory can easily be translated to the language of convex optimization theory. Examples include the evaluation of measures of entanglement that reasonably quantify the degree of entanglement of a given state, such as the distillable entanglement or the asymptotic relative entropy of entanglement [123, 124].

Also, it has been realized that while the complete solution of the question of separability is a NP-HARD problem in the system size [127], one can nevertheless find hierarchies of sufficient criteria for entanglement in the bi-partite setting. In each step, by solving an efficiently solvable convex optimization problem, one finds an answer to the problem in the form (i) one can assert that the state is entangled, or (ii) one cannot assert it, and has to go one (computationally more expensive) step further [36]. Even though this method provably detects every entangled state after some step [37], it will never stop for separable states. This problem was addressed recently in Ref. [130], where an algorithm is introduced which is designed to prove a given bipartite quantum state to be separable in a finite number of steps. It is based on the search for a decomposition via a countable subset of product states, which is dense within all product states.

The problem of testing for multi-partite entanglement has been related to robust semi-definite programming and a hierarchy of relaxations in Ref. [38], and the method of Ref. [36] has also been extended to the multipartite setting [39].

This chapter is concerned with a link of the theory of entanglement to the theory of convex optimization in a similar spirit. The central observation here is very simple yet potentially very useful: many problems related to entanglement can be cast into the form of optimization problems with polynomial constraints of degree three. This includes the question whether a state is entangled or not, notably not only in the bi-partite, but also for the several separability classes of the multi-partite setting. Then, the construction of non-decomposable witnesses introduced in section 2.2.3 involves a problem of this kind, as well as the evaluation of the geometric measure of entanglement to quantify multi-partite entanglement. This structure is due to the fact that in all these instances, one essentially minimizes over product state vectors of a multi-partite quantum system.

This polynomial part of the optimization problems is still non-convex and computationally expensive to solve. Yet, applying results from relaxation theory of non-convex problems [131–135], notably the method of Lasserre [133], we find hierarchies of solutions to our original problems, and each step is a better approximation than the previous one. Each step itself amounts to solving an efficiently implementable semi-definite program [129]. Moreover, the hierarchy is asymptotically complete, in the sense that the exact solution is asympotically attained. The increase of the size of the vector of objective variables of these semi-definite problems grows notably polynomially in the label of the hierarchy.

We will first give a short introduction to semi-definite programming. Then, we will state how one can introduce auxiliary variables to cast the considered problems from entanglement theory into the desired form. In the following sections, we will introduce the hierarchies of relaxations in detail, and study numerical examples. In particular, we find that the bounds on ϵ for the entanglement witness of section 3.4 are in excellent agreement with those found with the method introduced in this chapter.

4.2 Problems in entanglement theory as optimization problems

At the core of the problems we discuss in this chapter are minimizations over product vectors. Given a Hermitean operator W, we seek the minimum of

$$\mathrm{Tr}[|\psi_1\rangle\langle\psi_1| \otimes ... \otimes |\psi_n\rangle\langle\psi_n| W], \tag{4.1}$$

where the minimum is taken with respect to product state vectors of a composite quantum systems with parts labeled $1, \ldots, n$, in a Hilbert space $\mathcal{H} = \mathcal{H}_1 \otimes \ldots \otimes \mathcal{H}_n$. For instance, this is exactly the problem we faced in section 2.2.3 in the construction of PPT entanglement witnesses.

One way of solving this problem is to choose a specific basis for the Hilbert space and to explicitly parametrize the state vectors. This yields a polynomial in these parameters, in general of very high order, which is obviously not a convex problem in these variables: a solution can be found, albeit not in an efficient manner. For

small systems, algorithms such as simulated annealing may be employed, delivering upper bounds to the optimal solution, as no control is possible as to what extent one is far away from the global optimum.

The strategy we adopt here is in instances of the above type to introduce additional variables, giving rise to one vector $x \in \mathbb{R}^t$, $x^T = (x_1, ..., x_t)$, which is the objective variable, parametrizing the product states. Then we cast the problem into the form of a linear objective function, simply as

$$\text{minimize } c^T x \tag{4.2}$$

with a (fixed) $c \in \mathbb{R}^t$, subject to constraints which are polynomials in the objective variables:

$$g_l(x) \geq 0, \tag{4.3}$$

$l = 1, ..., L$, where $g_l : \mathbb{R}^t \to \mathbb{R}$ are real polynomials of a degree up to three.

Then, we apply recently found known results from non-convex relaxation theory [133], giving rise to a sequence of approximations to the problem above in the following form:

$$\begin{aligned} \text{minimize} \quad & d^T y, \\ \text{subject to} \quad & F^{[h]}(y) \geq 0, \\ & G_l^{[h]}(y) \geq 0,\ l = 1, ..., L \end{aligned} \tag{4.4}$$

with matrices $F^{[h]}(y)$ and $G_l^{[h]}(y)$ that are linear in the elements of y. For each l, the matrix $G_l^{[h]}(y)$ depends on the coefficients of the polynomial $g_l(x)$ from Eq. (4.3). The optimization problem is effectively turned into a problem of a larger vector y, which is growing in dimension with step size h, as well as the matrices $F^{[h]}(y)$ and $G_l^{[h]}(y)$. Hence the semi-definite program of the first step, which is denoted by $h_{\min}$ and grows with the maximal degree of the polynomial constraints, comes at the lowest computational cost. The objective function stays the same, but is uplifted, namely

$$y \longmapsto d^T y, \tag{4.5}$$

where

$$d^T = (0, c_1, ..., c_t, 0, ..., 0), \tag{4.6}$$

with $c \in \mathbb{R}^t$ being defined as above. The constraints are transformed into so-called semi-definite constraints in the matrices, because these are constrained to be positive semi-definite.

In fact, by transforming the constraints, the set over which the minimization is performed is effectively enlarged. In subsequent steps of the hierarchy, the constraints get more and more restrictive, until the convex hull of the set over which the minimization of the original problem is performed is reached. Hence in each step a lower bound to the solution of the original problem is obtained, and the bounds get better in subsequent steps. Furthermore, it can be shown that the approximations converge to the real solution asymptotically [133].

Optimization problems exhibiting a linear objective function and semi-definite constraints are called semi-definite programs [129]. Such instances of convex optimization problems can be efficiently solved, for example by means of interior-point methods [129]. As mentioned above, many problems in quantum information theory are easily written down in the form of a semi-definite program [36, 125]. In fact, it may be argued that to specify the solution of a problem in form of a semi-definite program has the same status as stating a result in terms of the spectrum of a matrix, as this again merely means that efficient methods are available to find the eigenvalues of a given matrix.

Hence by casting the problem into the form of a minimization of a linear objective function subject to polynomial constraints, and by applying the method of Ref. [133], we arrive at a hierarchy of efficiently solvable semi-definite programs approximating the solution of the problem from below with increasing accuracy, and the real solution is obtained asymptotically.

Before we show how the problem of Eq. (4.1) can be formulated as a problem involving a linear objective function subject to polynomial constraints, we would like to mention some useful facts and clarify further some of the notions introduced above.

4.2.1 Polynomial constraints, Lagrange duality, and relaxations

First, in order to make clear why polynomially constrained problems are hard to solve in general, let us consider as an example quadratic polynomial constraints which are of the form

$$x^T A_l x + b_l^T x + c_l \leq 0, \tag{4.7}$$

$l = 1, ..., L$. The matrices A_l are, however, not necessarily positive semi-definite. This is by no means a minor detail: if all matrices $A_1, ..., A_L$ were positive matrices, $A_l \geq 0$, then the constraints of Eq. (4.7) would be convex. This would yield a convex quadratic program, which can be efficiently solved. In fact, these are also instances of semi-definite programs, the constraint

$$(Ax + b)^T (Ax + b) - c^T x - d \leq 0 \tag{4.8}$$

is equivalent to[1]

$$\begin{pmatrix} \mathbb{1} & Ax + b \\ (Ax + b)^T & c^T x + d \end{pmatrix} \geq 0. \tag{4.9}$$

In contrast, if the matrices are not all positive semi-definite, one obtains a very hard, non-convex optimization problem. This structure is yet dictated by the problems from quantum information theory at hand.

[1] Given two real, symmetric, and positive matrices A and B of dimension $n \times n$ and $m \times m$, respectively, and a real matrix C of dimension $n \times m$, we define

$$M = \begin{pmatrix} A & C \\ C^T & B \end{pmatrix},$$

which is a symmetric matrix of dimension $(n + m) \times (n + m)$. Assuming that the kernel of B is a subset of the kernel of C, then M is positive iff the *Schur complement* $A - CB^{-1}C^T$ is positive. Here B^{-1} is the pseudoinverse of B, i.e., the inverse performed on the range of B [24].

Coming back to semi-definite programs, a very useful fact is that for any *primal* semi-definite program, in its most general form being given by

$$\begin{array}{cc} \text{minimize} & c^T x, \\ \text{subject to} & F(x) = F_0 + \sum_{s=1}^{t} x_s F_s \geq 0, \end{array} \tag{4.10}$$

one can formulate the Lagrange-dual problem, which is again a semi-definite problem. It is given by

$$\begin{array}{cc} \text{maximize} & -\mathrm{Tr}[F_0 Z], \\ \text{subject to} & \mathrm{Tr}[F_s Z] = c_s,\ s = 1, ..., t, \\ & Z \geq 0. \end{array} \tag{4.11}$$

Vectors x fulfilling the primal semi-definite constraint are called *primal feasible*, while matrices Z fulfilling the dual constraints are called *dual feasible*. For a primal feasible x and a dual feasible Z we obtain

$$c^T x + \mathrm{Tr} Z F_0 = \sum_{i=1}^{t} Z F_i x_i + \mathrm{Tr} Z F_0 = \mathrm{Tr} Z F(x) \geq 0, \tag{4.12}$$

where we used the fact that $\mathrm{Tr} AB \geq 0$ if both A and B are positive operators. From this it follows that

$$-\mathrm{Tr} F_0 Z \leq c^T x, \tag{4.13}$$

so any solution of the dual problem is a lower bound to the optimal solution of the primal problem. This is referred to as weak duality. Under certain conditions, in particular, if there is a solution x satisfying $F(x) > 0$, the optimal values of the dual and the primal problem are identical. This case, which is rather the typical one, is denoted as strong duality. The idea of Lagrange duality is a powerful tool to formulate rigorous lower bounds to solutions of optimization problems.

Finally, let us make clear what the notion of a relaxation means that appeared before. The idea of a relaxation is to introduce new variables and to formulate the problem as a convex problem in a larger space. This idea can be exemplified in the simplest form of a relaxation, the Shor relaxation [131]. For example, let A_1 in Ineq. (4.7) be a matrix which is not positive semi-definite, and let us assume that $b_1 = 0$ and $c_1 = 0$ for simplicity. Then, one can still write the constraint equivalently as

$$\mathrm{Tr}[X A_1] \leq 0,\ X = x x^T, \tag{4.14}$$

using a $t \times t$ symmetric matrix X. The equality $X = x x^T$ is equivalent to the constraints $X \leq x x^T$ and $X \geq x x^T$. The latter is a convex constraint and equivalent to

$$\begin{pmatrix} X & x \\ x^T & 1 \end{pmatrix} \geq 0, \tag{4.15}$$

while $X \leq x x^T$ is a non-convex constraint. Shor's relaxation amounts to taking only the convex part into account, thereby delivering an efficiently solvable convex problem which yields a lower bound to the original problem [131], because the set over which the minimization is performed is effectively enlarged. Such relaxations in terms of semi-definite constraints are employed in Lasserre's method, yet instead of one many such relaxations forming a complete hierarchy.

4.2.2 Polynomial constraints for product states

Now we will show that the encountered optimization problems can be written as polynomially constrained problems of a degree of at most three. That this is possible is based on the following observation: An operator O can be constrained to a projector onto a pure state $|\psi\rangle\langle\psi|$ by requiring

$$\mathrm{Tr}[O] = \mathrm{Tr}[O^2] = 1 \text{ and } O \geq 0. \tag{4.16}$$

But these constraints are in fact equivalent to the following ones

$$\mathrm{Tr}[O^2] = \mathrm{Tr}[O^3] = 1, \tag{4.17}$$

see also Ref. [136]. Hence an Hermitean operator can be constrained to a pure state by two polynomial constraints of degree two and three, respectively.

This follows from the fact that, the only decreasingly ordered vector of eigenvalues λ of O consistent with

$$\sum_i \lambda_i^2 = 1 \text{ and } \sum_i \lambda_i^3 = 1 \tag{4.18}$$

is the vector $\lambda = (1, 0, ..., 0)$. As the eigenvalues, as well as those of O^2 and O^3, are unitarily invariant, the above statements can be shown to be valid on the level of probability distributions. Essentially, $\sum_i \lambda_i^2 = 1$ already requires all absolute values of eigenvalues to be smaller than or equal to 1, such that the only ordered vector of real numbers consistent with $\sum_i \lambda_i^3 = 1$ becomes $(1, 0, ..., 0)$.

For a qubit system, the constraints can further be simplified by merely requiring as constraints $\mathrm{Tr}[O] = 1$, $\mathrm{Tr}[O^2] = 1$, as for Hermitean 2×2 matrices these conditions alone imply that $O = |\psi\rangle\langle\psi|$.

When applied to our specific problems at hand, these constraints will appear in the following form. We will require that Hermitean matrices P are, except from normalization, products of pure states with respect to all constituents. This will be incorporated as follows: Denoting with $I = \{1, ..., n\}$ the index set labeling the subsystems and with $\mathrm{Tr}_{I\setminus j}$ the partial trace with respect to all systems except the one with label j, the lines

$$\mathrm{Tr}[\mathrm{Tr}_{I\setminus j}[P]^2] = 1 \tag{4.19}$$

$$\mathrm{Tr}[\mathrm{Tr}_{I\setminus j}[P]^3] = 1 \tag{4.20}$$

for all $j \in I$ indeed enforce that all the reductions are pure states. If all reductions are pure, the global state must be a pure product state. This can be seen as follows. For states ρ, the only possibility for

$$\mathrm{Tr}[\mathrm{Tr}_{I\setminus j}[\rho]^2] = 1 \text{ and } \mathrm{Tr}[\mathrm{Tr}_{I\setminus j}[\rho]^3] = 1 \tag{4.21}$$

to hold for all $j \in I$ is that ρ is of the form of product pure states,

$$\rho = |\phi_1\rangle\langle\phi_1| \otimes ... \otimes |\phi_n\rangle\langle\phi_n|. \tag{4.22}$$

The constraints can be reduced to constraints of the degrees one and two for systems consisting only of qubits.

Having stated the general strategy, let us now look at the specific instances of problems in quantum information we will be considering in this chapter.

4.2.3 Non-decomposable witnesses

Problems of the type of the one in Eq. (4.1) appear in the construction of PPT entanglement witnesses [28], as we have seen in section 2.2.3. Further, this method can also be used to obtain a finer witness from a given one $\tilde{W}$, i.e., a witness that detects the same states as $\tilde{W}$ – and more. If $\tilde{\varepsilon} = \min_{|a,b\rangle} \mathrm{Tr}[|a,b\rangle\langle a,b|\tilde{W}] > 0$ then $\tilde{W} - \tilde{\varepsilon}\mathbb{1}$ is a finer witness than $\tilde{W}$. This is of use even if the local measurements available are fixed, because the observable $\mathbb{1}$ does not require a measurement.

This renders the method useful in the context of quantum cryptography. Witnesses are of practical interest here, since it has been shown that the provable presence of quantum correlations in such protocols is a necessary precondition for secure key distillation [20]. Furthermore, the set of local measurements available in a particular implementation of a quantum key distribution (QKD) scheme is naturally restricted by the protocol. In order to deliver the entanglement proof, it is sufficient to obtain one relevant entanglement witness as a first step towards the demonstration of the feasibility of the scheme, because this witness can then be optimized with the method presented above.

For a given entanglement witness, we want to solve the following optimization problem

$$\begin{aligned} \text{minimize} \quad & \mathrm{Tr}[WP] \\ \text{subject to} \quad & P \text{ is a projector onto a product state.} \end{aligned} \tag{4.23}$$

The task is to write this problem in terms of a polynomially constrained problem. Then we can approximate the problem by a hierarchy of semi-definite programs as mentioned above. In each step, we obtain a lower bound to the minimal expectation of W with respect to product vectors. If this lower bound is positive, then it is already possible to construct a proper PPT witness. Higher relaxations might improve the witness further.

Using the insights from the last section, we turn the constraint on P into the following polynomial constraints

$$\begin{aligned} \text{minimize} \quad & x, \\ \text{subject to} \quad & x \geq \mathrm{Tr}[WP], \\ & \mathrm{Tr}[\mathrm{Tr}_{I\setminus j}[P]^2] = 1, \text{ for all } j \in I, \\ & \mathrm{Tr}[\mathrm{Tr}_{I\setminus j}[P]^3] = 1, \text{ for all } j \in I. \end{aligned} \tag{4.24}$$

For multi-party qubit systems this can be written as a polynomially constrained problem with polynomials of degree two by simply replacing the last constraint by the linear constraint $\mathrm{Tr}[P] = 1$.

Finally, we would like to remark that one may trade in the constraint of degree three for a linear constraint and a semidefinite constraint, so that the problem takes the form

$$\begin{aligned} \text{minimize} \quad & x, \\ \text{subject to} \quad & x \geq \mathrm{Tr}[WP], \\ & \mathrm{Tr}[P] = 1, \\ & \mathrm{Tr}[\mathrm{Tr}_{I\setminus j}[P]^2] = 1, \text{ for all } j \in I, \\ & P \geq 0, \end{aligned} \tag{4.25}$$

and look for the intersection of the feasible sets of the semi-definite part and the constraint set of the relaxations. Again, a positive number leads to a proper witness, but in this case asymptotic convergence to the optimum cannot be guaranteed. This change in constraints is also possible in the other examples, but we do not treat it explicitly there.

Parametrization

In an implementation of this optimization problem, one has to choose a basis of Hermitean matrices for each Hilbert space,

$$\{\sigma_1, ..., \sigma_{d_j^2}\}, \tag{4.26}$$

for $j = 1, ..., N$, suppressing an additional index labelling the subsystems. The basis can be chosen such that the Hermitean matrices satisfy $\mathrm{Tr}[\sigma_1] = 1$ and

$$\mathrm{Tr}[\sigma_k] = 0, \ k = 2, ..., d_j^2, \tag{4.27}$$

and have a Hilbert-Schmidt scalar product

$$\mathrm{Tr}[\sigma_k \sigma_l] = \xi_{d_j} \delta_{kl} \tag{4.28}$$

with a dimension dependent constant ξ_{d_j} (and similarly for terms of third order). For the case of qubit subsystems, the appropriately normalized familiar Pauli matrices can be taken. In terms of this basis of Hermitean matrices, the matrix P can be written as

$$P = \sum_{\kappa=(k_1,...,k_N)} p_\kappa \Sigma_\kappa, \tag{4.29}$$

where $\kappa = (k_1, ..., k_N)$, is a multi-index, with $k_j = 1, ..., d_j^2$ for $j \in I$, and

$$\Sigma_\kappa = \sigma_{k_1} \otimes \sigma_{k_2} \otimes ... \otimes \sigma_{k_N}. \tag{4.30}$$

If the number of parties is small, it might be useful to reduce the number of variables at the expense of an increase of the lowest relaxation step $h_{\min}$ as follows: The tensor p_κ can be directly assumed to be of rank one, $p_\kappa = \Pi_{i=1}^N a_{k_i}^i$. In this case, the rhs of Eq. (4.29) is a product of the single particle density matrices

$$\rho_i = \sum_{j=0}^{d_i^2-1} a_j^i \sigma_j. \tag{4.31}$$

These correspond to pure states if the constraints $\mathrm{Tr}[\rho_i^2] = \mathrm{Tr}[\rho_i^3] = 1$ are fulfilled. The number of variables is reduced from $\Pi_{i=1}^N d_i^2$ to $\sum_{i=0}^N d_i^2$, while the degree of the polynomial $\mathrm{Tr}[WP]$ is increased from 1 to N, which increases $h_{\min}$, as will become clear in section 4.3 below. This is the parametrization that will be used for the numerical examples.

Before we present the hierarchy of relaxations explicitly, we discuss the other applications which are similar in structure from the point of view taken here.

4.2.4 Estimating the geometric entanglement to quantify multi-particle entanglement

The same tools can be used in order to quantify multi-particle entanglement for pure quantum states. Needless to say, the question of quantifying multi-particle entanglement is much more involved that the analogous question in the bi-partite setting.

One of the reasonable quantities to quantify multi-particle entanglement is the geometric measure of entanglement [48, 49, 138]: for a given state vector $|\psi\rangle \in \mathcal{H}_1 \otimes ... \otimes \mathcal{H}_N$, essentially, entanglement is then quantified in terms of the solution of the maximization problem

$$\Lambda^2 = \max_{|\phi\rangle \in S} |\langle\psi|\phi\rangle|^2, \tag{4.32}$$

where S is the class of fully separable states, such that the geometric measure of entanglement becomes

$$E(|\psi\rangle\langle\psi|) = 1 - \Lambda^2. \tag{4.33}$$

Setting $\rho = |\psi\rangle\langle\psi|$ and $P = |\phi\rangle\langle\phi| = |\phi_1\rangle\langle\phi_1| \otimes ... \otimes |\phi_N\rangle\langle\phi_N|$, we arrive at

$$\begin{aligned} \text{minimize} \quad & x, \\ \text{subject to} \quad & \mathrm{Tr}[P\rho] + x \geq 1, \\ & \mathrm{Tr}[\mathrm{Tr}_{I\setminus j}[P]^2] = 1, \text{ for all } j \in I, \\ & \mathrm{Tr}[\mathrm{Tr}_{I\setminus j}[P]^3] = 1, \text{ for all } j \in I, \end{aligned} \tag{4.34}$$

which is the same optimization as in the previous subsection, except from the first line in the list of constraints, which enables to write the problem such that objective variable x corresponds to E from Eq. (4.33).

It is even possible to go further and perform optimizations over all separable *mixed* states, hereby constructing sufficient criteria for entanglement.

4.2.5 Tests for bi-partite and multi-partite entanglement

The approach is here to consider for a given state $\rho \in \mathcal{S}(\mathcal{H}_1 \otimes ... \otimes \mathcal{H}_N)$ the minimal Hilbert-Schmidt norm with respect to the set of separable states. For simplicity of

notation, we explicitly formulate the optimization problem for the instance of full separability, without loss of generality. That is, we test whether ρ can be written as

$$\rho = \sum_{i=1}^{n} p_i \rho_1^{(i)} \otimes ... \otimes \rho_N^{(i)}, \tag{4.35}$$

with $\{p_i\}_i$ forming a probability distribution. The question whether a state is fully separable is hence equivalent to asking whether a state is an element of the convex hull of product vectors with respect to all subsystems. According to Caratheodory's theorem [137], for any k-dimensional subset $S \subset \mathbb{R}^m$, any point of the convex hull of S can be written as a convex combination of at most $k+1$ points from S. Hence, the number of elements in the convex combination given by Eq. (4.35) can be restricted to $n = \prod_{j=1}^{N} d_j^2$, again without loss of generality. However, for two qubits it is known that any separable state can be written as a mixture of only 4 projectors onto product vectors instead of 16 [99], so that the bound from Caratheodory can probably be undercut in general.

To decide whether a state ρ is fully separable or not, we may solve the following optimization problem,

$$\begin{aligned} \text{minimize} \quad & \|\rho - P\|_2^2 = \mathrm{Tr}[(\rho - P)^2], \\ \text{subject to} \quad & P \text{ is fully separable.} \end{aligned} \tag{4.36}$$

We make use of the Hilbert-Schmidt norm as it is quadratic in the matrix entries.

When written as a polynomially constrained problem, each relaxation gives a lower bound of the Hilbert Schmidt distance to the set of fully separable states. Hence, asserting that the state is not fully separable whenever we obtain a value larger than the one that we accept as accuracy of the computation[2], each step delivers a sufficient criterion for multi-partite entanglement in its own right, and the hierarchy is complete in the sense that each entangled state is detected by some step. The

[2]This notion can be sharpened by employing the notion of weak membership [41]. This can be phrased as follows: we denote for any convex set $S \subset \mathbb{Q}^m$ and any rational $\delta > 0$ with $B(S, \delta)$ the set of all $x \in \mathbb{Q}^m$ for which there exists a $y \in S$ such that

$$\|x - y\|_2 \leq \delta, \tag{4.37}$$

and with $B(S, -\delta)$ the set of all $x \in S$ for which $y \in S$ for all $y \in \mathbb{Q}^m$ with $\|x - y\|_2 \leq \delta$. So clearly, S is a strict subset of $B(S, \delta)$, and $B(S, -\delta)$ is a strict subset of S. The weak membership problem allows for two alternatives: given a rational element $x \in \mathbb{Q}^m$ and a rational number $\delta > 0$ either (i) assert that $x \in B(S, \delta)$, or (ii) assert that $x \neq B(S, -\delta)$.

associated optimization problem can now be written as

$$\begin{aligned} \text{minimize} \quad & x, \\ \text{subject to} \quad & x \geq \mathrm{Tr}(\rho - P)^2, \\ & P - \sum_{i=1}^{n} P^{(i)} = 0, \\ & \mathrm{Tr}[\mathrm{Tr}_{I\setminus j}[P^{(i)}]^2] = (\mathrm{Tr}[P^{(i)}])^2, \\ & \qquad \text{for all } i = 1, ..., n,\ j \in I. \\ & \mathrm{Tr}[\mathrm{Tr}_{I\setminus j}[P^{(i)}]^3] = (\mathrm{Tr}[P^{(i)}])^3, \\ & \qquad \text{for all } i = 1, ..., n,\ j \in I. \\ & \sum_{i=1}^{n} \mathrm{Tr}[P^{(i)}] = 1. \end{aligned} \tag{4.38}$$

This is a global optimization problem with polynomial constraints of degree three. Again, for qubits, polynomial constraints of degree two are sufficient. In the parametrization, P is a sum of projectors onto product states with a weight different from 1. If $|\psi\rangle\langle\psi|$ is a normalized projector, then $\alpha|\psi\rangle\langle\psi|$ fulfills

$$\mathrm{Tr}[(\alpha|\psi\rangle\langle\psi|)^2] = (\mathrm{Tr}[\alpha|\psi\rangle\langle\psi|])^2 \text{ and } \mathrm{Tr}[(\alpha|\psi\rangle\langle\psi|)^3] = (\mathrm{Tr}[\alpha|\psi\rangle\langle\psi|])^3. \tag{4.39}$$

This explains the terms of the rhs of the lines restricting the $P^{(i)}$. Further, the weights have to add up to one in order to ensure that $\mathrm{Tr}[P] = 1$, which is the reason of the last line.

Here one tests the hypothesis that the state is fully separable against the alternative that the state is entangled in some sense. To assert that the state is multi-particle entangled and not separable with respect to any separability class, several tests are hence required. In this way, the various classes of genuine multi-particle entanglement can be detected. Note that even when applied to the bi-partite case, the resulting hierarchy of semi-definite relaxations is inequivalent to the one in Ref. [36, 37], and also inequivalent to the robust semi-definite programming approach in Refs. [38]. The above formulation in the optimization problem in terms of full separability still does not constitute a restriction of generality, as this includes all separability classes with respect to all possible splits.

Alternatively to the above approach, one may write each test in the form of a *feasibility problem*, a problem with a vanishing objective function,

$$\begin{aligned} \text{minimize} \quad & 0, \\ \text{subject to} \quad & \rho \text{ satisfies the test of step} \\ & \quad h = h_{\min}, h_{\min+1}, ... \text{ in the hierarchy.} \end{aligned} \tag{4.40}$$

Either one finds no solution (which is to say, the problem is not primal feasible), and one can assert that the state is not fully separable, or one has to go on one step in the hierarchy. In each step of the hierarchy, forming a semi-definite problem, the dual problem can then be employed to prove the infeasibility of the above primal

problem serving as a certificate [129] (see also Ref. [36]). Any feasible solution of the dual problem with $-\mathrm{Tr}[ZF_0] > 0$ proves the infeasibility of the primal (original) problem, because the dual problem gives lower bounds on the primal problem. That is, we can use the dual problem to prove properties of our original problem at hand.

4.3 Complete hierarchies of relaxations to approximate the solutions

In this section, we will define the hierarchy of relaxations introduced in section 4.2. This method is based on recent results in real algebraic geometry, see also Ref. [135].

Before defining the matrices appearing in the semi-definite relaxations, cf. Eq. (4.4), we need to introduce some notation. Even though the highest degree of the occuring contraints in the problems discussed above is three, it will be convenient to formulate the sequence of semi-definite programs in terms that formally involve higher-order polynomials.

For any $r \in \mathbb{N}$, we consider the basis of polynomials of degree r in the variables $x_1, ..., x_t$ as

$$(1; x_1, ..., x_t; x_1^2, x_1x_2, ..., x_1x_t; x_2^2, x_2x_3, ..., x_t^r), \tag{4.41}$$

in this ordering. The dimension of this basis will be denoted as D_r. We drop the index t, as this will stay the same throughout the procedure. Any polynomial of degree of at most r can then be identified with a vector $p \in \mathbb{R}^{D_r}$. It is convenient to introduce two labelings, connected with each other by a function

$$f_r : \{1, ..., D_r\} \to \left\{\alpha = (\alpha_1, ..., \alpha_t) : \sum_{s=1}^{t} \alpha_s \leq r\right\}, \tag{4.42}$$

such that the i-th element, $i = 1, ..., D_r$, of the basis given by Eq. (4.41) is written as

$$\prod_{i=1}^{t} x_i^{\alpha_i}, \tag{4.43}$$

characterized by $\alpha = (\alpha_1, ..., \alpha_t) \in \mathbb{N}_0^t$. Note that for a given $k \in \mathbb{N}$ there are $\binom{t+k-1}{k}$ possible vectors α such that $\sum_{s=1}^{t} \alpha_s = k$. It follows that the dimensions D_h are given by

$$D_h = \sum_{k=0}^{h} \binom{t+k-1}{k}. \tag{4.44}$$

In the following we give the required matrices from Lasserre's method for general polynomials [133] and discuss the cases occuring in the paper explicitly afterwards. Let δ_l be the degree of the polynomial constraint $l \in \{1, ..., L\}$ and $\lceil \delta_l/2 \rceil$ be the smallest integer greater than or equal to $\delta_l/2$. We assume that the objective function is linear, which is no restriction of generality, as other polynomials can always be incorporated in the constraints as in section 4.2.5. Then the first possible relaxation

step of Lasserre's method is $h_{\min} = \max_l \lceil \delta_l/2 \rceil$. For $h \geq h_{\min}$ the matrix $F^{[h]}(y)$ is of dimension $D_h \times D_h$ and linear in a vector $y \in \mathbb{R}^{D_{2h}}$,

$$\left[F^{[h]}(y)\right]_{i,j} = y_{f_{2h}^{-1}(f_h(i)+f_h(j))}. \tag{4.45}$$

For instance,

$$F^{[1]}(y) = \begin{pmatrix} y_{00} & y_{10} & y_{01} \\ y_{10} & y_{20} & y_{11} \\ y_{01} & y_{11} & y_{02} \end{pmatrix}. \tag{4.46}$$

In turn, the matrices $G_l^{[h]}(y)$, one for each of the constraint polynomials, $l = 1, ..., L$, are of dimension $D_{\tilde{h}_l} \times D_{\tilde{h}_l}$, where $\tilde{h}_l = h - \lceil \delta_l/2 \rceil$. Each polynomial g_l is characterized according to the above procedure by a vector v_l. The matrices $G_l^{[h]}(y)$ are then defined as

$$\left[G_l^{[h]}(y)\right]_{i,j} = \sum_\alpha v_{f_{\delta_l}^{-1}(\alpha)} y_{(f^{-1}_{\delta_l+2\tilde{h}_l}(f_{\tilde{h}_l}(i)+f_{\tilde{h}_l}(j))+\alpha)}. \tag{4.47}$$

For example, let the degree of the polynomial constraint g_l be given by $\delta_l = 2$ and let $h = 1$, so that $\tilde{h}_l = 0$. Then, the matrix has only a single entry which is given by

$$G_l^{[1]}(y) = \sum_{i=1}^{D_2} v_i y_i = g_l, \tag{4.48}$$

so that g_l is recovered. In the next relaxation step, the matrix would be of the size $D_1 \times D_1 = 3 \times 3$.

For qubits, $h_{\min} = 1$, because the maximal degree of the constraint polynomials is 2. For higher dimensional systems, the highest occuring order is 3 due to the positivity constraints. In this case, $h_{\min} = 2$.

Convergence to the optimum

In Ref. [133] convergence to the solution of the original problem, cf. Eqs. (4.2) and (4.3), in the limit $h \to \infty$ is guaranteed if there exist polynomials, $u_0, u_1, ..., u_L$, all sums of squares, such that the set

$$\{x \in \mathbb{R}^t : u_0(x) + \sum_{l=1}^{L} u_l(x) g_l(x) \geq 0\} \tag{4.49}$$

is compact. This is, however, the case in all of the specific situations from entanglement theory considered above. The set in Eq. (4.49) is compact if there exists an $l \in \{1, ..., L\}$ such that the set

$$\{x \in \mathbb{R}^t : g_l(x) \geq 0\} \tag{4.50}$$

is compact. In each of the discussed cases, we find that due to the linear constraints incorporating the trace requirement and the quadratic constraints coming from the purity of the reduced states, there exists an $a > 0$ such that $a^2 - \|x\|^2 \geq 0$ for all

$x \in \mathcal{M}$. This follows from the fact that for each of the involved matrices, the trace is bounded from above, and positivity of the matrices enforces boundedness of all elements. Hence, to ensure asymptotic completeness, we may add the constraint $g_{L+1}(x) = a^2 - \|x\|^2 \geq 0$ to the list of quadratic constraints, such that the condition in Eq. (4.50) is certainly satisfied. Hence, one can conclude that

$$\min_{y \in \mathcal{M}^{[h]}} d^T y \to \min_{x \in \mathcal{M}} c^T x \tag{4.51}$$

for $h \to \infty$, and for

$$\begin{aligned} \mathcal{M}^{[h]} &= \{y \in \mathbb{R}^{D_{2h}} : F^{[h]}(y) \geq 0, \\ &\qquad G_l^{[h]}(y) \geq 0,\, l = 1, ..., L+1\}. \end{aligned} \tag{4.52}$$

This is not only meant as a numerical procedure: instead, in each step a semi-definite program is given explicitly, which can also be assessed with analytic means in principle. Moreover, symmetries of the involved states under certain groups can be carried over to symmetries in the Hermitean matrices in the semi-definite programs, similarly to the strategy employed in Ref. [124] for semi-definite programs, and in Ref. [123] for convex but not semi-definite programs.

After having discussed the asymptotic converges, let us see why the convergence cannot be guaranteed if the problem is formulated with both semi-definite as well as polynomial constraints. In section 4.2.3, we saw that it is possible to trade in extra semi-definite constraints for a lower maximal degree of the polynomial constraints, see Eq. (4.25). In this case, we may either express the semi-definite constraint with polynomials of higher order. This is possible because a Hermitean matrix is positive iff the determinants of all its submatrices are positive (see also Ref. [24]). However, this will most probably spoil the benefit of having reduced the maximal degree of the polynomial constraints in the first place, and might only be of interest in situations where the problem cannot be formulated involving only polynomial constraints, a situation occuring, e.g., in an entanglement test for continuous variable systems [X].

Another possibility is to combine the semi-definite relaxations with the semi-definite constraint itself. This gives rise to a hierarchy of sufficient tests, without the property of asymptotic completeness. To see how they can be combined, let us consider an additional semi-definite constraint. In terms of the $y \in \mathbb{R}^{D_{2h}}$, we have the feasible set of the additional semi-definite constraint

$$\mathcal{F} = \{y \in \mathbb{R}^{D_{2h}} : F_0 + \sum_{s=1}^{t} y_{s+1} F_s \geq 0\}, \tag{4.53}$$

with appropriate matrices $F_0, ..., F_t$. Therefore, we can write the full hierarchy of semi-definite programs as

$$\begin{aligned} \text{minimize} \quad & d^T y, \\ \text{subject to} \quad & F^{[h]}(y) \geq 0, \\ & G_l^{[h]}(y) \geq 0,\, l = 1, ..., L \\ & F_0 + \sum_{s=1}^{t} y_{s+1} F_s \geq 0, \end{aligned} \tag{4.54}$$

$h = h_{\min}, h_{\min+1}, ...$ being the label of the element of the hierarchy. The projection of the feasible sets $\mathcal{M}^{[h]}$ onto the plane of first order moments, i.e., onto the plane

$$\{y \in \mathbb{R}^{D_{2h}} : y = (0, y_2, ..., y_{t+1}, 0, ..., 0)\}, \tag{4.55}$$

conceived as a subset of $\mathbb{R}^t$, converges (pointwise) to the convex hull of $\mathcal{M}$ [132–134]. Therefore, we have that

$$\min_{y \in \mathcal{M}^{[h]} \cap \mathcal{F}} d^T y \leq p^* \tag{4.56}$$

for all $h \to \infty$. Moreover, $\min_{y \in \mathcal{M}^{[h]} \cap \mathcal{F}} d^T y$ is a monotone increasing sequence in h, such that the sufficient criteria become more powerful with an increasing order of the hierarchy.

Size of the relaxations

A relevant issue is how large the semi-definite relaxations are in each step of the hierarchy. The matrix $F^{[h]}$ is of dimension $D_h \times D_h$, with D_h given by Eq. (4.44) For example,

$$D_2 = 1 + t + \frac{t(t+1)}{2}. \tag{4.57}$$

In the number of variables t, this is a manifestly polynomial expression. In step h the vector y is of the length D_{2h}. Notably, in each of the steps, the effort of a numerical solution of the associated semi-definite program is polynomial in the dimension of the matrices [129]. Hence, each problem can be solved in an efficient manner.

In terms of the step h in the hierarchy, it turns out that the scaling is also polynomial. Approximating the above sum by an integral expression, we arrive at

$$D_h = O(h^t). \tag{4.58}$$

That is, for a fixed number of variables (which is the setting considered here), the size of the vector of the objective variables increases also only polynomially in the step h in the hierarchy.

4.4 Numerical Examples

In order to show that the approach is also feasible in practice, we present in this section three numerical examples, two for the geometric measure of entanglement and one for the construction of entanglement witnesses for bound entangled three-qubit states.

4.4.1 Geometric measure for three-qubit states

Let us start with the calculation of the geometric measure of entanglement for three-qubit states. As we have shown in section 4.2 the computation of the geometric measure of entanglement for a given pure three-qubit state vector $|\psi\rangle$ requires essentially the calculation of

$$\Lambda^2 = \max_{|a,b,c\rangle} |\langle a,b,c|\psi\rangle|^2 \tag{4.59}$$

We use here the second parametrization described in section 4.2.3. In terms of the Pauli matrices forming a basis of Hermitean matrices, we can write

$$|\psi\rangle\langle\psi| = \frac{1}{8}\sum_{i,j,k=0}^{3} \lambda_{ijk}(\sigma_i \otimes \sigma_j \otimes \sigma_k), \tag{4.60}$$

$$|a,b,c\rangle\langle a,b,c| = \frac{1}{8}\sum_{i,j,k=0}^{3} a_i b_j c_k(\sigma_i \otimes \sigma_j \otimes \sigma_k), \tag{4.61}$$

where $\lambda_{000} = a_0 = b_0 = c_0 = 1$. The coefficients λ_{ijk}, $i,j,k = 0,...,3$, are determined from the known state vector $|\psi\rangle$.

We have to impose constraints that guarantee that ρ_A is a pure state on the coefficients (a_1, a_2, a_3) describing the state $\rho_A = 1/2\sum_{i=0}^{3} a_i\sigma_i$ (and similarly (b_1, b_2, b_3) and (c_1, c_2, c_3)). We have seen before that for qubit systems, instead of requiring $\mathrm{Tr}[\rho_A^2] = 1$ and $\mathrm{Tr}[\rho_A^3] = 1$, we may alternatively merely require that $\mathrm{Tr}[\rho_A] = 1$ and $\mathrm{Tr}[\rho_A^2] = 1$ (where $\mathrm{Tr}[\rho_A] = 1$ is already a consequence of the parametrization).

So we arrive at the optimization problem

$$\begin{aligned} \text{maximize}_{a_i,b_j,c_k} \quad & \frac{1}{8}\sum_{i,j,k=0}^{3} \lambda_{ijk} a_i b_j c_k, \\ \text{subject to} \quad & a_0 = b_0 = c_0 = 1 \\ & a_1^2 + a_2^2 + a_3^2 = 1, \\ & b_1^2 + b_2^2 + b_3^2 = 1, \\ & c_1^2 + c_2^2 + c_3^2 = 1. \end{aligned} \tag{4.62}$$

This polynomial optimization problem can be solved with the help of Lasserre's method, see section 4.3. For the numerical calculations we used the freely available package GloptiPoly [139] which is based on SeDuMi [140]. The package GloptiPoly has a number of desirable features, in particular, it provides a certificate for global optimality.

First, we present a nontrivial example for the calculation of the geometric measure of entanglement, in a case where its value is already known. In this way we can test our methods. We aim at computing the geometric measure of entanglement for state vectors of the form

$$|\psi(s)\rangle = \sqrt{s}|W\rangle + \sqrt{1-s}|\tilde{W}\rangle, \tag{4.63}$$

$s \in [0,1]$, where $|W\rangle$ and $|\tilde{W}\rangle$ are state vectors of three-qubit W states in different bases,

$$|W\rangle = \frac{1}{\sqrt{3}}\Big(|001\rangle + |010\rangle + |100\rangle\Big), \tag{4.64}$$

$$|\tilde{W}\rangle = \frac{1}{\sqrt{3}}\Big(|011\rangle + |101\rangle + |110\rangle\Big). \tag{4.65}$$

For the geometric measure of entanglement of $|\psi(s)\rangle$ a formula has been developed in Ref. [49], exploiting the permutation symmetry of the states. The comparison between the theoretical value and the numerical calculation using Lasserre's method for $h = 2$ is shown in Fig. 4.1. Details of the performance are summarized in Table 4.1. Note that $h_{\min} = 2$ here because of the parametrization that we used. The results indicate clearly the usefulness of the presented approach. As a matter of fact, this is a case where already a very small number of steps in the hierarchy detects the global optimum, as is typical for this method [139].

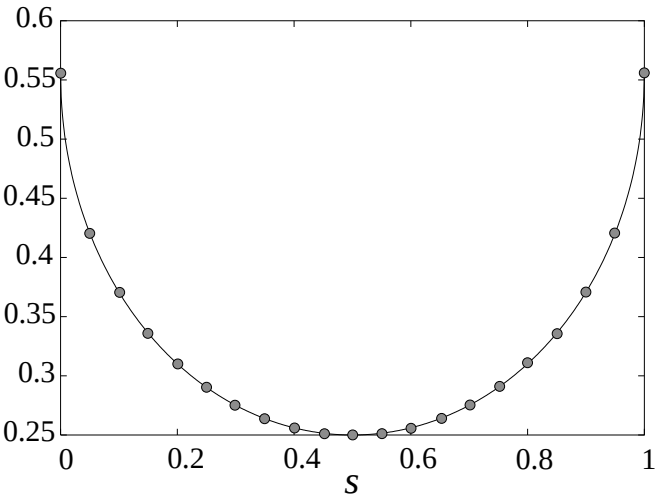

Figure 4.1. *The numerical values of the geometric measure of entanglement E of the family of states of Eq. (4.63), plotted on top of the analytical values of Ref. [49].*

4.4.2 Geometric measure for 4-qubit states

We calculate the geometric measure of entanglement also for the following one parameter family of state vectors

$$|\psi_4(p)\rangle = \sqrt{p}|\mathrm{GHZ}'\rangle - \sqrt{1-p}\,|\psi^+\rangle \otimes |\psi^+\rangle, \tag{4.66}$$

where

$$|\mathrm{GHZ}'\rangle = (|0011\rangle + |1100\rangle)/\sqrt{2}, \tag{4.67}$$

Subsection	Relaxation h	# variables	dim(y)	CPU time
4.3.1	2	9	714	10.92 s
4.3.2	2	12	1819	103.97 s
4.3.3	2	9	714	6.14 s

Table 4.1. *Details of the relaxations in the three numerical examples discussed above for one point of each example. The provided CPU time refers to a machine with a Intel Xeon Processor, 2.2 GHz, 1GB Ram, using GloptiPoly 2.2e [139], SeDuMi 1.05 [140], and MatLab 6.5.1.199709 (release 13). In all cases $h = h_{\min} = 2$, so that the result was obtained after the first relaxation step.*

$|\psi^+\rangle = (|01\rangle + |10\rangle)/\sqrt{2}$, and $p \in [0, 1]$. The state vector $|\psi_4(2/3)\rangle$ corresponds to the 4-qubit singlet state, i.e., the state vector satisfying

$$U^{\otimes 4}|\psi\rangle = |\psi\rangle \tag{4.68}$$

for all unitary U [141]. For the two individual states in the above superpositions in Eq. (4.66), the geometric measure can be directly evaluated [49]: For $p = 1$ we find $\Lambda^2 = 1/2$, and for $p = 0$ we obtain $\Lambda^2 = 1/4$ from $\Lambda^2_{\psi^+} = 1/2$. The numerical results for the geometric measure of entanglement for other values of p are plotted in Fig. 4.2. It is interesting to note that at the singlet value $p = 2/3$, the behavior

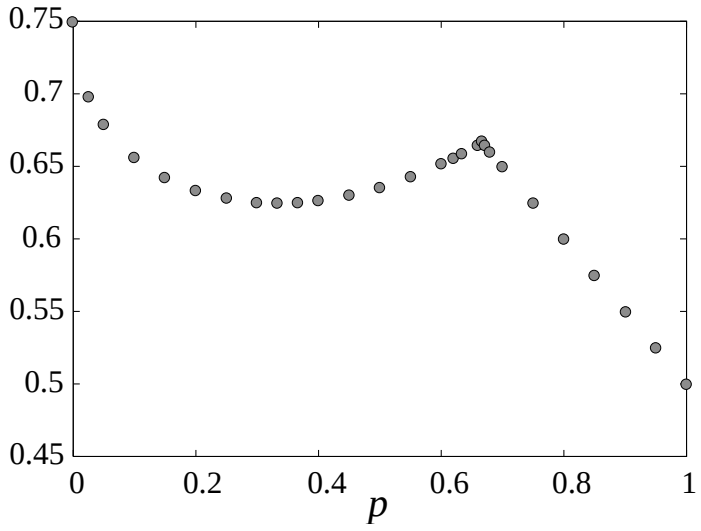

Figure 4.2. *The numerical values of the geometric measure of entanglement E of the family of states of Eq. (4.66).*

of the geometric measure changes. From there up to $p = 1$ the optimum is attained for the choices $|0011\rangle$ or $|1100\rangle$ of the product state which gives rise to the linear behavior.

The family of states specified in Eq. (4.66) is invariant under the exchange $(AB) \leftrightarrow (CD)$. Because of this symmetry, one may without loss of generality assume that the product state vector leading to the maximal value of Λ^2 is given by $|\phi_1, \phi_2, \phi_1, \phi_2\rangle$, where $|\phi_{1,2}\rangle = e^{i\chi_{1,2}} \cos\theta_{1,2}|0\rangle + e^{i\eta_{1,2}} \sin\theta_{1,2}|1\rangle$, where the optimal phases can be shown to be all vanishing. This gives rises to an optimization problem with polynomial constraints with only two variables which can be solved exactly by GloptiPoly. The results coincide with the results above.

4.4.3 Witness for 3-qubit PPT entangled states

Employing the same strategy, we would like to calculate the value of ϵ as defined in section 4.2 for the family of witnesses constructed for the three qubit PPT entangled states introduced in section 2.5.4. In section 3.4, we obtained upper bounds for the values of ϵ by using a multi-variable minimization routine [106] for the parameter range $a = b = 1/c \in]0, 1[$, which are plotted in Fig. 3.3. The minimization has to be performed with respect to W from Eq. (2.100) without the term $\epsilon\mathbb{1}$, and the substitution $c \leftrightarrow 1/c$ has to be done. The numerical results obtained with the methods from this section are plotted in Fig. 4.3 on top of the former results. Again, the global optimum is achieved, and the values agree with the values found in 3.4. For details concerning the relaxations, see Table 4.1.

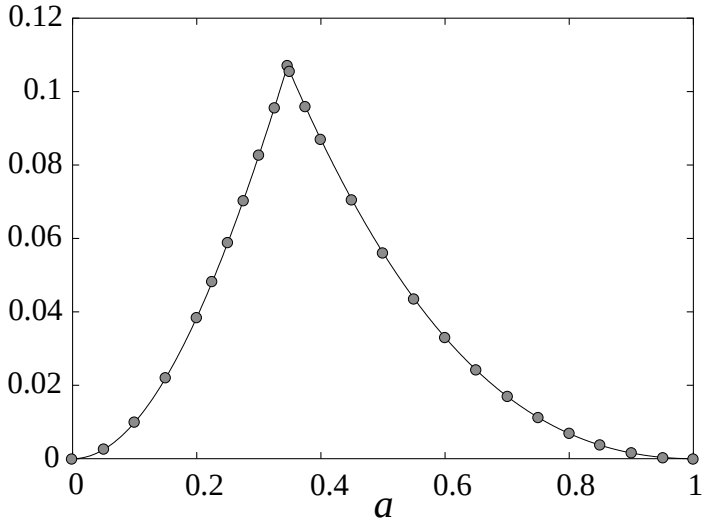

Figure 4.3. *The numerical values of ϵ for $\bar{W} = W + \epsilon\mathbb{1}$, with W given by Eq. (2.100), plotted on top of the results from Fig 3.3.*

4.5 Conclusions

In this chapter, we have revisited several problems in entanglement theory with the tools and language of convex optimization. The central point was that many problems, where a minimization over pure product vectors is required, can be written as instances of certain optimization problems involving polynomial constraints of degree two or three, or with additional semi-definite constraints. For such polynomially constrained problems, which are generally instances of non-convex optimization problems, hierarchies of semi-definite relaxations can be found. One arrives at hierarchies of more and more refined tests detecting entangled or separable states, or better and better lower bounds to optimization problems.

In all instances, recently achieved known results from semi-algebraic geometry guarantee that asymptotically, the achieved minimum is indeed approaching the globally optimal one. In this sense, the statements are similar in spirit with yet more versatile than the ones presented in Refs. [36, 37, 39]. Moreover, we have seen that the size of the optimization problems to be solved in each test grows polynomially with the steps in the hierarchy, and that for small problems, often only a small number of steps is required to find the exact solution.

The presented method is on the one hand meant as a numerical method to achieve good bounds to problems that are of relevance in the study of multi-particle entanglement, in the construction of entanglement witnesses in the bi-partite and multi-partite case, and in the context of quantum key distribution. Other application include the construction of entanglement witnesses based on second moments and the assessment of maximal output purities, see Ref. [X]. On the other hand, each instance of the hierarchy delivers a semi-definite program which is readily accessible with analytical methods, and where properties of the Lagrange-dual can be exploited. It is hoped that these techniques shed new light on the structure of optimization problems underlying the questions of entanglement and separability of several constituents.

Chapter 5

Entanglement witnesses vs. Bell inequalities

5.1 Overview

In this chapter, we turn our attention again to witness operators and investigate their relation to the CHSH inequality introduced in chapter 1. Many facts about Bell inequalities and witness operators have been named already in the first two chapters. Let us recall those which are most important in the context of this chapter.

First of all, Bell inequalities correspond to witness operators in the framework of quantum mechanics. The relation between the two entities was first studied in Ref. [40], where a weak form of Bell inequality was introduced. This Bell inequality is not obeyed by *all* classical vectors of probabilities for the local measurements involved, but only by those that are consistent with quantum mechanics. It turns out that these vectors correspond to quantum mechanical product states, hence all convex combinations are separable states and the weak Bell inequalities can detect all entangled states. However, this does *not* mean that there is no LHV model for entangled states, an example being the family of entangled states for which Werner constructed LHV models explicitly [26].

Further, we remarked in the introduction of chapter 2 that the Bell inequalities correspond to non-optimal witness operators in general because no known Bell inequality is violated by PPT entangled states [70] and because there exist entangled states already in systems of two qubits for which a LHV model exists for any number of local measurements [26], as also mentioned above. Hence the question about the relation of witness operators and Bell inqualities concerns the relation of the border between separable and entangled states and the border between LHV and non-LHV states.

The main difficulty here is the very large number of degrees of freedom of the Bell inequalities, because only the number of measurement settings per site and the number of measurement outcomes on each site is fixed, but not the measurement settings themselves. In contrast, if all the measurement settings are fixed, then it is possible to directly apply the formalism of Ref. [43]. This was used in Ref. [142] to show that

for certain fixed settings Bell inequalities for systems of two qutrits [76] correspond to decomposable witnesses and are hence not violated by PPT entangled states. For fixed measurement settings, it is also possible to relate Bell inequalities for the class of so-called graph states of several qubits to the projector based witnesses of section 2.2.2, as shown in Ref. [XI].

In this chapter, however, we will neither restrict the LHV models nor fix the settings when treating the CHSH inequality. In section 5.2, we recall some facts related to witnesses and the CHSH inequality and show how to write a CHSH inequality as a CHSH witness. Then we transform optimal witnesses to witnesses detecting only states that violate a CHSH inequality by shifting them with the identity in section 5.3. In section 5.4, we transform CHSH witness in the same spirit by substracting the identity, bringing them closer to the set of separable states. Then we use another approach to relate the CHSH witnesses to optimal witnesses directly, by considering the diagonalized CHSH witness in section 5.5. Finally, we conclude and name open questions in section 5.6.

5.2 Some useful facts and definitions

In this section, we concentrate on the relation between witness operators and CHSH inequalities. Before we start, let us recall some facts that we will use in this section, first about entanglement witnesses and then about CHSH inequalities. We will again omit the tensor product signs when there is no danger of confusion.

In chapter 2 we already noted that in systems of two qubits all entanglement witnesses are decomposable, i.e., of the form

$$W = P + Q^{T_A}, \tag{5.1}$$

where P and Q are positive semi-definite operators. The optimal entanglement witnesses are of the form $W = |\phi\rangle\langle\phi|^{T_A}$, where $|\phi\rangle$ is an entangled state vector. An optimal witness detecting the entangled state ρ can be constructed from the eigenvector of ρ^{T_A} with negative eigenvalue. Further, writing $|\phi\rangle$ in the Schmidt form

$$|\phi\rangle = \alpha|00\rangle + \beta|11\rangle, \quad \alpha, \beta > 0, \quad \alpha^2 + \beta^2 = 1, \tag{5.2}$$

the witness can be locally decomposed as

$$W_\alpha = \frac{1}{4}\Big(\mathbb{1}\mathbb{1} + \sigma_z\sigma_z + (\alpha^2 - \beta^2)(\sigma_z\mathbb{1} + \mathbb{1}\sigma_z) + 2\alpha\beta(\sigma_x\sigma_x + \sigma_y\sigma_y)\Big), \tag{5.3}$$

cf. Eq. (2.26).

Let us now recall some facts about the CHSH inequalities. Within quantum mechanics, the CHSH inequalities can be described by an operator $\mathcal{B}$ such that

$$|\mathrm{Tr}[\mathcal{B}\rho_{\mathrm{LHV}}]| \leq 2 \tag{5.4}$$

is fulfilled for all states ρ_{LHV} admitting a local hidden variable model for the measurements of the CHSH operator

$$\mathcal{B} = \mathbf{a}\cdot\boldsymbol{\sigma}\otimes(\mathbf{b}+\mathbf{b}')\cdot\boldsymbol{\sigma} + \mathbf{a}'\cdot\boldsymbol{\sigma}\otimes(\mathbf{b}-\mathbf{b}')\cdot\boldsymbol{\sigma}. \tag{5.5}$$

Here $\mathbf{a}$, $\mathbf{a}'$, $\mathbf{b}$, and $\mathbf{b}'$ are unit vectors describing the measurements that the parties A and B perform. There exists a necessary and sufficient criterion for the violation of a CHSH inequality found by the Horodeckis [143]. For this we need that any two qubit state can be written as

$$\rho = \frac{1}{4}\sum_{i=0}^{3}\lambda_{ij}\sigma_i\otimes\sigma_j. \tag{5.6}$$

In the following, we will refer to the 3×3 dimensional subtensor $\lambda_{i>0,j>0}\equiv T_\rho$ as the correlation tensor. This tensor holds all the information that is needed to decide whether a state violates a CHSH inequality: A state ρ violates a CHSH inequality iff $u_1+u_2>1$, where u_1 and u_2 are the two largest eigenvalues of $U_\rho = T_\rho^T T_\rho$ [143].

From the definition of optimal witnesses and the CHSH operator it follows directly that CHSH witnesses cannot be optimal witnesses. The latter can be constructed as

$$W_{\mathrm{CHSH}} = 2\cdot\mathbb{1}+\mathcal{B}, \tag{5.7}$$

they are positive on all LHV states, in particular on all separable states. The partially transposed witness $W_{\mathrm{CHSH}}^{T_A}$ is still a CHSH witness. However, for every optimal witness, $W_{\mathrm{opt}}^{T_A}$ is a positive operator. Hence W_{CHSH} cannot be optimal. In the following, we will investigate the relation between optimal witnesses and CHSH witnesses in detail.

5.3 From optimal witnesses to CHSH inequalities

First, we pose the following question: Given an optimal entanglement witness $W=|\phi\rangle\langle\phi|^{T_A}$, how much do we have to shift it by adding the identity such that it is positive on all states admitting a local hidden variable model? In other words, for which $\gamma>0$ is $W+\gamma\mathbb{1}$ a CHSH witness? We calculate bounds on γ, first considering witnesses with maximally entangled states $|\phi\rangle=|\phi^+\rangle$ and then optimal witnesses constructed with arbitrary entangled states.

For $|\phi\rangle=|\phi^+\rangle$, the optimal witnesses take the simple local form

$$W = \frac{1}{4}\Big(\mathbb{1}\mathbb{1}+\sigma_x\sigma_x+\sigma_y\sigma_y+\sigma_z\sigma_z\Big). \tag{5.8}$$

Now we can use the observation that

$$\begin{aligned}\sigma_x\sigma_x+\sigma_y\sigma_y &= \frac{1}{\sqrt{2}}\Big[\sigma_x\Big(\frac{\sigma_x+\sigma_y}{\sqrt{2}}\Big)+\sigma_x\Big(\frac{\sigma_x-\sigma_y}{\sqrt{2}}\Big)\\ &\quad+\sigma_y\Big(\frac{\sigma_x+\sigma_y}{\sqrt{2}}\Big)-\sigma_y\Big(\frac{\sigma_x-\sigma_y}{\sqrt{2}}\Big)\Big]\equiv\frac{1}{\sqrt{2}}\mathcal{B}_{x,y}\end{aligned} \tag{5.9}$$

to write the witness in terms of CHSH operators as follows

$$W = \frac{1}{4}\Big(\mathbb{1}\mathbb{1} + \frac{1}{2\sqrt{2}}\big(\mathcal{B}_{x,y} + \mathcal{B}_{x,z} + \mathcal{B}_{y,z}\big)\Big). \tag{5.10}$$

The expectation value of each of these CHSH operators is bounded by -2 from below for states admitting a local hidden variable model, so that we can estimate

$$\mathrm{Tr}[W\rho_{\mathrm{LHV}}] \geq \frac{1}{4}\Big(1 + \frac{1}{2\sqrt{2}}(-3\cdot 2)\Big) = \frac{\sqrt{2}-3}{4\sqrt{2}} \equiv -\gamma. \tag{5.11}$$

Hence, $W' = W + \gamma\cdot\mathbb{1}$ corresponds to a CHSH witness, being positive not only on separable, but more general on all states fulfilling the CHSH inequality. The spectral decomposition of W' is given by

$$W' = \Big(\frac{1}{2}+\gamma\Big)\Big[|00\rangle\langle 00| + |\psi^+\rangle\langle\psi^+| + |11\rangle\langle 11|\Big] - \Big(\frac{1}{2}-\gamma\Big)|\psi^-\rangle\langle\psi^-|, \tag{5.12}$$

and since $1/2-\gamma \approx 0.220 > 0$, W' is still detecting states.

Let us estimate the strength of the witness by looking at the following family of states

$$\rho_p = p|\psi\rangle\langle\psi| + \frac{(1-p)}{4}\mathbb{1}, \tag{5.13}$$

where $|\psi\rangle = a|01\rangle - b|10\rangle$, and $a, b \geq 0$. In the following, we abbreviate $x = ab$. For this family of states, the only eigenvector with possibly negative eigenvalue is $|\phi^+\rangle$,

$$\rho_p^{T_A}|\phi^+\rangle = \Big(-px + \frac{(1-p)}{4}\Big)|\phi^+\rangle. \tag{5.14}$$

Hence, following chapter 2, the original witness $|\phi^+\rangle\langle\phi^+|^{T_A}$ is a good witness for these states. The states are entangled provided that

$$p > p_e = \frac{1}{(1+4x)}, \tag{5.15}$$

while the witness W' detects the states provided that

$$p > p_w = \frac{3}{\sqrt{2}(1+4x)}. \tag{5.16}$$

The rhs is larger than or equal to one for $x \geq \gamma$, so that the witness does not detect any states for this range of parameters.

Let us compare the witness with the Horodecki criterion from above [143]: For the states ρ_p we have

$$T_{\rho_p} = \begin{pmatrix} -2xp & 0 & 0 \\ 0 & -2xp & 0 \\ 0 & 0 & -p \end{pmatrix}. \tag{5.17}$$

Since $x \leq 1/2$ the states are violating all CHSH inequalities if

$$p^2(1+4x^2) > 1 \Leftrightarrow p > p_h = \frac{1}{\sqrt{1+4x^2}}. \tag{5.18}$$

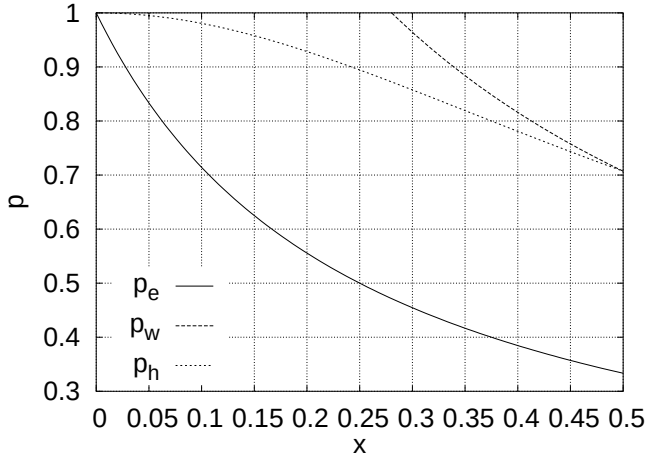

Figure 5.1. *The graphs show values of p above which the three criteria detect the states ρ_p from Eq. (5.13) depending on $x = ab$. The lowest line corresponds to the PPT criterion, the middle line to the Horodecki criterion, and the top line to the shifted witness W'.*

For $a = b = 1/\sqrt{2}$, when the states correspond to the Werner states [26], both W' and the Horodecki criterion detect the states if $p > 1/\sqrt{2}$ which is equivalent to the value found by Werner [26], indicating that γ is a good bound. For other values of a, however, the bounds differ, see Fig. 5.3. Still, the witness detects a rather large proportion of the states detected by some CHSH inequality.

Let us now consider the general optimal witnesses given by $W = |\phi\rangle\langle\phi|^{T_A}$ with $|\phi\rangle = \alpha|00\rangle + \beta|11\rangle$. We can rewrite Eq. (5.3) in the same way as above:

$$W_\alpha = \frac{1}{4}\Big[\mathbb{1}\mathbb{1} + (\alpha^2-\beta^2)(\sigma_z\mathbb{1}+\mathbb{1}\sigma_z) + (\alpha-\beta)^2\sigma_z\sigma_z + \frac{\alpha\beta}{\sqrt{2}}\Big(\mathcal{B}_{x,y}+\mathcal{B}_{x,z}+\mathcal{B}_{y,z}\Big)\Big] \quad (5.19)$$

Again, we would like to find a lower bound for this expression with respect to states not violating a CHSH inequality. The CHSH contribution is $\geq -3\sqrt{2}\alpha\beta$ for states admitting a LHV model.

The expectation value of the other terms $(\alpha^2 - \beta^2)(\sigma_z\mathbb{1} + \mathbb{1}\sigma_z) + (\alpha - \beta)^2\sigma_z\sigma_z$ is bounded from below by the minimal eigenvalue. Assuming that $\alpha \geq \beta$, this is given by $-2(\alpha^2 - \beta^2) + (\alpha - \beta)^2$ because $(\alpha^2 - \beta^2) - (\alpha - \beta)^2 = 2(\alpha\beta - \beta^2) \geq 0$. Hence we obtain for states ρ_{LHV} obeying all CHSH inequalities the bound

$$\mathrm{Tr}[W_\alpha\rho_{\mathrm{LHV}}] \geq \frac{1}{4}\Big(1 - 2(\alpha^2 - \beta^2) + (\alpha - \beta)^2 - 3\sqrt{2}\alpha\beta\Big) \equiv -\gamma_\alpha \quad (5.20)$$

which reduces to γ from Eq. (5.11) for $\alpha = 1/\sqrt{2}$. The operator $W'_\alpha = W_\alpha + \gamma_\alpha \cdot \mathbb{1}$ is positive on states admitting a local hidden variable model. However, in order to detect states, it must not be positive on all states. The eigenvector with negative eigenvalue of the witness W_α is again the state $|\psi^-\rangle$, and hence also for W'_α, with the eigenvalue $-\alpha\beta + \gamma_\alpha$. This is negative for $\alpha \leq [8/(19 - 6\sqrt{2})]^{1/2} \approx 0.872$ only, hence W'_α does not detect any states for a larger value of α.

5.4 From CHSH inequalities to witnesses

Now we adress the opposite question: how much can we shift a CHSH witness towards the set of separable states by substracting the identity so that it remains a witness? In other words, for which $\delta > 0$ is $2\mathbb{1} + \mathcal{B} - \delta\mathbb{1}$ still a witness? We calculate δ depending on the parameters of $\mathcal{B}$ and relate the CHSH witness to optimal witnesses from a restricted class of witness operators.

First let us parametrize the CHSH operator from Eq. (5.5) such that all measurements vectors lie in the $x-z$ plane. This can be done without loss of generality, because the only free parameters are the angles between the two measurement directions on each side. In particular, we can choose $\mathbf{a} = \mathbf{b} = \hat{\mathbf{z}}$ and $(\mathbf{a}, \mathbf{b})' = (\sin(\theta_{a,b}), 0, \cos(\theta_{a,b}))$. The operator takes the form

$$\mathcal{B} = -s_a s_b \cdot \sigma_x\sigma_x + s_a(1-c_b)\cdot\sigma_x\sigma_z + (1-c_a)s_b\cdot\sigma_z\sigma_x + (1+c_a+c_b-c_ac_b)\cdot\sigma_z\sigma_z, \quad (5.21)$$

where we abbreviated $\sin(\theta_{a,b}) \equiv s_{a,b}$ and in analogy for the cosine terms. This is already written in the basis of products of Pauli matrices. We can perform a singular value decomposition of the matrix of coefficients and obtain

$$\mathcal{B} = \lambda_+ \tilde{\sigma}_x\bar{\sigma}_x + \lambda_- \tilde{\sigma}_z\bar{\sigma}_z \quad (5.22)$$

$$\lambda_\pm = \Big(2\big(1 \pm \sqrt{1 - s_a^2 s_b^2}\,\big)\Big)^{\frac{1}{2}}. \quad (5.23)$$

We will now estimate the maximal expectation value that this operator can attain with respect to product states with the help of the following proposition. This will directly provide the desired bound.

Proposition 5.1. The maximal expectation value of an operator $\mathcal{A} = \alpha\sigma_x\sigma_x + \beta\sigma_z\sigma_z$ with respect to product states is given by $\max(\alpha, \beta)$. The minimal value is the maximum with opposite sign.

Proof. Using the Cauchy-Schwartz inequality, we can estimate

$$|\langle a,b|\Big(\alpha\sigma_x\sigma_x + \beta\sigma_z\sigma_z\Big)|a,b\rangle| = |\langle\sqrt{\alpha}\sigma_x\rangle_a\langle\sqrt{\alpha}\sigma_x\rangle_b + \langle\sqrt{\beta}\sigma_z\rangle_a\langle\sqrt{\beta}\sigma_z\rangle_b|$$
$$\leq \sqrt{(\alpha\langle\sigma_x\rangle_a^2 + \beta\langle\sigma_z\rangle_a^2)(\alpha\langle\sigma_x\rangle_b^2 + \beta\langle\sigma_z\rangle_b^2)}.$$

The maximum of the two terms on the rhs will surely be attained for vectors in the $x-z$ plane, for which $\langle\sigma_x\rangle^2 + \langle\sigma_z\rangle^2 = 1$ holds. With the help of Lagrange multipliers, we obtain $\max_{[x^2+z^2=1]} \alpha x^2 + \beta z^2 = \max(\alpha, \beta)$. This holds for both terms below the square root. For $\alpha > \beta$, the maximum is attained for the eigenstates of $\sigma_x\sigma_x$, with positive or negative sign, and an analogous result holds for $\alpha < \beta$. □

Hence the minimal expectation value of the CHSH witness $2\mathbb{1} + \mathcal{B}$ with respect to product states is $2 - \lambda_+$, which follows from proposition 5.1 and from $\lambda_+ \geq \lambda_-$. This means that $2\mathbb{1} + \mathcal{B} - (2 - \lambda_+)\mathbb{1} = \lambda_+\mathbb{1} + \mathcal{B}$ is still a witness.

We will show now how this witness can be related to optimal witness of the class of witnesses that can be written as

$$W = \sum_{i,j=\{0,x,z\}} c_{ij}\sigma_i \otimes \sigma_j \quad (5.24)$$

which have the property that $W = W^T = W^{T_A}$. The CHSH witness $2\mathbb{1} - \mathcal{B}$ with $\mathcal{B}$ from Eq. (5.21) belongs to this class, which we will refer to as EW_4 in the following. In Ref. [20] it was shown that the optimal witness of this class are given by

$$W_e = \frac{1}{2}\big(|\phi_e\rangle\langle\phi_e| + |\phi_e\rangle\langle\phi_e|^{T_A}\big), \tag{5.25}$$

where $|\phi_e\rangle$ is a real entangled state. Choosing $|\phi_e\rangle$ to be the Bell state $|\phi^+\rangle$ in the basis of Eq. (5.22), the corresponding witness in local form is given by $W_+ = (\mathbb{1} + \tilde{\sigma}_x\bar{\sigma}_x + \tilde{\sigma}_z\bar{\sigma}_z)/4$. Then we can write the shifted witness from above as

$$\lambda_+\mathbb{1} + \mathcal{B} = 4\lambda_- W_+ + (\lambda_+ - \lambda_-)(\mathbb{1} + \tilde{\sigma}_x\bar{\sigma}_x), \tag{5.26}$$

i.e., even after the shift the resulting witness is still given by the sum of an optimal witnesses from the class EW_4 and a positive definite operator. However, if we choose $\theta_a = \theta_b = \pi/2$, then $\lambda_+ = \lambda_- = \sqrt{2}$, and the shifted witness $\lambda_+ + \mathcal{B}$ is equal to the optimal witness from the restricted class. Still, the result indicates that the substraction of the identity might not be the optimal strategy for the optimization of the CHSH witness. In the following section, we will use a different approach.

5.5 CHSH inequalities written as non-optimal witnesses

In this section, we show explicitly how any CHSH inequality can be decomposed into a sum of an optimal witness and a general positive operator, starting from the diagonalized CHSH witness. First, we find such decompositions into an optimal witness and a positive operator, and then decompositions involving optimal witnesses W_e from the restricted class of witnesses from above.

The Bell operator of Eq. (5.5) in diagonal form is given by

$$W_{\mathrm{CHSH}} = 2\cdot\mathbb{1} + \mu_+(|\psi_1\rangle\langle\psi_1| - |\psi_2\rangle\langle\psi_2|) + \mu_-(|\psi_3\rangle\langle\psi_3| - |\psi_4\rangle\langle\psi_4|), \tag{5.27}$$

where $\mu_\pm = 2\sqrt{1 \pm s_a s_b}$ and all the eigenstates $|\psi_i\rangle$ are maximally entangled [144]. Choosing convenient local bases, these can be brought to the form

$$|\psi_1\rangle = |\phi^+\rangle, \quad |\psi_2\rangle = |\psi^-\rangle, \quad |\psi_3\rangle = |\tilde{\phi}^+\rangle, \text{ and } \quad |\psi_4\rangle = |\tilde{\psi}^-\rangle \tag{5.28}$$

where the local bases of the latter two vectors are different from the local bases of the first two vectors, while all vectors still form an orthonormal set. Note that for $\theta_a = \theta_b = \pi/2$, the eigenvalue μ_- vanishes, while μ_+ reaches its maximal value $2\sqrt{2}$, so that also W_{CHSH} has a maximal negative eigenvalue for this choice of settings. In the following, we will refer to these settings as optimal, which is further motivated by the results of the previous section.

We first write the witness W_{CHSH} directly as the sum of an optimal witness and a positive operator, i.e.,

$$W_{\mathrm{CHSH}} = \chi|\phi\rangle\langle\phi|^{T_A} + P, \quad P \geq 0, \quad \chi \geq 0, \tag{5.29}$$

where $|\phi\rangle$ is an entangled vector. We start by rewriting

$$\begin{aligned} W_{\text{CHSH}} &= 2\cdot\Big(\mathbb{1}-|\psi^-\rangle\langle\psi^-|\Big)+\mu_+|\phi^+\rangle\langle\phi^+|+\mu_-\Big(|\tilde{\phi}^+\rangle\langle\tilde{\phi}^+|-|\tilde{\psi}^-\rangle\langle\tilde{\psi}^-|\Big) \\ &\quad +(2-\mu_+)|\psi^-\rangle\langle\psi^-| \end{aligned}$$

where all the terms in the rhs of the first line are orthogonal to $|\psi^-\rangle$. We use as the entangled vector in Eq. (5.29) $|\phi\rangle=\alpha|00\rangle+\beta|11\rangle$ with $\alpha\geq\beta$, because the vector corresponding to the negative eigenvalue $-\alpha\beta$ of $|\phi\rangle\langle\phi|^{T_A}$ is $|\psi^-\rangle$. We substitute

$$|\psi^-\rangle\langle\psi^-|=-\frac{1}{\alpha\beta}\Big[|\phi\rangle\langle\phi|^{T_A}-\alpha^2|00\rangle\langle 00|-\beta^2|11\rangle\langle 11|-\alpha\beta|\psi^+\rangle\langle\psi^+|\Big], \tag{5.30}$$

arriving at

$$W_{\text{CHSH}}=\chi|\phi\rangle\langle\phi|^{T_A}+O,\ \text{where}\quad \chi\equiv\frac{\mu_+-2}{\alpha\beta}. \tag{5.31}$$

This is already of the desired form provided that O is a positive operator. An easy bound on the positivity of O can be obtained as follows: the three terms on the rhs of Eq. (5.30) appear with negative sign in O. If we put all the coefficients to α^2, then in the resulting operator O' the weight of the negative terms is increased, and from positivity of O' the positivity of O follows. Because we substract the identity in the subspace orthogonal to $|\psi^-\rangle$, we can estimate

$$\begin{aligned} O &\geq (2-\chi\alpha^2+\mu_-)|\tilde{\phi}^+\rangle\langle\tilde{\phi}^+|+(2-\chi\alpha^2+\mu_+)|\phi^+\rangle\langle\phi^+| \\ &\quad +(2-\chi\alpha^2-\mu_-)|\tilde{\psi}^-\rangle\langle\tilde{\psi}^-|, \end{aligned}$$

so that a sufficient condition for the positivity of O is $\chi\alpha^2+\mu_-\leq 2$.

Let us investigate the maximal values χ and α can attain. First we will maximize χ. Since $\text{Tr}[W_{\text{CHSH}}]=\chi+\text{Tr}[P]=8$, the decomposition with maximal weight of the partially transposed projector corresponds to a maximal χ. This is bounded by

$$\chi\leq\frac{2-\mu_-}{\alpha^2}\leq 2(2-\mu_-)\leq 4, \tag{5.32}$$

hence χ is maximized by choosing $\alpha^2=1/2$, corresponding to the maximal entangled state. The highest relative weight of $1/2$ is reached for the optimal settings, where $\mu_-=0$. Maximizing α instead, we obtain the bound

$$\alpha^2\leq\frac{y^2}{1+y^2}\in[\frac{1}{2},\frac{1}{4-2\sqrt{2}}\approx 0.854] \tag{5.33}$$

where $y=(2-\mu_-)/(\mu_+-2)$. The maximal bound is again reached for the optimal settings.

Let us now relate the diagonalized CHSH witness from Eq. (5.27) to the optimal witnesses of the class EW_4, cf. Eq. (5.25). At this point we can make use of the choice of bases leading to the eigenbasis of the CHSH witness from Eq. (5.28). As noted in the proof of Lemma 2.4, we can use that $\mathbb{1}-|\psi^-\rangle\langle\psi^-|$ is the projector onto the symmetric subspace to rewrite

$$\mathbb{1}-|\psi^-\rangle\langle\psi^-|=\frac{1}{2}(\mathbb{1}+2|\phi^+\rangle\langle\phi^+|^{T_A})\Leftrightarrow-|\psi^-\rangle\langle\psi^-|=|\phi^+\rangle\langle\phi^+|^{T_A}-\frac{\mathbb{1}}{2}. \tag{5.34}$$

Using this identity, the CHSH witness in the form of Eq. (5.27) can be written as

$$W_{\text{CHSH}} = 2\mu_+ W_e + 2\mu_- \tilde{W}_e + (2 - \frac{\mu_+ + \mu_-}{2})\mathbb{1}, \tag{5.35}$$

where $W_e = (|\phi^+\rangle\langle\phi^+| + |\phi^+\rangle\langle\phi^+|^{T_A})/2$, and in analogy for $\tilde{W}_e$. This is a good decomposition since, using the abbreviation $x = s_a s_b$,

$$2 \geq \sqrt{1+x} + \sqrt{1-x} \Leftrightarrow 4 \geq 2(1 + \sqrt{1-x^2}) \tag{5.36}$$

is always fulfilled, hence the term proportional to the identity is positive or vanishes. From this decomposition we see directly that

$$W_{\text{CHSH}} - \big(2 - \frac{\mu_+ + \mu_-}{2}\big) \cdot \mathbb{1} \tag{5.37}$$

is still a witness, but not a CHSH witness anymore. In fact, this bound is equivalent to the bound obtained with the help of proposition 5.1 from the last section, because $\mu_+ + \mu_- = 2\lambda_+$. Hence the CHSH witness can be written in a very natural way as a superposition of two optimal witness from the restricted class EW_4 and the identity. For the optimal settings, the weight of one of these witnesses vanishes, and we recover the result from the end of the preceding section.

5.6 Conclusions

In this chapter, we investigated the relation between optimal witness operators and the CHSH inequality in detail. We estimated how much optimal witnesses have to be shifted by the identity to make them positive on all states admitting a LHV model.

Then we considered the opposite question and obtained tight bounds for how much the identity can be substracted from a CHSH witness, preserving the witness properties. We further related this witness to an optimal witness of the class EW_4 of witnesses which are invariant with respect to partial as well as complete transposition. The CHSH witness in the parametrization that we used is part of that class. Finally, we diagonalized the witness and related it to general optimal witnesses, as well as to optimal witnesses of the class EW_4. We found a natural decomposition into two such optimal witnesses and the identity, where the weight of the identity matched the bound that we had obtained before on how much the identity can be substracted from the CHSH witness.

A natural next step would be to investigate the relationship between witnesses and more complex Bell inequalities, for instance, the inequality involving three dichotomic measurements per site for two parties found by Śliwa [79], Collins, and Gisin [80]. Even more fascinating would be the step to more parties or to systems of higher dimension, because of Peres' conjecture that PPT entangled states do not violate any Bell inequality [87]. If it would be possible to show that all Bell inequality correspond to decomposable witnesses for any choice of the measurements, then the conjecture would proven. However, the investigation will become increasingly difficult with the increasing degrees of freedom of the Bell inqualities in higher dimensions.

Bibliography

[1] A. Einstein, N. Podolski, and N. Rosen, Phys. Rev. **47**, 777 (1935). German translation in [4].

[2] E. Schrödinger, Naturwissenschaften **23**, 807 (1935); **23**, 823 (1935); **23**, 844 (1935). Reprinted in [4].

[3] J.S. Bell, Physics **1**, 195 (1964). Reprinted in [5].

[4] K. Baumann and R.U. Sexl, *Die Deutungen der Quantentheorie*, (Vieweg, Braunschweig, 1987).

[5] J.S. Bell, *Speakable and Unspeakable in Quantum Mechanics*, (Cambridge University Press, Cambridge 1988).

[6] R. Feynman, Int. J. Theor. Phys. **21**, 467 (1982).

[7] D. Deutsch, Proc. Roy. Soc. A **400**, 96 (1985).

[8] P. Shor, SIAM J. Comp. **26**, 1484 (1997). Preprint: quant-ph/9508027.

[9] L. Grover, Phys. Rev. Lett. **79**, 325 (1997). Preprint: quant-ph/9706033.

[10] C.H. Bennett and G. Brassard, in *Proc. IEEE Int. Conference on Computers, Systems, and Signal Processing*, (IEEE, New York, 1984).

[11] A.K. Ekert, Phys. Rev. Lett. **67**, 661 (1991).

[12] C.H. Bennett, G. Brassard, C. Crépeau, R. Jozsa, A. Peres and W.K. Wootters, Phys. Rev. Lett. **70**, 1895 (1993).

[13] C.H. Bennett and S.J. Wiesner, Phys. Rev. Lett. **69**, 2881 (1992).

[14] C. Brukner, M. Żukowski, J.W. Pan, and A. Zeilinger, Phys. Rev. Lett. **92**, 127901 (2004). Preprint: quant-ph/0210114.

[15] L.M.K. Vandersypen, M. Steffen, G. Breyta, C.S. Yannoni, M.H. Sherwood and I.L. Chuang, Nature **414**, 883 (2001).

[16] S. Gulde, M. Riebe, G.P.T. Lancaster, C. Becher, J. Eschner, H. Häffner, F. Schmidt-Kaler, I.L. Chuang and R. Blatt, Nature **421**, 48 (2003).

[17] D. Bouwmeester, A. Ekert, and A. Zeilinger (Eds.), *The Physics of Quantum Information,* (Springer, Berlin, 2000).

[18] M. Nielsen and I. Chuang, *Quantum Computation and Quantum Information,* (Cambrigde University Press, Cambridge, 2001).

[19] G. Alber, R. Beth, M. Horodecki, P. Horodecki, R. Horodecki, M. Rötteler, H. Weinfurter, R. Werner, and A. Zeilinger, *Quantum Information,* (Springer, Berlin, 2001).

[20] M. Curty, M. Lewenstein, and N. Lütkenhaus, Phys. Rev. Lett. **92**, 217903 (2004). Preprint: quant-ph/0307151.

[21] M. Curty, O. Gühne, M. Lewenstein and N. Lütkenhaus, unpublished. Preprint: quant-ph/0409047.

[22] J.W. Pan, D. Bouwmeester, M. Daniell, H. Weinfurter, A. Zeilinger, Nature **403**, 515 (2000).

[23] C.A. Sackett, D. Kielpinski, B.E. King, C. Langer, V. Meyer, C.J. Myatt, M. Rowe, Q.A. Turchette, W.M. Itano, D.J. Wineland, I.C. Monroe, Nature **404** 256, (2000).

[24] R.A. Horn and C.R. Johnson, *Matrix analysis*, (Cambridge University Press, Cambridge, 1985).

[25] R.A. Horn and C.R. Johnson, *Topics in matrix analysis*, (Cambridge University Press, Cambridge, 1991).

[26] R.F. Werner, Phys. Rev. A **40**, 4277 (1989).

[27] A. Peres, Phys. Rev. Lett. **77**, 1413 (1996). Preprint: quant-ph/9604005.

[28] M. Horodecki, P. Horodecki, and R. Horodecki, Phys. Lett. A **223**, 1 (1996). Preprint: quant-ph/9605038.

[29] P. Horodecki, Phys. Lett. A **232**, 333 (1997). Preprint: quant-ph/9703004.

[30] O. Rudolph, J. Phys. A **36**, 5825 (2003); Phys. Rev. A **67**, 032312 (2003). Preprint: quant-ph/0202121.

[31] K. Chen and L. Wu, Quant. Inf. Comp. **3**, 193 (2003). Preprint: quant-ph/0205017.

[32] M. Horodecki and P. Horodecki, Phys. Rev. A **59**, 4206 (1999). Preprint: quant-ph/9708015.

[33] M.A. Nielsen and J. Kempe, Phys. Rev. Lett. **86**, 5184 (2001). Preprint: quant-ph/0011117.

[34] B. Kraus, J.I. Cirac, S. Karnas, and M. Lewenstein, Phys. Rev. A **61**, 062302 (2000). Preprint: quant-ph/9912010.

[35] S. Karnas and M. Lewenstein, Phys. Rev. A **64**, 042313 (2001). Preprint: quant-ph/0102115.

[36] A.C. Doherty, P.A. Parrilo, and F.M. Spedalieri, Phys. Rev. Lett. **88**, 187904 (2002). Preprint: quant-ph/0112007.

[37] A.C. Doherty, P.A. Parrilo, and F.M. Spedalieri, Phys. Rev. A **69**, 022308 (2004). Preprint: quant-ph/0308032.

[38] F.G.S.L. Brañdao and R.O. Vianna, unpublished. Preprints: quant-ph/0405008, quant-ph/0405063, quant-ph/0405096.

[39] A.C. Doherty, P.A. Parrilo, and F.M. Spedalieri, unpublished. Preprint: quant-ph/0407143.

[40] B.M. Terhal, Phys. Lett. A **271**, 319 (2000). Preprint: quant-ph/9911057.

[41] R.T. Rockefellar, *Convex analysis* (Princeton University Press, Princeton, 1970).

[42] W. Rudin, *Functional Analysis*, (McGraw-Hill, Singapore, 1991).

[43] M. Lewenstein, B. Kraus, J.I. Cirac, and P. Horodecki, Phys. Rev. A **62**, 052310 (2000). Preprint: quant-ph/0005014.

[44] C.H. Bennett, G. Brassard, S. Popescu, B. Schumacher, J.A. Smolin, and W.K. Wootters, Phys. Rev. Lett. **76**, 722 (1996). Preprint: quant-ph/9511027.

[45] C.H. Bennett, H.J. Bernstein, S. Popescu, and B. Schumacher, Phys. Rev. A **53**, 2046 (1996).

[46] C.H. Bennett, S. Popescu, D. Rohrlich, J.A. Smolin, and A.V. Thapliyal, Phys. Rev. A **63**, 012307 (2001).

[47] N. Linden, S. Popescu, B. Schumacher, and M. Westmoreland, quant-ph/9912039; E.F. Galvao, M.B. Plenio, and S. Virmani, J. Phys. A **33**, 8809 (2000); S. Wu and Y. Zhang, Phys. Rev. A **63**, 012308 (2001); A. Acín, G. Vidal, and J.I. Cirac, Quant. Inf. Comp. **3**, 55 (2003).

[48] H. Barnum and N. Linden, J. Phys. A **34**, 6787 (2001).

[49] T.-C. Wei and P.M. Goldbart, Phys. Rev. A **68**, 042307 (2003). Preprint: quant-ph/0307219.

[50] V. Coffman, J. Kundu, and W.K. Wootters, Phys. Rev. A **61**, 052306 (2000); J. Eisert and H.J. Briegel, Phys. Rev. A **64**, 022306 (2001); D.A. Meyer and N.R. Wallach, J. Math. Phys. **43**, 4273 (2002); F. Verstraete, J. Dehaene, and B. De Moor Phys. Rev. A **68**, 012103 (2003); A.J. Scott, Phys. Rev. A **69**, 052330 (2004); M. Hein, J. Eisert, and H.J. Briegel, Phys. Rev. A **69**, 062311 (2004).

[51] M. Horodecki, P. Horodecki, and R. Horodecki, Phys. Rev. Lett. **80**, 5239 (1998). Preprint: quant-ph/9801069.

[52] W. Dür, J.I. Cirac, M. Lewenstein, and D. Bruß, Phys. Rev. A **61**, 062313 (2000). Preprint: quant-ph/9910022.

[53] D.P. DiVincenzo, P.W. Shor, J.A. Smolin, B.M. Terhal, and A.V. Thapliyal, Phys. Rev. A **61**, 062312 (2000). Preprint: quant-ph/9910026.

[54] P. Horodecki, M. Horodecki, and R. Horodecki, Phys. Rev. Lett. **82**, 1056 (1999).

[55] K.G.H. Vollbrecht and M.M. Wolf, Phys. Rev. Lett. **88**, 247901 (2002). Preprint: quant-ph/0201103.

[56] K. Horodecki, M. Horodecki, P. Horodecki, and J. Oppenheim, unpublished. Preprint: quant-ph/0309110.

[57] A. Peres, Phys. Lett. A **202**, 16 (1995).

[58] H.A. Carteret, A. Higuchi, and A. Sudbery, J. Math. Phys. **41**, 7932 (2000). Preprint: quant-ph/0006125.

[59] A. Acín, A. Andrianov, L. Costa, E. Jane, J.I. Latorre, and R. Tarrach, Phys. Rev. Lett. **85**, 1560 (2000). Preprint: quant-ph/0003050.

[60] W. Dür, J.I. Cirac, and R. Tarrach, Phys. Rev. Lett. **83**, 3562 (1999). Preprint: quant-ph/9903018.

[61] W. Dür and J.I. Cirac, Phys. Rev. A **61**, 042314 (2000). Preprint: quant-ph/9911044.

[62] W. Dür and J.I. Cirac, J. Phys. A **34**, 6837 (2001). Preprint: quant-ph/0011025.

[63] W. Dür, G. Vidal, and J.I. Cirac, Phys. Rev. A **62**, 062314 (2000). Preprint: quant-ph/0005115.

[64] D.M. Greenberger, M. Horne, and A. Zeilinger, *Bell's theorem, quantum theory, and conceptions of the unviverse*, ed. M. Kafatos, (Kluwer, Dordrecht, 1999).

[65] B.M. Terhal and P. Horodecki, Phys. Rev. A **61**, R040301 (2000). Preprint: quant-ph/9911117.

[66] A. Miyake and F. Verstraete, Phys. Rev. A **69**, 012101 (2004). Preprint: quant-ph/0307067.

[67] F. Verstraete, J. Dehaene, B. De Moor, and H. Verschelde, Phys. Rev. A **65**, 052112 (2002). Preprint: quant-ph/0109033.

[68] J.F. Clauser, M.A. Horne, A. Shimony, and R.A. Holt, Phys. Rev. Lett. **23**, 880 (1969). Erratum: *ibid.* **24**, 549 (1970).

[69] D. Kaszlikowski and M. Żukowski, Int. J. Theor. Phys. **42**, 1023 (2003). Preprint: quant-ph/0302165.

[70] R.F. Werner and M.M. Wolf Quant. Inf. Comp. **1**, 1 (2001). Preprint: quant-ph/0102024.

[71] N.D. Mermin, Phys. Rev. Lett. **65**, 1838 (1990).

[72] M. Ardehali, Phys. Rev. A **46**, 5375 (1992).

[73] A.V. Belinskii and D.N. Klyshko, Sov. Phys. Usp. **36**, 653 (1993).

[74] R.F. Werner and M.M. Wolf, Phys. Rev. A **64**, 032112 (2001). Preprint: quant-ph/0102024.

[75] M. Żukowski and C. Brukner, Phys. Rev. Lett. **88**, 210401 (2002). Preprint: quant-ph/0102039.

[76] D. Collins, N. Gisin, N. Linden, S. Massar, and S. Popescu, Phys. Rev. Lett. **88**, 040404 (2002). Preprint: quant-ph/0106024.

[77] X.-H. Wu and H.-S. Zong, Phys. Lett. A **307**, 262 (2003).

[78] W. Laskowski, T. Paterek, M. Żukowski, and C. Brukner, Phys. Rev. Lett. **93**, 200401 (2004). Preprint: quant-ph/0411066.

[79] C. Śliwa, unpublished. Preprint: quant-ph/0305190.

[80] D. Collins and N. Gisin, J. Phys. A **37**, 1775 (2004). Preprint: quant-ph/0306129.

[81] A. Aspect, J. Dalibard, and G. Roger, Phys. Rev. Lett. **49**, 1804 (1982).

[82] M.A. Rowe, D. Kielpinski, V. Meyer, C.A. Sackett, W.M. Itano, C. Monroe, and D.J. Wineland, Nature **409**, 791 (2001).

[83] T. Jennewein, G. Weihs, J.-W. Pan, and A. Zeilinger, Phys. Rev. Lett. **88**, 017903 (2002). Preprint: quant-ph/0201134.

[84] S. Popescu, Phys. Rev. Lett **74**, 2619 (1995).

[85] N. Gisin, Phys. Lett. A **210**, 151 (1996).

[86] W. Dür, Phys. Rev. Lett. **87**, 230402 (2001). Preprint: quant-ph/0107050.

[87] A. Peres, Found. Phys. **29**, 589 (1999). Preprint: quant-ph/9807017.

[88] A. Ácin, Phys. Rev. Lett. **88**, 027901 (2002). Preprint: quant-ph/0108029.

[89] R.F. Werner and M.M. Wolf, Phys. Rev. A **61**, 062102 (2000). Preprint: quant-ph/9910063.

[90] J.M.G. Sancho and S.F. Huelga, Phys. Rev. A **61**, 042303 (2000). Preprint: quant-ph/9910041.

[91] H. Aschauer, J. Calsamiglia, M. Hein, H.J. Briegel, Quant. Inf. Comp. **4**, 383 (2004). Preprint: quant-ph/0306048.

[92] A. Ekert and P. Horodecki, Phys. Rev. Lett. **89**, 127902 (2002). Preprint: quant-ph/0111064.

[93] C. Moura Alves, P. Horodecki, D.K.L. Oi, L.C. Kwek, and A.K. Ekert, Phys. Rev. A **68**, 032306 (2003). Preprint: quant-ph/0304123.

[94] H.A. Carteret, unpublished. Preprint: quant-ph/0309216.

[95] B.M. Terhal, Theor. Comput. Sci. **287**, 313 (2002). Preprint: quant-ph/0101032.

[96] C.H. Bennett, D.P. DiVincenzo, T. Mor, P.W. Shor, J.A. Smolin, and B.M. Terhal, Phys. Rev. Lett. **82**, 5385 (1999). Preprint: quant-ph/9808030.

[97] D. Bruß and A. Peres, Phys. Rev. A **61**, 30301 (2000). Preprint: quant-ph/9911056.

[98] A. Acin, D. Bruß, M. Lewenstein and A. Sanpera, Phys. Rev. Lett. **87**, 40401 (2001). Preprint: quant-ph/0103025.

[99] A. Sanpera, R. Tarrach, and G. Vidal, Phys. Rev. A **58**, 826 (1998). Preprint: quant-ph/9801024.

[100] A. Sanpera, D. Bruß, and M. Lewenstein, Phys. Rev. A **63**, 050301 (2001). Preprint: quant-ph/0009109.

[101] M. Lewenstein and A. Sanpera, Phys. Rev. Lett. **80**, 2261 (1998). Preprint: quant-ph/9707043.

[102] M. Barbieri, F. De Martini, G. Di Nepi, P. Mataloni, G.M. D'Ariano, and C. Macchiavello, Phys. Rev. Lett. **91**, 227901 (2003). Preprint: quant-ph/0307003.

[103] J. Schlienz and G. Mahler, Phys. Rev. A **52**, 4396 (1995).

[104] A. O. Pittenger and M. H. Rubin Phys. Rev. A 67, 012327 (2003). Preprint: quant-ph/0207024.

[105] B.M. Terhal, Lin. Alg. Appl. **323**, 61 (2000).

[106] We used the an implementation of the Fletcher-Reeves conjugate gradient algorithm from the GNU scientific library which is available at http://sources.redhat.com/gsl.

[107] O. Gühne, PhD thesis,*Detecting Quantum Entanglement: Entanglement witnesses and uncertainty relations*, Hannover 2004.

[108] N. Kiesel, M. Bourennane, C. Kurtsiefer, H. Weinfurter, D. Kaszlikowski, W. Laskowski, and M. Żukowski, J. Mod. Optics **50**, 1131 (2003).

[109] P.G. Kwiat, K. Mattle, H. Weinfurter, and A. Zeilinger, Phys. Rev. Lett. **75**, 4337 (1995).

[110] H. Weinfurter and M. Żukowski, Phys. Rev. A, **64**, 010102 (2001). Preprint: quant-ph/0103049.

[111] M. Eibl, N. Kiesel, M. Bourennane, C. Kurtsiefer, and H. Weinfurter Phys. Rev. Lett. **92**, 077901 (2004).

[112] M. Eibl, S. Gaertner, M. Bourennane, C. Kurtsiefer, M. Żukowski, and H. Weinfurter, Phys. Rev. Lett. **90**, 200403 (2003). Preprint: quant-ph/0302042.

[113] G. Tóth and O. Gühne, unpublished. Preprint: quant-ph/0405165.

[114] G. Tóth, unpublished. Preprint: quant-ph/0406061.

[115] We thank R.F. Werner for raising this point.

[116] A. Peres, Fortschritte der Physik **48**, 397 (2000).

[117] W. van Dam, R.D. Gill, and P.D. Grünwald, *The statistical strength of nonlocality proofs*, unpublished. Available at http://www.math.uu.nl/people/gill.

[118] A. Acin, J. I. Cirac, and Ll. Masanes, Phys. Rev. Lett. **92**, 107903 (2004). Preprint: quant-ph/0311064.

[119] L.M. Vandersypen M. Steffen, G. Breyta, C.S. Yannon, R. Cleve, and I.L. Chuang, Phys. Rev. Lett. **85**, 5452 (2000). Preprint: quant-ph/0007017.

[120] F. Schmidt-Kaler, H. Häffner, M. Riebe, S. Gulde, G.P.T. Lancaster, T. Deuschle, C. Becher, C.F. Roos, J. Eschner, and R. Blatt, Nature **422**, 408 (2003).

[121] A. Barenco, C.H. Bennett, R. Cleve, D.P. DiVincenzo, N. Margolus, P. Shor, T. Sleator, J. Smolin, and H. Weinfurter, Phys. Rev. A **52**, 3457 (1995). Preprint: quant-ph/9503016.

[122] M. Mottonen, J. J. Vartiainen, V. Bergholm, and M. M. Salomaa, Phys. Rev. Lett. 93, 130502 (2004).

[123] K. Audenaert, J. Eisert, E. Jane, M.B. Plenio, S. Virmani, and B. de Moor, Phys. Rev. Lett. **87**, 217902 (2001). Preprint: quant-ph/0103096.

[124] E.M. Rains, IEEE Trans. Inf. Theory **47**, 2921 (2001). Preprint: quant-ph/0008047.

[125] M. Jezek, J. Rehacek, and J. Fiurasek, Phys. Rev. A **65**, 060301 (2002); K. Audenaert and B. De Moor, Phys. Rev. A **65**, 030302 (2002); F. Verstraete and H. Verschelde, Phys. Rev. Lett. **90**, 097901 (2003); K. Audenaert, M.B. Plenio, and J. Eisert, Phys. Rev. Lett. **90**, 027901 (2003); Y.C. Eldar, M. Stojnic, and B. Hassabi, Phys. Rev. A **69**, 062318 (2004); B. Synak, K. Horodecki, and M. Horodecki, unpublished. Preprint: quant-ph/0405149.

[126] K. Audenaert, in *Proceedings of Quantum Theory and Global Optimisation*, Sixteenth International Symposium on Mathematical Theory of Networks and Systems (MTNS2004), Catholic University of Leuven, Belgium, 5-9 July 2004. Preprint: quant-ph/0402076.

[127] L. Gurvits, Annual ACM Symposium on Theory of Computing, Proceedings of the thirty-fifth ACM symposium on theory of computing, San Diego, CA, USA (2003).

[128] L.M. Ioannou, B.C. Travaglione, D.C. Cheung, and A.K. Ekert, unpublished. Preprint: quant-ph/0403041.

[129] L. Vandenberghe and S. Boyd, *Semidefinite programming.* SIAM Review **38**, 49 (1996); C. Helmberg, *Semidefinite programming*, European Journal of Operational Research **137**, 461 (2002).

[130] F. Hulpke and D. Bruss, unpublished. Preprint: quant-ph/0407179.

[131] N.Z. Shor, Soviet Journal of Circuits and Systems Sciences **25**, 1 (1987).

[132] M. Kojima and L. Tuncel, Mathematical Programming **89**, 79 (2000); A. Takeda, K. Fujisawa, Y. Fukaya, and M. Kojima, Journal of Global Optimization **24**, 237 (2002); M. Kojima, S. Kim, and H. Waki, J. Ope. Res. Soc. Japan **46**, 2 (2003).

[133] J.B. Lasserre, SIAM J. Optimization **11**, 796 (2001).

[134] D. Henrion and J.B. Lasserre, IEEE Control Systems Magazine **24**, 72 (2004).

[135] P.A. Parrilo, *Structured semi-definite programs and semi-algebraic geometry methods in robustness and optimization* (PhD thesis, California Institute of Technology, Pasadena, 2000).

[136] N.S. Jones and N. Linden, unpublished. Preprint: quant-ph/0407117.

[137] M. Grötschel, L. Lovasz, and A. Schrijver, *Geometric algorithms and combinatorical optimization* (Springer, Heidelberg, 1988).

[138] A. Shimony, Ann. NY Acad. Sci. **755**, 675 (1995).

[139] D. Henrion and J.B. Lasserre, ACM Transactions on Mathematical Software **29**, 165 (2003). See also the web page www.laas.fr/∼henrion/software/gloptipoly/gloptipoly.html.

[140] J.F. Sturm, Optimization Methods and Software **11**, 625 (1999); see also the documentation of the software on the web page fewcal.kub.nl/sturm/software/sedumi.html.

[141] J. Kempe, D. Bacon, D.A. Lidar, and K.B. Whaley, Phys. Rev. A **63**, 042307 (2001).

[142] A. Acín, T. Durt, N. Gisin, and J.I. Latorre, Phys. Rev. A **65**, 052325 (2002). Preprint: quant-ph/0111143.

[143] R. Horodecki, P. Horodecki, and M. Horodecki, Phys. Lett. A **200**, 340 (1995).

[144] V. Scarani, and N. Gisin, J. Phys. A **34**, 6043 (2001). Preprint: quant-ph/0103068.

List of Publications

[I] P. Hyllus and E. Sjöqvist: *Comment on "Complementarity between Local and Nonlocal Topological Effects".* Phys. Rev. Lett. **89**, 198901 (2002).

[II] O. Gühne, P. Hyllus, D. Bruß, A. Ekert, M. Lewenstein, C. Macchiavello, and A. Sanpera: *Detection of entanglement with few local measurements.* Phys. Rev. A **66**, 062305 (2002). Preprint: quant-ph/0205089.

[III] P. Hyllus and E. Sjöqvist: *Precession and interference in the Aharonov-Casher and scalar Aharonov-Bohm effect.* Foundations of Physics **33**, 1085 (2003). Preprint: quant-ph/0210070.

[IV] K. Eckert, O. Gühne, F. Hulpke, P. Hyllus, J. Korbicz, J. Mompart, D. Bruß, M. Lewenstein, and A. Sanpera: *Entanglement properties of composite quantum systems.* In: G. Leuchs, T. Beth (Hrsg.): *Quantum information processing*, Wiley-VCH (Berlin) 2003, ISBN 3-527-40371-X. Preprint: quant-ph/0210107.

[V] O. Gühne, P. Hyllus, D. Bruß, A. Ekert, M. Lewenstein, C. Macchiavello, and A. Sanpera: *Experimental detection of entanglement via witness operators and local measurements.* J. Mod. Opt. **50**, 1079 (2003). Preprint: quant-ph/0210134.

[VI] O. Gühne and P. Hyllus: *Investigating three qubit entanglement with local measurements.* Int. J. Theor. Phys. **42**, 1001 (2003). Preprint: quant-ph/0301162.

[VII] M. Bourennane, M. Eibl, C. Kurtsiefer, S. Gaertner, H. Weinfurter, O. Gühne, P. Hyllus, D. Bruß, M. Lewenstein, and A. Sanpera: *Witnessing multipartite entanglement.* Phys. Rev. Lett. **92**, 087902 (2004). Preprint: quant-ph/0309043.

[VIII] M. Bourennane, M. Eibl, S. Gaertner, C. Kurtsiefer, H. Weinfurter, A. Cabello, O. Gühne, P. Hyllus, D. Bruß, M. Lewenstein, and A. Sanpera: *Four Photon Polarization Entanglement: Tests and Applications.* Int. J. Quant. Inf. **2**, 133 (2004).

[IX] Philipp Hyllus, Carolina Moura Alves, Dagmar Bruß, Chiara Macchiavello: *Generation and detection of bound entanglement.* Phys. Rev. A **70**, 032316 (2004). Preprint: quant-ph/0405164.

[X] J. Eisert, P. Hyllus, O. Gühne, M. Curty: *Complete hierarchies of efficient approximations to problems in entanglement theory.* Phys. Rev. A **70**, 062317 (2004). Preprint: quant-ph/0407135.

[XI] O. Gühne, G. Tóth, P. Hyllus, and H.J. Briegel, *Bell inequalities for graph states.* Preprint: quant-ph/0410059.

Acknowledgements

I would like to thank all the people who contributed to the making of this thesis in one way or another.

- First of all I would like to thank my supervisor Prof. Dr. Maciej Lewenstein who is the father of it all: the nice atmosphere within the group and the splendid possibilities for research and for exchange with other institutions. I benefitted a lot from his inspiration and his vast knowledge.
- Then I am indepted to Prof. Dr. Dagmar Bruß and Prof. Dr. Anna Sanpera who helped me getting started with doing research in Hannover, and I would like to thank them for illuminating discussions and for drawing my attention to interesting topics. I am further grateful to Dagmar Bruß for refereeing this thesis.
- In particular, I enjoyed the close collaboration with Otfried Gühne on entanglement witnesses and Bell inequalities, and would like to thank him for his help in many calculations, and for all the discussions about physics and about life in general. I hope that this will continue. I also owe him thanks for the proofreading of parts of this thesis.
- In the same way I would like to thank Kai Eckert for the collaboration on Sagnac atom interferometry, and all the help with computer problems and other matters. I would also like to send my thanks to the people from the IQO who took part in this enterprise: Dr. Ernst Rasel, Dr. Christian Jentsch, and Tobias Müller.
- Then I am grateful to all the other people of the group of Maciej Lewenstein and the people from the ITP who created a nice atmosphere, especially all the PhD students who started at the same time as I did. In particular, I would like to thank Florian Hulpke for his help with "unimportant" questions about quantum theory and for correcting part of the thesis, Jarek Korbicz, as well as Ujjwal Sen, whose help in proofreading I also acknowledge. Finally, I would like to thank the people that were physically closest to me during my time as a PhD student, and who contributed a lot to the nice atmosphere: Łukasz Dobrek, Alem Mebrahtu, and Marco Krohn.
- The collaboration with the people from the experimental group of Harald Weinfurter in Munich was very stimulating. I would like to thank especially

Mohamed Bourennane, who was the driving force behind the experimental part of the joint project. The same is true regarding the collaborations on non-convex optimization with Jens Eisert and Marcos Curty, as well as on Bell inequalities for graph states with Géza Tóth.

- I am indebted to the group of Prof. Dr. Artur Ekert in Cambridge and to the group of Prof. Dr. Hans J. Briegel in Innsbruck for their hospitality and for stimulating discussions during my visits at their institutions.
- My research was supported by the *Schwerpunktprogramm Quanteninformationsverarbeitung* SPP 1078 of the *Deutsche Forschungsgemeinschaft* which I appreciated very much.

Zum Schluß möchte ich meinen Eltern und meinem Bruder Karsten danken für ihre unbedingte Unterstützung, sowie meinen Freunden in Hannover, die die Zeit hier so schön gemacht haben. Und Andrea, ohne die das alles nicht möglich gewesen wäre.

Lebenslauf

4. Mai 1976	Geboren in Wolfenbüttel.
August 1982 - Juli 1986	Besuch der Grundschule Sutthausen in Osnabrück.
August 1986 - Juli 1988	Besuch der Orientierungsstufe „In der Wüste“ in Osnabrück.
August 1988 - Juni 1995	Besuch des Gymnasiums „In der Wüste“ in Osnabrück.
Juni 1995	Abitur am Gymnasiums „In der Wüste“ in Osnabrück.
August 1995 - Sept. 1996	Zivildienst in der Sozialstation des Rotes Kreuzes Osnabrück.
April 1997	Beginn des Studiums der Physik an der Universität Hamburg.
August 1999	Vordiplom in Physik.
September 1999 - Dez. 2000	ERASMUS-Student an der Universität Uppsala in Schweden.
Januar 2001 - Juli 2001	Studium am Fysikum der Universität Stockholm.
August 2001	„Master of Science with a major in physics“.
Oktober 2001 - März 2005	Wissenschaftlicher Mitarbeiter in der Arbeitsgruppe von Prof. Dr. Maciej Lewenstein. Arbeitsgebiet: Quanteninformationsverarbeitung.