# Detecting Quantum Entanglement: Entanglement Witnesses and Uncertainty Relations

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

**Dipl.-Phys. Otfried Gühne**
geboren am 15. Mai 1975 in Münster in Westfalen

2004

Referent: Prof. Dr. M. Lewenstein
Korreferentin: Priv.-Doz. Dr. D. Bruß

Tag der Promotion: 1. Juli 2004

# Abstract

This thesis deals with methods of the detection of entanglement. After recalling some facts and definitions concerning entanglement and separability, we investigate two methods of the detection of entanglement.

In the first part of this thesis we consider so-called entanglement witnesses, mainly in view of the detection of multipartite entanglement. Entanglement witnesses are observables for which a negative expectation value indicates entanglement. We first present a simple method to construct these witnesses. Since witnesses are nonlocal observables, they are not easy to measure in a real experiment. However, as we will show, one can circumvent this problem by decomposing the witness into several local observables which can be measured separately. We calculate the local decompositions for several interesting witnesses for two, three and four qubits. Local decompositions can be optimized in the number of measurement settings which are needed for an experimental implementation. We present a method to prove that a given local decomposition is optimal and discuss with this the optimality of our decompositions. Then we present another method of designing witnesses which are by construction measurable with local measurements. Finally, we shortly report on experiments where some of the witnesses derived in this part have been used to detect three- and four-partite entanglement of polarized photons.

The second part of this thesis deals with separability criteria which are written in terms of uncertainty relations. There are two different formulations of uncertainty relations since one can measure the uncertainty of an observable by its variance as well as by entropic quantities. We show that both formulations are useful tools for the derivation of separability criteria for finite-dimensional systems and investigate the resulting criteria. Our results in this part exhibit also some more fundamental properties of entanglement: We show how known separability criteria for infinite-dimensional systems can be translated to finite-dimensional systems. Furthermore, we prove that any entropic uncertainty relation on one part of the system gives rise to a separability criterion on the composite system.

Keywords: Entanglement, Entanglement witnesses, Uncertainty relations

# Zusammenfassung

Diese Arbeit behandelt Methoden zum Nachweis von Verschränkung. Nach einer kurzen Einführung der wichtigsten Definitionen und Resultate des Separabilitätsproblems werden zwei Möglichkeiten zum Nachweis von Verschränkung untersucht.

Im ersten Teil der Arbeit behandeln wir sogenannte Verschränkungszeugen, hauptsächlich im Hinblick auf den Nachweis von Mehrparteienverschränkung. Verschränkungszeugen sind Observable, bei denen ein negativer Erwartungswert ein Anzeichen für Verschränkung ist. Wir stellen zuerst eine einfache Methode zur Konstruktion solcher Verschränkungszeugen vor. Da die so konstruierten Observablen nichtlokal sind, sind sie experimentell nicht leicht zugängig. Wie wir dann zeigen, läßt sich diese Problem jedoch durch eine Zerlegung des Zeugen in mehrere lokale Observable lösen, diese lokalen Observablen können dann einzeln gemessen werden. Wir berechnen die lokalen Zerlegungen verschiedener Verschränkungszeugen für interessante Zustände von zwei, drei und vier Qubits. Diese Zerlegungen können in dem Sinne, daß sie möglichst wenige lokale Messungen erfordern, optimiert werden. Wir geben ein Verfahren an, mit dem man für eine gegebene Zerlegung untersuchen kann, ob sie optimal ist und untersuchen die vorher berechneten Zerlegungen damit. Dann wird noch ein anderes Verfahren zur Konstruktion von Zeugenoperatoren eingeführt, bei dem die Zeugen automatisch lokal meßbar sind. Schließlich wird noch von Experimenten berichtet, in denen einige der in dieser Arbeit berechneten Zeugen zum Nachweis von Polarisationsverschränkung in Systemen von drei und vier Photonen benutzt wurden.

Der zweite Teil dieser Arbeit behandelt den Nachweis von Verschränkung mithilfe von Kriterien, die auf Unschärferelationen basieren. Es gibt zwei verschiedene Formulierungen von Unschärferelationen, entweder wird die Varianz oder eine Entropie als Maß für die Unschärfe einer Observable genommen. Wir zeigen, daß beide Formulierungen zur Herleitung von Separabilitätskriterien für endlichdimensionale Systeme geeignet sind und untersuchen die sich ergebenden Kriterien. Ferner zeigen die Resultate dieses Teils der Arbeit einige grundlegende Zusammenhänge auf. So zeigen wir, wie einige bekannte Separabilitätskriterien für unendlichdimensionale Systeme auf endlichdimensionale Systeme übertragen werden können. Außerdem zeigen wir, daß aus jeder entropischen Unschärferelation für ein Teilsystem ein Separabilitätskriterium für das Gesamtsystem folgt.

Schlagworte: Verschränkung, Verschränkungszeugen, Unschärferelationen

*Es gibt nicht Ödes, nichts Unfruchtbares, nichts Totes in der Welt;*
*kein Chaos, keine Verwirrung, außer einer scheinbaren;*
*ungefähr wie sie in einem Teiche zu herrschen schiene,*
*wenn man aus einiger Entfernung eine verworrene Bewegung*
*und sozusagen ein Gewimmel von Fischen sähe,*
*ohne die Fische selbst zu unterscheiden.*

Gottfried Wilhelm Leibniz

# Contents

# Introduction

The study of the phenonemon of entanglement goes back to the thirties of the last century. Albert Einstein, Boris Podolski and Nathan Rosen were the first who studied the counterintuitive features of quantum mechanical correlations in composite systems [1]. Inspired by this, Erwin Schrödinger coined the term „Verschränkung“ (translated as “entanglement”) to describe these correlations [2]. However, all these authors wanted to express their disapproval with the consequences of quantum theory. They viewed entanglement as a property of quantum theory contradicting the intuition so much that they concluded that quantum theory cannot be a fundamental and complete theory.

In the following years the physicists did not pay much attention to the study of entanglement. It was known to be a bizarre phenomenon, but to the majority of physicists it did not seem to be an interesting or useful topic for research. This situation changed dramatically in the last twenty years, mainly for two reasons.

One of the reasons is of theoretical nature. It has turned out that quantum mechanics enables to perform tasks which are not possible within classical mechanics. These new possibilities concern mainly computational tasks [3–6], quantum cryptography [7, 8] and quantum teleportation [9]. It has also become more and more clear that with respect to these new possibilities entanglement plays a crucial role.

The other reason lies in the experimental progress in the last decades. Entanglement is not a purely theoretical concept anymore, it has been produced and investigated in many experimental situations [10, 11]. The experimental progress also allowed to realize some of the new protocols mentioned above [12, 13].

Although there has been a lot of effort to characterize entanglement in the last years, it is still not fully understood. For instance, for the simple question whether a given state of a bipartite system is entangled or not, no general answer is known. The situation gets even more complicated when more then two parties are involved.

The goal of this thesis is to characterize entanglement under a certain perspective, namely the perspective of the detection of entanglement. All the results presented here can be viewed as attempts to answer the question: *What shall we measure to prove that a given state is entangled?* To this aim we develop criteria for entanglement which are directly related to measurement data. Our results will, however, as a byproduct also establish some facts on a fundamental level. For instance, we will investigate the connection between entanglement and uncertainty relations.

This thesis is structured as follows. In the first chapter we will give a brief introduction into the notions of entanglement and separability. It is not our aim to give a complete overview there, we only want to introduce the facts that are needed for the understanding of the rest of this thesis. After this introduction, this thesis is divided into two parts. As we will see, there are connections between them, but the main ideas are completely different.

The first part, in the Chapter 2, investigates the possibility of detecting entanglement using the method of entanglement witnesses. These are special observables which are designed for the detection of entanglement. We will mainly study their application to multipartite systems. We will first provide methods to construct them. Since they are nonlocal observables, it seems on the first view that they are difficult to implement experimentally. However, we will show that this difficulty can be circumvented by using local decompositions of witnesses. We will calculate these local decompositions for several types of witnesses and investigate the question whether these decompositions are optimal in a sense to be defined. Finally, we will also shortly report on recent experiments performed by Mohamed Bourennane and coworkers in the group of Harald Weinfurter in Garching, where some of the witnesses presented in this chapter have been implemented in order to detect true multipartite entanglement. The results presented in Chapter 2 are based on the Refs. [I, III, IV, VI, VII][1].

The second part, the Chapter 3, investigates whether criteria based on uncertainty relations can be used for the detection of entanglement. There are two different formulations of the uncertainty principle. The first one uses variances as the measure of the uncertainty and the second one uses entropies for this task. We will consider both formulations and show that both can be used to detect entanglement. We will also establish some fundamental connections between uncertainty relations and entanglement. For instance we will show how any entropic uncertainty relation on on part of the system can be used to derive a separability criteria on the composite system. The results in this chapter origin from the Refs. [V, VIII].

[1] References in Roman numerals refer to the publication list on page 89.

# Chapter 1

# Entanglement

In this chapter we want to give a short introduction into the notions of entanglement, separability, and the so-called separability problem. Again, we want to remind the reader that we do not aim to give a complete overview of the subject, we only want to introduce the concepts and definitions which are necessary for the understanding of the further course of this thesis. This approach implies that concerning proofs we mostly refer to the literature. We also want to keep this introduction as non-technical as possible. Some more technical facts are introduced in the following chapters at the place where they are needed. Overviews of the topics presented here can be found in Refs. [14–17].

## 1.1 Bipartite entanglement

In this section, we want to introduce some facts about bipartite entanglement. We only look at finite-dimensional systems here. The properties of infinite-dimensional subsystems will be discussed later in Section 3.7. Let us start our discussion with pure states.

### 1.1.1 Pure states

Let us assume that we have given two quantum systems. The first one is owned by one physicist, called Alice, and the second one by another one, called Bob. The physical states of Alice's system may be described by states in a Hilbert space $\mathcal{H}_A$ of dimension $N$, and in Bob's system in a Hilbert space $\mathcal{H}_B$ of dimension $M$. The composite system of both parties is then described by vectors in the tensor-product of the two spaces $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B$. This is defined as follows: Let $|a_i\rangle$ be a basis of Alice's space and $|b_j\rangle$ be a basis of Bob's space. Then $\mathcal{H}_A \otimes \mathcal{H}_B$ is the set of all linear combinations of the states $|a_i\rangle \otimes |b_j\rangle$. Thus any vector in $\mathcal{H}_A \otimes \mathcal{H}_B$ can be written as

$$|\psi\rangle = \sum_{i,j=1}^{N,M} c_{ij} |a_i\rangle \otimes |b_j\rangle \in \mathcal{H}_A \otimes \mathcal{H}_B, \tag{1.1}$$

with a complex $N \times M$ matrix $C = (c_{ij})$. From this it directly follows that the dimension of $\mathcal{H}$ equals $N \cdot M$. To keep the notation short, we often write tensor products of vectors as $|a\rangle \otimes |b\rangle \equiv |a\rangle|b\rangle \equiv |ab\rangle$.

The measurement of observables[1] can be defined in a similar way: If $A$ is an observable on Alice's space and $B$ on Bob's space, the expectation value of $A \otimes B$ is defined for the basis vectors by $Tr((A \otimes B)(|a_i\rangle\langle a_i| \otimes |b_j\rangle\langle b_j|)) = Tr(A|a_i\rangle\langle a_i|)Tr(B|b_j\rangle\langle b_j|)$. This definition is extended by linearity to other observables and states. We often write the expectation value of an observable $M$ in the state $\varrho$ as $Tr(M\varrho) \equiv \langle M\rangle_\varrho$. Now we can define separability and entanglement for pure states.

**Definition 1.1.** A pure state $|\psi\rangle \in \mathcal{H}$ is called a *product state* or *separable* if we can find states $|\phi^A\rangle \in \mathcal{H}_A$ and $|\phi^B\rangle \in \mathcal{H}_B$ such that

$$|\psi\rangle = |\phi^A\rangle \otimes |\phi^B\rangle \tag{1.2}$$

holds. Otherwise the state $|\psi\rangle$ is called *entangled.*

Physically, the definition of product states means that the state is uncorrelated. A product state can thus easily be prepared in a local way: Alice produces the state $|\phi^A\rangle$ and Bob produces independently $|\phi^B\rangle$. If Alice measures any observable $A$ and Bob measures $B$, then the probabilities of the different outcomes factorize. Thus, the measurement outcomes for Alice do not depend on the outcomes on Bob's side.

The question, whether a given pure state is entangled or not, is easy to decide: $|\psi\rangle$ is a product state, if and only if the rank of the matrix $C = (c_{ij})$ in Eq. (1.1) equals one. This is due to the fact that a matrix $C$ is of rank one, if and only if there exist two vectors $\mathfrak{a}$ and $\mathfrak{b}$ such that $c_{ij} = \mathfrak{a}_i\mathfrak{b}_j$. In this case one can write $|\psi\rangle = (\sum_i \mathfrak{a}_i|a_i\rangle) \otimes (\sum_j \mathfrak{b}_j|b_j\rangle)$, which proves the claim.

Before proceeding to the definition of entanglement for mixed states we shall mention a very useful tool in the description of entanglement for bipartite systems. This is the so-called *Schmidt decomposition.* Some of its properties will play an outstanding role in this thesis thus we will give a proof of it and its important properties later in Section 2.6.

**Lemma 1.2.** Let $|\psi\rangle = \sum_{i,j=1}^{N,M} c_{ij}|a_ib_j\rangle \in \mathcal{H}_A \otimes \mathcal{H}_B$ be a vector in the tensor product of two Hilbert spaces. Then there exist an orthonormal basis $|\alpha_i\rangle$ of $\mathcal{H}_A$ and an orthonormal basis $|\beta_j\rangle$ of $\mathcal{H}_B$ such that

$$|\psi\rangle = \sum_{k=1}^{R} \lambda_k|\alpha_k\beta_k\rangle \tag{1.3}$$

holds, with positive real coefficients $\lambda_k$. The $\lambda_k$ are uniquely determined as the square roots of the eigenvalues of the matrix $CC^\dagger$, where $C = (c_{ij})$. The number $R$ is called the Schmidt rank of $|\psi\rangle$. If the $\lambda_k$ are pairwise different, also the $|\alpha_k\rangle$ and $|\beta_k\rangle$ are unique up to a phase.

[1] By the term "observables" we denote Hermitean operators acting in a Hilbert space $\mathcal{H}$. We also denote the set of all bounded linear operators by $\mathcal{B}(\mathcal{H})$.

Note that the pure product states correspond to the states of Schmidt rank one. In this sense, the Schmidt rank of a state can be used to extend the notion of product states and to measure the entanglement of a non-separable state.

### 1.1.2 Mixed states

In a more general situation we do not know the exact state of a quantum system. We only know that it is, with some probabilities $p_i$, in one of some states $|\phi_i\rangle \in \mathcal{H}$. This situation is described by a density matrix

$$\varrho = \sum_i p_i |\phi_i\rangle\langle\phi_i|, \quad \text{with} \quad \sum_i p_i = 1. \tag{1.4}$$

The density matrix is an element of $\mathcal{B}(\mathcal{H})$. In a given basis this density matrix or state is represented by a complex matrix. This matrix is positive semidefinite[2] and thus Hermitean, since all the operators $|\phi_i\rangle\langle\phi_i|$ are positive and Hermitean. Due to the condition on the $p_i$ also $Tr(\varrho) = 1$ holds. Conversely, any positive semidefinite operator of trace one can be interpreted as a density matrix of some state. This leads to a geometrical picture of the set of all states as a convex set. By this we mean that given two states $\varrho_1$ and $\varrho_2$, their convex combination $\varrho = \alpha\varrho_1 + (1-\alpha)\varrho_2$ with $\alpha \in [0;1]$ is again a state. This property holds also for combinations of more than two states. Given some $p_i \geq 0$ with $\sum_i p_i = 1$ then the convex combination $\sum_i p_i \varrho_i$ of some states is again a state. We call coefficients $p_i \geq 0$ with the property $\sum_i p_i = 1$ often *convex weights.* Now we can define separability and entanglement for mixed states according to Ref. [18].

**Definition 1.3.** Let $\varrho \in \mathcal{B}(\mathcal{H}_A \otimes \mathcal{H}_B)$ be a density matrix for a composite system. We say that $\varrho$ is a *product state* if there exist states $\varrho^A$ for Alice and $\varrho^B$ for Bob such that

$$\varrho = \varrho^A \otimes \varrho^B. \tag{1.5}$$

The state is called *separable*, if there are convex weights $p_i$ and product states $\varrho_i^A \otimes \varrho_i^B$ such that

$$\varrho = \sum_i p_i \varrho_i^A \otimes \varrho_i^B \tag{1.6}$$

holds. Otherwise the state is called *entangled.*

Physically, this definition discriminates between three scenarios: A product state is an uncorrelated state, where Alice and Bob own each a separate state. For non-product states there are two different kinds of correlations: Separable states are classically correlated: This means that for the production of a separable state only local operations and classical communication (LOCC) are necessary. Alice and Bob can, by classical communication, share a random number generator which produces the outcomes $i$ with probabilities $p_i$. For each of the outcomes, they can agree to produce the state $\varrho_i^A \otimes \varrho_i^B$ locally. By this procedure they produce the state

[2] A matrix $C = (c_{ij})$ is called positive semidefinite, written as $C \geq 0$, if for all vectors $\mathfrak{a}$ and $\mathfrak{b}$ the relation $\sum_{i,j} \mathfrak{a}_i c_{ij} \mathfrak{b}_j \geq 0$ holds.

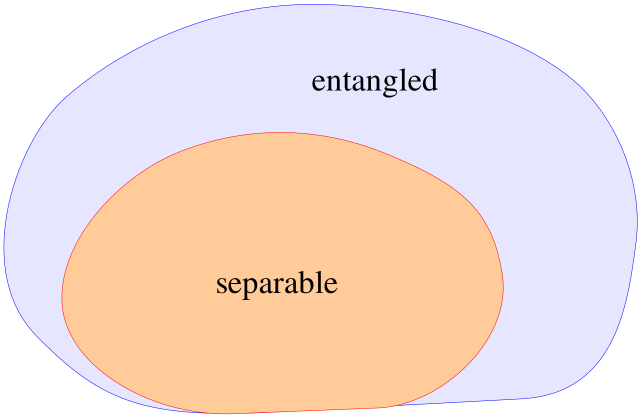

Figure 1.1. *Schematic picture of the set of all states as a convex set and the set of separable states as a convex subset.*

$\varrho = \sum_i p_i \varrho_i^A \otimes \varrho_i^B$. This procedure is by no means specific for quantum theory, which justifies the terminology. Otherwise, if a state is entangled, the correlations cannot origin from the classical procedure described above. In this sense entangled states are a typical feature of quantum mechanics.

For the theme of this thesis it is very important to note that the set of separable states is a convex set. This is clear from the definition of separability, obviously a convex combination of two separable states is again separable. Furthermore, the definition of separability implies that any separable state can be written as a convex combination of pure product states. Thus the set of separable states is the so called convex hull of the pure product states. This scheme is also shown in Figure 1.1.

Given the definition of entanglement and separability, it is very natural to ask whether a given density matrix is separable or entangled. This is the so-called separability problem. There are several criteria known, which imply separability or entanglement of a state. However, up to now, no general solution for the separability problem is known. It is one of the main aims of this thesis to develop criteria for entanglement, which are directly related to measurement data, and are thus implementable in experiments.

Before proceeding with the presentation of different separability criteria, it is useful to define some notations for the simplest bipartite quantum system. Let us assume that Alice owns a two-level quantum system, called a qubit, with the basis vectors $|0\rangle$ and $|1\rangle$[3]. If Bob owns also a qubit the whole space is four-dimensional, a basis is given by $\{|00\rangle, |01\rangle, |10\rangle, |11\rangle\}$. An important set of entangled states are the

[3] We always denote by $|0\rangle = |z^+\rangle$ (resp. $|1\rangle = |z^-\rangle$) the eigenvectors of the Pauli matrix $\sigma_z$ corresponding to the eigenvalue $+1$ $(-1)$. The eigenvectors of the other Pauli matrices, $\sigma_x$ and $\sigma_y$, are denoted by $|x^\pm\rangle = (|0\rangle \pm |1\rangle)/\sqrt{2}$ and $|y^\pm\rangle = (|0\rangle \pm i|1\rangle)/\sqrt{2}$.

so-called Bell states. They form an orthonormal basis and are given by

$$\begin{aligned} |\psi^+\rangle &= \frac{1}{\sqrt{2}}(|01\rangle + |10\rangle), \qquad |\psi^-\rangle = \frac{1}{\sqrt{2}}(|01\rangle - |10\rangle), \\ |\phi^+\rangle &= \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle), \qquad |\phi^-\rangle = \frac{1}{\sqrt{2}}(|00\rangle - |11\rangle). \end{aligned} \tag{1.7}$$

These states are maximally entangled, in the sense that any other entangled state can be produced from each of them by local means. The state $|\psi^-\rangle$ is also often called the singlet state. An important separable state it the maximally mixed state $\varrho = \mathbb{1}/4$, which represents white noise.

### 1.1.3 Criteria for entanglement

Here we want to give some criteria for entanglement or separability, which play a crucial role in this thesis.

Let us start with the criterion of the *partial transposition.* We can expand any density matrix of a composite quantum system in a chosen product basis:

$$\varrho = \sum_{i,j}^{N} \sum_{k,l}^{M} \rho_{ij,kl} |i\rangle\langle j| \otimes |k\rangle\langle l|. \tag{1.8}$$

Given this decomposition, we can define the partial transposition of $\varrho$ as the transposition with respect to one subsystem. Thus we have two partial transpositions. The partial transposition with respect to Alice is given by

$$\varrho^{T_A} = \sum_{i,j}^{N} \sum_{k,l}^{M} \rho_{ji,kl} |i\rangle\langle j| \otimes |k\rangle\langle l| \tag{1.9}$$

and similarly we can define $\varrho^{T_B}$ by exchanging $k$ and $l$ instead of the $i$ and $j$. Note that the partial transposition is related to the usual transposition by $\varrho^T = (\varrho^{T_A})^{T_B}$ and thus $\varrho^{T_B} = (\varrho^{T_A})^T$.

It is also worth mentioning that the partial transposition depends on the product basis in which it is performed. But one can show that its spectrum does not depend on the basis[4]. We say a density matrix $\varrho$ has a *positive partial transpose* (or: the matrix is PPT) if

$$\varrho^{T_A} \geq 0 \Leftrightarrow \varrho^{T_B} \geq 0 \tag{1.10}$$

holds. If a matrix is not PPT, we call it NPT. Now we can formulate the PPT criterion, originally introduced in Ref. [19].

**Theorem 1.4** (PPT Criterion). Let $\varrho$ be a bipartite separable state. Then the partial transposition with respect to each subsystem is positive.

[4] Note that this also holds for the normal transposition. The transpose of a matrix depends on the basis, in which it is transposed, but the spectrum of the transpose does not.

*Proof.* This fact follows directly from the definition of separability in Eq. (1.6). □

This theorem endows us with a very strong criterion for the detection of entanglement. For a given density matrix we can easily calculate the partial transpose and compute its spectrum. If we find negative eigenvalues we can conclude that the state is entangled. Given this result, the question arises if this criterion is also sufficient for separability, *i.e.*, whether $\varrho^{T_A} \geq 0$ implies separability. As it was shown already shortly after the discovery of the PPT criterion, this is the case only in low dimensional systems:

**Theorem 1.5.** If $\varrho$ is a state in a $2 \times 2$ or $2 \times 3$ system, then $\varrho^{T_A} \geq 0$ implies that $\varrho$ is separable. In other dimensions this is not the case.

*Proof.* For the proof of the $2 \times 2$ or $2 \times 3$ case, see [20]. For the first counterexample in an $2 \times 4$ system see [21]. We will discuss this in the context of bound entanglement later in Section 1.1.5. □

Now we can mention two other criteria, which play a role in this thesis. The first one is the *range criterion:*

**Theorem 1.6** (Range criterion). A state $\varrho$ is entangled if there is no family of product vectors $|a_i b_i\rangle$ such that the set $\{|a_i b_i\rangle\}$ spans the range of $\varrho$ as well as the set $\{|a_i^* b_i\rangle\}$[5] spans the range of $\varrho^{T_A}$.

*Proof.* It is easy to see that for separable states such a basis by definition exists. See also Ref. [21]. □

Another important criterion is the *majorization criterion:*

**Theorem 1.7** (Majorization criterion). Let $\varrho$ be separable and let $\varrho_A = Tr_B(\varrho)$ be the reduced state with respect to Alice. Denote by $\mathcal{P} = (p_1, p_2, ...)$ the decreasingly ordered eigenvalues of $\varrho$ and by $\mathcal{Q} = (q_1, q_2, ...)$ the decreasingly ordered eigenvalues of $\varrho_A$. Then

$$\sum_{i=1}^{k} p_i \leq \sum_{i=1}^{k} q_i \tag{1.11}$$

holds for all $k$. The same inequality holds, when $\varrho_A$ is replaced by the reduced density matrix of the second system $\varrho_B = Tr_A(\varrho)$.

*Proof.* This fact was proven in Ref. [22]. □

Later in this thesis we will discuss how this criterion can be extended if not the eigenvalues of the density matrix, but the probabilities of the outcomes of a measurement are taken into account.

Of course, there are much more entanglement criteria than the three presented here. But they do not play such an important role in the present thesis. However, many of them are worth to be mentioned: For two qubits also the reduction criterion is necessary and sufficient for separability [23]. It is also possible to characterize

[5] The symbol $|a^*\rangle$ denotes the vector resulting when all coefficients of $|a\rangle$ in a certain basis are complex conjugated. Note that $|a^*\rangle\langle a^*| = |a\rangle\langle a|^T$.

separability completely in terms of positive, but not completely positive maps [20]. However, the characterization of all positive, but not completely positive maps is still an open problem in mathematics [17]. For the detection of bound entanglement the so-called cross norm or realignment criterion has turned out to be useful [24, 25]. Recently, there have been several approaches to detect entanglement with an algorithmic approach, using the technique of semidefinite programming [26, 27].

Other approaches to detect entanglement use special observables, they are called either Bell inequalities, or entanglement witnesses. Bell inequalities will be discussed in Section 1.3. Entanglement witnesses will be discussed now.

### 1.1.4 Entanglement witnesses

The criteria from the previous section have all something in common: they all assume that the density matrix is already known. They all require applying certain operations to a density matrix, to decide whether the state is entangled or not[6]. There is, however, a necessary and sufficient entanglement criterion in terms of directly measurable observables. These are the so called entanglement witnesses [20, 29], which we introduce now.

**Definition 1.8.** An observable $\mathcal{W}$ is called an *entanglement witness* (or witness for short), if

$$\begin{aligned} Tr(\mathcal{W}\varrho_s) &\geq 0 \text{ for all separable } \varrho_s \\ Tr(\mathcal{W}\varrho_e) &< 0 \text{ for at least one entangled } \varrho_e \end{aligned} \tag{1.12}$$

holds. Thus, if we measure $Tr(\mathcal{W}\varrho) < 0$ we know for sure that the state $\varrho$ is entangled. We call a state with $Tr(\mathcal{W}\varrho) < 0$ to be detected by $\mathcal{W}$.

This definition has a clear geometrical meaning. The expectation value of an observable depends linearly on the state. Thus, the set of states where $Tr(\mathcal{W}\varrho) = 0$ holds is a hyperplane in the set of all states, cutting this set into two parts. In the part with $Tr(\mathcal{W}\varrho) > 0$ lies the set of all separable states, the other part (with $Tr(\mathcal{W}\varrho) < 0$) is the set of state detected by $\mathcal{W}$. This scheme is depicted in Figure 1.2.

From this geometrical interpretation it follows that all entangled states can be detected by witnesses:

**Theorem 1.9.** For each entangled state $\varrho_e$ there exist an entanglement witness detecting it.

*Proof.* This theorem was proved in [20]. The idea for the proof relies on the fact that the set of separable states is a convex and closed. Thus, for a point outside this set, there is a hyperplane separating this point from the convex set[7]. □

[6] But note that in principle one can decide whether a state violates the PPT criterion or not without knowing the state completely. For a scheme to do that see Ref. [28].

[7] For finite-dimensional spaces with a scalar product this is easy to derive. The statement holds also for infinite-dimensional Banach spaces, where the proof requires the Hahn-Banach-theorem [30].

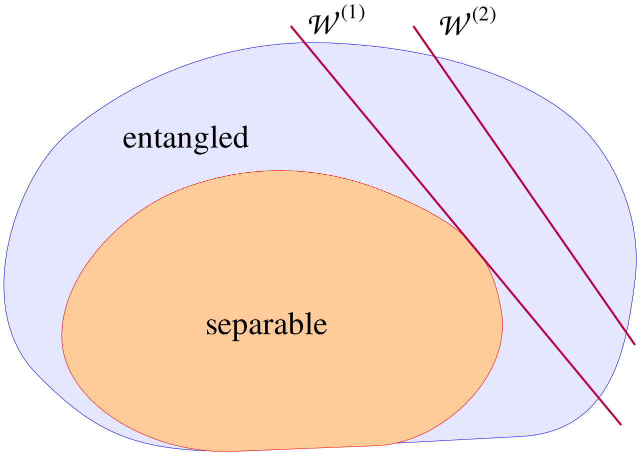

Figure 1.2. *Schematic picture of the set of all states and the set of separable states as nested convex sets and two witnesses, $\mathcal{W}^{(1)}$ and $\mathcal{W}^{(2)}$. The red lines represent the hyperplanes, where $Tr(\mathcal{W}\varrho) = 0$. The witness $\mathcal{W}^{(1)}$ is finer than the witness $\mathcal{W}^{(2)}$.*

Although this theorem ensures us that any state can be detected with an entanglement witness, the task remains to construct witnesses. This is not an easy task, and the characterization of all witnesses is, in general, a very demanding and unsolved problem. Solving this problem would also solve the separability problem.

To give an example how a witness can be constructed in a special case, let us take a state $\varrho_e$ which is NPT. Then there is a negative eigenvalue $\lambda_- < 0$ of $\varrho_e^{T_A}$ and a corresponding eigenvector $|\eta\rangle$. Now $\mathcal{W} = |\eta\rangle\langle\eta|^{T_A}$ is a witness, detecting $\varrho_e$ : We have $Tr(\mathcal{W}\varrho_e) = Tr(|\eta\rangle\langle\eta|^{T_A}\varrho_e) = Tr(|\eta\rangle\langle\eta|\varrho_e^{T_A}) = \lambda_- < 0$ and $Tr(\mathcal{W}\varrho_s) = Tr(|\eta\rangle\langle\eta|\varrho_s^{T_A}) \geq 0$ for all separable $\varrho_s$, since they are PPT.

Every entanglement witness can, by definition, detect some entangled states. However, some witnesses are better in this task, *i.e.*, one can *optimize* entanglement witnesses: A witness $\mathcal{W}^{(1)}$ is called *finer* than another witness $\mathcal{W}^{(2)}$ if $\mathcal{W}^{(1)}$ detects all the states detected by $\mathcal{W}^{(2)}$ and also some states in addition [31]. A witness $\mathcal{W}^{(1)}$ is called optimal, if there is no other witness finer than $\mathcal{W}^{(1)}$. The fact that a witness can be finer then another one has again a geometrical meaning, illustrated in Figure 1.2. Note that a necessary condition for a witness to be optimal is that it "touches" the set of separable states, *i.e.*, there exist a separable state $\varrho$ with $Tr(\mathcal{W}\varrho) = 0$. To optimize a given witness is, in general, a formidable task.

Entanglement witnesses provide us with an entanglement criterion, which is directly related to expectation values of observables. Thus they are good tools to detect entanglement in real experimental situations. Since they are nonlocal observables, some modifications are necessary, before they can be implemented in an real experiment. To close this gap between the witnesses as theoretical constructs and as observables measured in a laboratory is one of the main topics of this thesis.

### 1.1.5 Bound entanglement

For many tasks in quantum information theory like teleportation or cryptography one needs maximally entangled two-qubit states, *i.e.*, singlets. In the real world, however, noise is unavoidable, thus only mixed states are available. So one has to deal with the question, if and how one can create a singlet state out of some given mixed state. This leads to the problem of the so-called distillation of entanglement. It can be posed a follows: Assume that there is an arbitrary, but finite, number of copies of an entangled quantum state $\varrho$ distributed between Alice and Bob. Can they perform local operations on the states, assisted by classical communication, such that at the end they share a singlet state? If this is the case, we call the state $\varrho$ distillable, otherwise we call it undistillable or *bound entangled.*

It is not clear from the beginning that one can distill quantum states at all. However, the first distillation protocols were derived in [32, 33], showing that it is in principle possible. The question whether a given state is distillable is of course difficult to decide, since there is no restriction to a specific kind of local operations, as well as to the number of copies. The failure of a special protocol is not enough to conclude that a state is undistillable. However, some simple criteria is known:

**Theorem 1.10.** If a bipartite state is PPT, then the state is undistillable. If a state violates the majorization criterion, then the state is distillable.

*Proof.* For the sufficient condition to be undistillable, see Ref. [34]. For the sufficient condition to be distillable, see Ref. [35]. □

Although this gives a simple necessary criterion for distillability, it leaves two questions: First, the question arises, if the property of a state to be NPT already suffices to be distillable. This problem is note finally decided. However, it has been conjectured that this is not the case. Some evidence for this conjecture has given in [36, 37].

The other interesting question lies in the characterization of PPT entanglement. We know already, that PPT entangled states cannot exist in $2 \times 2$ and $2 \times 3$ systems. But in other dimensions they exists, the first example of them has been given in [21], for further examples see Refs. [38, 39]. Physically, the Theorem 1.10 states that PPT entangled states represent a weak form of entanglement, since bound entangled states are useless for certain tasks. However, for some tasks they can be used: In Ref. [40] an example has been given how bound entangled states can be used for establishing a secret key in quantum cryptography.

Nevertheless, all these results indicate that bound entangled states are a interesting object of study. Their construction, characterization and detection is a challenging and important task.

## 1.2 Multipartite entanglement

Here, we want to discuss the structure of entanglement when more than two parties are involved. It will turn out that this structure is much more rich than the structure of entanglement in the bipartite case. Especially, the difference between *genuine multipartite* entanglement and lower classes of entanglement becomes important. We first discuss the case of three qubits and then the case of general multipartite systems.

### 1.2.1 Three qubits

Let us first consider pure states. There are two classes of pure states which are not genuinely tripartite entangled: The *fully separable* states, which can be written as

$$|\phi^{\mathrm{fs}}\rangle_{A|B|C} = |\alpha\rangle_A \otimes |\beta\rangle_B \otimes |\gamma\rangle_C, \tag{1.13}$$

and the *biseparable* states which can be written as a product state in the bipartite system, which is created, if two of the three qubits are grouped together to one party. There are three possibilities of grouping two qubits together, hence there are three classes of biseparable states. One example is

$$|\phi^{\mathrm{bs}}\rangle_{A|BC} = |\alpha\rangle_A \otimes |\delta\rangle_{BC}. \tag{1.14}$$

The other possibilities read $|\phi^{\mathrm{bs}}\rangle_{B|AC} = |\beta\rangle_B \otimes |\delta\rangle_{AC}$ and $|\phi^{\mathrm{bs}}\rangle_{C|AB} = |\gamma\rangle_C \otimes |\delta\rangle_{AB}$. Here, $|\delta\rangle$ denotes a two party state, which might be entangled. The genuine tripartite entangled states are the states which are neither fully separable nor biseparable. But still these states can be divided into two inequivalent classes in the following way:

Given two three-qubit states, $|\phi\rangle$ and $|\psi\rangle$, one can ask whether it is possible to transform $|\phi\rangle$ into $|\psi\rangle$ with local operations and classical communication, without requiring that this can be done with unit probability. These operations are called stochastic local operations and classical communication (SLOCC). It turns out [41] that $|\phi\rangle$ can be transformed into $|\psi\rangle$ iff there exist invertible operators $A, B, C$, acting on the space of one qubit with

$$|\psi\rangle = A \otimes B \otimes C|\phi\rangle. \tag{1.15}$$

Since the operators $A, B, C$ are invertible, this defines an equivalence relation with a clear physical meaning. Surprisingly, it was proved in [41] that there are two different equivalence classes of genuine tripartite entangled states which cannot be transformed into another by SLOCC. One class, the class of GHZ states is represented by the GHZ state

$$|GHZ_3\rangle = \frac{1}{\sqrt{2}}(|000\rangle + |111\rangle), \tag{1.16}$$

the other class, the class of W states can be transformed via SLOCC into

$$|W\rangle = \frac{1}{\sqrt{3}}(|100\rangle + |010\rangle + |001\rangle. \tag{1.17}$$

In this sense there are two different classes of tripartite entanglement. There are much more pure GHZ class states than W class states: By local unitary operations one can transform any pure three-qubit state into[8]

$$|\psi\rangle = \lambda_0|000\rangle + \lambda_1 e^{i\theta}|100\rangle + \lambda_2|101\rangle + \lambda_3|110\rangle + \lambda_4|111\rangle, \tag{1.18}$$

where $\lambda_i \geq 0, \sum_i \lambda_i^2 = 1$ and $\theta \in [0;\pi]$, see Ref. [43]. Thus, six real parameters are necessary to characterize the nonlocal properties of a pure state. For th W class states, however, $\theta = \lambda_4 = 0$ holds, which shows that they are a set of measure zero in the set of all pure states.

Physically, there are also differences between the two classes: On the one hand, the GHZ state $|GHZ_3\rangle$ is maximally entangled and a generalization of the Bell states of two qubits. For instance, for the most known Bell inequalities the violation is maximal for GHZ states [44]. On the other hand, the entanglement of the W state $|W\rangle$ is more robust against particle losses: If one particle is lost in the GHZ state, the state $\varrho_{AB} = Tr_C(|GHZ_3\rangle\langle GHZ_3|)$ is separable, for the W state the resulting reduced density matrix $\varrho_{AB} = Tr_C(|W\rangle\langle W|)$ is entangled.

Before turning to mixed states, it is important to note that the definition of true tripartite entanglement has again a physical interpretation: For the creation of pure true tripartite entangled states the participation and common interaction of all three parties is necessary. This is not true for biseparable states, there only two parties have to interact.

The classification of mixed states according to Refs. [45, 46] is similar to the definition of bipartite entanglement via convex combinations. We define a mixed state $\varrho^{\text{fs}}$ as fully separable if $\varrho^{\text{fs}}$ can be written as a convex combination of fully separable pure states, *i.e.* if there are convex weights $p_i$ and fully separable states $|\phi_i^{\text{fs}}\rangle$ such that we can write

$$\varrho^{\text{fs}} = \sum_i p_i |\phi_i^{\text{fs}}\rangle\langle\phi_i^{\text{fs}}|. \tag{1.19}$$

A state $\varrho^{\text{bs}}$ which is not fully separable is called biseparable if it can be written as a convex combination of biseparable pure states:

$$\varrho^{\text{bs}} = \sum_i p_i |\phi_i^{\text{bs}}\rangle\langle\phi_i^{\text{bs}}|. \tag{1.20}$$

Note that the biseparable states $|\phi_i^{\text{bs}}\rangle$ might be biseparable with respect to different partitions. Of course, one can define three classes of biseparable mixed states which are biseparable with respect to a fixed partition.

Finally, $\varrho$ is fully entangled if it is neither biseparable nor fully separable. There are again two classes of fully entangled mixed states: As fully entangled state belongs to the W class if it can be written as a convex combination of W-type pure states

$$\varrho^{\text{w}} = \sum_i p_i |\phi_i^{\text{w}}\rangle\langle\phi_i^{\text{w}}|, \tag{1.21}$$

otherwise it belongs to the GHZ class. One can show that the W class forms a convex set inside the GHZ class. Also, in contrast to the case of pure states, the set of mixed W class states is not of measure zero compared to the GHZ class [46].

[8] This is a generalization of the Schmidt decomposition to three qubits, see [42] for generalizations to more than three parties.

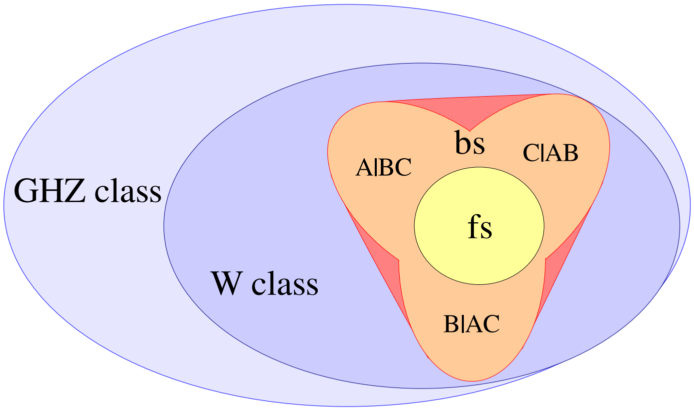

Figure 1.3. *Schematic picture of the structure of mixed states for three qubits. The convex set of all fully separable states (fs) is a subset of the set of all biseparable states (bs). The biseparable states are the convex combinations of the biseparable states with respect to fixed partitions, adumbrated by the three different leafs. Outside are the tripartite entangled states, the W class and the GHZ class. See text for further details.*

Again, this classification can be drawn in a schematic picture of nested convex sets, see Figure 1.3. Here, it is important to note that this picture is, of course, only a schematic picture which does not take all properties into account. For instance, it has been shown that there exist states which are biseparable with respect to each fixed partition, however, they are not fully separable [38, 47]. This counterintuitive behavior leads also to some surprising applications, like the possibility of "distributing entanglement using separable states" [48].

The question to which class a given mixed state belongs, is, as the separability problem for the bipartite case, not easy to decide. Necessary criteria for a state to be biseparable have been given in Refs. [45, 49, 50]. Methods to distinguish between the mixed W class and GHZ class are, however, very rare. As we will see later, witnesses give the possibility to prove that a state belongs to the GHZ class. However, it is not clear how one can show that a state is tripartite entangled and belongs to the W class. This can not be done with witnesses, since they are designed to show that a state lies *outside* a convex set, they fail to prove that a state is *inside* a convex set.

Let us finally mention that it is also possible to define distillability for three and more qubits. In this case, one aims at the distillation of the GHZ state $|GHZ_3\rangle$. Criteria and protocols for distillation have been given in Refs. [45, 49–51].

### 1.2.2 General classes

Let us turn to the classification of entanglement for general multipartite systems. The way to identify different classes is similar to the three-qubit case: First, we distinguish different types of entanglement for pure states. Then we extend this definition to mixed states by considering convex combinations of pure states.

Let us assume that a pure $n$-partite state $|\psi\rangle$ is given. We call this state *fully separable* if it is a product state of all parties, *i.e.*, if

$$|\psi\rangle = \bigotimes_{i=1}^{n} |\phi_i\rangle \tag{1.22}$$

holds. A mixed state is called fully separable if it can be written as a convex combination of pure fully separable states. If a pure state is not fully separable, it contains some entanglement. Again, as in the three-qubit case, this does not have to be true $n$-partite entanglement. Thus we call a pure state *$m$-separable*, with $1 < m < n$, if there exist a splitting of the $n$ parties into $m$ parts $P_i$ such that

$$|\psi\rangle = \bigotimes_{i=1}^{m} |\phi_i\rangle_{P_i} \tag{1.23}$$

holds. Note that an $m$-separable state still may contain some entanglement. Again, we call mixed states $m$-separable, if they can be written as convex combinations of pure $m$-separable states, which might belong to different partitions. Finally, we call a state truly $n$-partite entangled when it is neither fully separable, nor $m$-separable, for any $m > 1$. Here it it worth mentioning that a further classification of the pure truly $n$-partite entangled states like in the three-qubit case is not straightforward. At least a classification via equivalence classes under SLOCC is not very useful, since it has been shown that already for four qubits there are infinitely many equivalence classes under SLOCC [52]. Another approach tries to identify different classes via a generalization of the Schmidt rank, but here is is difficult to determine in which class a pure state lies [53].

### 1.2.3 Witnesses for detection problems

Here, we want to emphasize that witnesses are again good tools to distinguish the different classes of multipartite entanglement. This is due to the fact that all the classes defined above are defined via convex combinations, thus they are convex sets.

Let us explain these construction in the case of three qubits, the generalization to more parties is straightforward. In the case of three qubits on can define witnesses for several detection problems. There are GHZ class witnesses $\mathcal{W}_{GHZ}$ which allow to detect mixed states belonging of the GHZ class. Thus their expectation value is positive on all fully separable, biseparable, and W-type states:

$$\begin{aligned} Tr(\mathcal{W}_{GHZ}\varrho) &< 0 \Rightarrow \varrho \text{ is in the GHZ class.} \\ Tr(\mathcal{W}_{GHZ}\varrho) &\geq 0 \Rightarrow \varrho \text{ is not detected.} \end{aligned} \tag{1.24}$$

Also other witnesses can be defined: Witnesses for tripartite entanglement, denoted by $\mathcal{W}_3$, have a positive expectation value on all biseparable states thus a negative

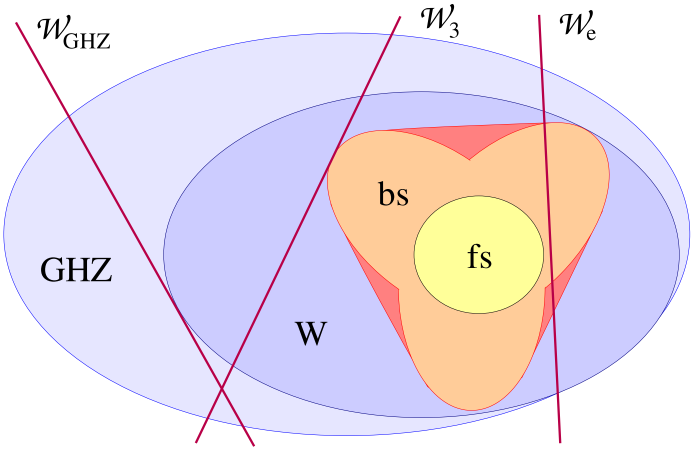

Figure 1.4. *Schematic picture of the set of mixed three-qubit states with three witnesses: The witness $\mathcal{W}_{GHZ}$ detects GHZ-type states, the witness $\mathcal{W}_3$ detects three-qubit entanglement, and the witness $\mathcal{W}_e$ rules out full separability.*

mean value indicates the presence of true tripartite entanglement. Witnesses for biseparable entanglement $\mathcal{W}_e$ have a positive mean value on all fully separable states, thus a negative expectation value is still a signature of entanglement, which might be only biseparable entanglement. All the types of witnesses can be cast into a schematic picture, this is shown in Figure 1.4. Finally, it is worth to stress that the interesting case in experiments is, of course, to detect true tripartite and GHZ class entanglement. Thus the construction of witnesses for this task is the main problem this thesis deals with.

## 1.3 Bell inequalities

Finally, we want to recall some facts about Bell inequalities, since formal similar, but conceptual different inequalities play an important role in the course of this thesis. For reviews on this subject see [54–56]

First, it is important to stress that Bell inequalities do not rely on the formalism and the physical content of quantum theory. They are independent of quantum mechanics. Bell inequalities are inequalities for the probabilities of the outcomes of some measurements, which have to be obeyed when the physical processes are describable by a local realistic or, synonymously, by a local hidden variable model. Let us discuss the assumptions which underly these models.

Consider that we have a system of two parties shared by Alice and Bob. We assume that these systems are space-like separated. Alice and Bob can perform some measurements $\tilde{A}_1, \tilde{A}_2, ..., \tilde{A}_N$, respectively, $\tilde{B}_1, \tilde{B}_2, ..., \tilde{B}_N$. Each of these measurements may have $M$ different outcomes $\tilde{A}_j^k$, respectively, $\tilde{B}_j^k$. A physical theory has to predict the probabilities $P(\tilde{A}_j^k)$ and $P(\tilde{B}_j^k)$ of these outcomes, as well as the probabilities $P(\tilde{A}_j^k; \tilde{B}_l^m)$ of the correlated outcomes, if Alice and Bob measure simultaneously. These probabilities depend, of course, on the state, which is shared by Alice and Bob. Now there are three assumptions on these predictions, which characterize a local realistic theory:

1. Realism: The system Alice and Bob share contains the full information about the results of all possible experiments that can be performed on it. Thus, the system can be described by some hidden variables which determines the measurement results. The values of the hidden variables may be unknown, or only some probability distribution of them may be known. Then the predictions of the measurement results are only probabilistic. However, in principle the information about all results is contained in the system.
2. Locality: The probabilities of the outcomes of a measurement on Alice's side do not depend on the measurement Bob chooses. Of course, there might be correlations between the subsystems, *i.e.*, the outcomes of Alice may depend on the outcomes of Bob, if he has chosen one certain measurement. But the probabilities of the outcomes for Alice should not change, when Bob chooses a different measurement.
3. Free will: Alice and Bob can choose their measurements randomly and in an arbitrary manner. The physical theory does not determine the experiments they choose.

Under these assumptions, on can derive inequalities for the probabilities of the outcomes of the measurements. These are the so called Bell inequalities. The first equation of this type was derived in 1964 [57]. Nowadays, other formulations than the original ones are used[9]. To give an example, let us assume two measurements $\tilde{A}_1, \tilde{A}_2$ resp. $\tilde{B}_1, \tilde{B}_2$ on each party, with two outcomes $+1$ and $-1$ for each measurement. If we denote $E(\tilde{A}; \tilde{B}) = P(\tilde{A}^+, \tilde{B}^+) + P(\tilde{A}^-, \tilde{B}^-) - P(\tilde{A}^+, \tilde{B}^-) - P(\tilde{A}^-, \tilde{B}^+)$, then the so-called Clauser-Horne-Shimony-Holt inequality (CHSH inequality)

$$|E(\tilde{A}_1; \tilde{B}_1) + E(\tilde{A}_2; \tilde{B}_1) + E(\tilde{A}_1; \tilde{B}_2) - E(\tilde{A}_2; \tilde{B}_2)| \leq 2 \tag{1.25}$$

has to hold for all local realistic theories [58]. This inequality can be translated into quantum theory as follows: A measurement $\tilde{A}_i$ can be realized by a measurement of a spin in some direction $\vec{a}_i$. This is represented by the observable $A_i = \vec{a}_i \cdot \vec{\sigma}$. Thus the CHSH inequality reads:

$$|\langle A_1 \otimes B_1 \rangle_\varrho + \langle A_2 \otimes B_1 \rangle_\varrho + \langle A_1 \otimes B_2 \rangle_\varrho - \langle A_2 \otimes B_2 \rangle_\varrho| \leq 2. \tag{1.26}$$

One interesting fact is, that quantum theory violates this inequality: There are quantum states, like the two-qubit singlet state, where this inequality is violated

[9] The original Bell inequality assumes that a special quantum mechanical state, the singlet, is given and that this state allows a local hidden variable model. This leads to the Bell inequality, imposing a constraint on the correlations, which turns out to be violated by the singlet. However, this inequality is difficult to test in experiments, since one would have to be sure that one has produced a perfect singlet state. For other states, a violation of the original Bell inequality does not say anything about the existence or non-existence of a local hidden variable model.

for an appropriate choice of the measurement directions $A_i$ and $B_i$. Even more interesting is the fact that nature itself violates this inequality: the probabilities occurring in Eq. (1.25) are directly measurable quantities and violations of the CHSH inequality have been found for photonic systems [10, 59] as well as for atomic systems [60]. Thus, even if quantum theory is replaced by another theory, this can not be a local realistic theory.

What are the connections between Bell inequalities and entanglement? One can show that all separable states obey all Bell inequalities. Thus, a violation of a Bell inequality implies that the state is entangled. However, violation of a Bell inequality and entanglement are not the equivalent. It has been shown that there exist entangled states, which admit a local hidden variable model [18]. Thus, these states do not violate any Bell inequality. Also, it is not clear, whether Bell inequalities can detect bipartite bound entanglement, as a typical form of weak entanglement. No Bell inequality for these states is known up to now, and it has even been conjectured that these states always admit a local hidden variable model [54]. In a certain sense, Bell inequalities can be viewed as non optimal entanglement witnesses: they can detect entanglement, but they fail to detect all states.

In the multipartite case the discussion of Bell inequalities is more involved, but also for this case Bell inequalities are known. One example, for three parties is the Mermin inequality [61], which reads in its quantum mechanical formulation

$$\langle A_1 \otimes B_1 \otimes C_2 \rangle_\varrho + \langle A_1 \otimes B_2 \otimes C_1 \rangle_\varrho + \langle A_2 \otimes B_1 \otimes C_1 \rangle_\varrho - \langle A_2 \otimes B_2 \otimes C_2 \rangle_\varrho \leq 2 \quad (1.27)$$

where the $A_i, B_i$ and $C_i$ are again observables with two different outcomes ±1, e.g. spin measurements in a three-qubit system. This inequality has to be obeyed for all states allowing a local hidden variable model, where the locality conditions hold with respect to all three parties. Thus, a violation of this inequality implies that the state is not fully separable.

One could, however, weaken the locality requirement by assuming only locality with respect to a bipartite partition of the three parties. This is of course a weaker condition then full locality, thus the bounds on the correlation get weaker [62]: For this case, the right hand side of Eq. (1.27) is in general bounded only by $2\sqrt{2}$. Thus this bound has to be violated, in order to prove that a state carries true tripartite entanglement. This makes it very difficult to detect true $n$-partite entanglement with Bell inequalities [63].

# Chapter 2

# Witnessing multipartite entanglement

## 2.1 Overview

In this part of the thesis we want to show that entanglement witnesses are useful tools for the detection of true multipartite entanglement. To this aim we will construct witnesses and decompose them in a way that allows to measure them easily with the present technology. We will also report on the experimental implementation of witnesses for multipartite systems.

Before we start our discussion of witnesses, let us discuss what we understand by *detecting entanglement.* In the most general case we have many copies of an unknown composite quantum system with several parties in laboratory. By performing measurements on these systems we gain some information about the state given. From this information we want to conclude that the state is entangled. Thus, an entanglement detection scheme has to tell us the measurements which we should perform, as well as how to decide from the measured data whether the state is entangled or not. Since many experiments aim at the generation of entanglement and not at the production of separable states, we allow some asymmetry: We want to detect entanglement, not separability. Thus the main task is to give conditions on the measurement data which imply that the state was entangled.

There are some natural properties a scheme for the detection of entanglement should have. So let us start with listing them:

1. The scheme should allow to detect entanglement unambiguously. By this we mean that the conclusion that a given state is entangled does not rely in any sense on assumptions on the state given. Only measurement results are taken into account.

2. The scheme should be adaptable to special detection problems. In a real experiment one usually does not produce an arbitrary state. Typically, one tries to produce a special state, but this state will be affected by some unknown and unavoidable noise. Thus the measurements on has to perform should be suited for the special state one intends to produce.

3. The scheme should allow to detect even weakly entangled states. This implies that it should also be robust against noise.
4. The scheme should also be able to detect true multipartite entanglement and to distinguish between the different classes of multipartite entanglement.
5. For an experimental demonstration the scheme should be implementable with simple means. It should require as few measurements as possible.

This list is by no means exhaustive and only a small collection of natural requirements. Other conditions may be added, e.g. in experiments when only a restricted set of measurements can be performed. Let us discuss some often used schemes for the detection of entanglement in view of these conditions.

The simplest idea to detect the entanglement for an unknown state, is, of course, to determine the state via state tomography completely. Then one can try to apply the already known separability criteria. The tomography can be done with local measurements on each party, thus there is no limitation of this method in principle. But it is not very flexible: It does not take into account the requirement (2), *i.e.*, that an experimentalist has some clue about the state which is prepared. State tomography is also not economic: It requires a lot of measurements. For instance to determine a four-qubit state requires 81 correlated measurements which needs a huge experimental effort.

Bell inequalities are widely used to detect entanglement [11, 64]. But, as we have discussed already in the previous chapter, violation of a Bell inequality and entanglement are not the same. Especially there are entangled states which can be described by a local hidden variable model and can thus never be detected with a Bell inequality [18]. We also mentioned that Bell inequalities for multipartite states are not very robust against noise: For states affected with some noise, they are often only capable of detecting bipartite entanglement, but not true multipartite entanglement [63]. Even worse, there exist some pure, true multipartite entangled states for which no Bell inequality is known so far, which can prove the multipartite entanglement. Thus, in experiments which are aimed to produce these states, Bell inequalities are unable to show that one really produced the desired class of entanglement. Another disadvantage of Bell inequalities is that up to now there is no canonical way to construct a Bell inequality for a given state. All these facts clearly indicate that Bell inequalities have problems to fulfill the requirements (2), (3) and (4). A clear advantage of Bell inequalities lies in the fact that they are relatively easy to implement. Only local measurements are needed. Thus they fulfill requirement (5).

Another method for detecting entanglement is the so called entanglement visibility, which is also sometimes used in experiments [65, 66]. Here, the visibility of two or more particle interferences is measured, in order to check whether a state is a coherent superposition or a mixture of two terms. If this visibility is high enough, one one can conclude that the state is entangled [67]. However, it is not clear that this methods can be applied to all entangled states. Also, the proofs that a certain visibility guarantees entanglement, usually rely on the assumption that only white noise is mixed to the state under consideration and, again, on Bell inequalities [67].

Recently, there have been several proposals for the detection of entanglement without estimating the whole density matrix [28, 68]. These proposals are very attractive from a theoretical point of view, nevertheless, they are very difficult to realize experimentally. Their implementation requires collective measurements on several qubits or the construction of quantum gates and networks. This is not easy to fulfill with the present experimental techniques.

As already mentioned, the main purpose of this chapter is to show that entanglement witnesses are good tools for the detection of entanglement in the multipartite case. Thus, we have to argue that entanglement witnesses fulfill all the requirements from above. It is clear from the definition of an entanglement witness that a measured negative expectation value guarantees that the state is entangled, thus requirement (1) is fulfilled. Also, witnesses are very flexible and obey condition (2), since one can construct them for a given state. We will provide a method for this in the following. Then we will also see that the conditions (3) and (4) are satisfied.

At the first view, the last condition seems to be problematic. Witnesses are in general nonlocal observables, thus it seems that they are not easy to implement. But, as we will show, this problem can be solved by decomposing them into local observables. These local observables can be measured individually, and from these results the mean value of the witnesses can be calculated.

We proceed as follows:

In the first section we develop a simple method to construct an entanglement witness detecting the genuine multipartite entanglement of a given pure state. This is not the only method of constructing witnesses, but, as we will see, it is very well suited in many situations.

In the second section we describe how a given entanglement witness can be implemented with local measurements. The core of our idea is the decomposition of a nonlocal observable into local observables. We state the problem of decomposing an witness into an *optimal* local decomposition.

The third section deals with explicit examples of witnesses. We construct witnesses for interesting states of two, three and four qubits. We also calculate their local decompositions. Since the emphasis of this chapter lies in the detection of multipartite entanglement, we mainly discuss three- and four-qubit witnesses.

As already mentioned, the local decompositions can be optimized in the number of measurement settings. To prove that a given decomposition is optimal is, in general, very demanding. In Section 2.5 we introduce a general technique for these optimality proofs. With this we can prove the optimality for our two- and three-qubit witnesses.

The following Section 2.6 is of a technical nature. There, we discuss the singular value decomposition, the Schmidt decomposition and its usefulness for the calculation of overlaps. These results are needed for the construction of the witnesses as well as in the derivation of uncertainty relations in the next chapter.

In Section 2.7 we present a different method of constructing entanglement witnesses. These witnesses are by construction implementable with local measurements, thus no decomposition is needed. They also have the advantage that they usually require only two measurements for an implementation. They are, however, usually not so robust against noise as the other witnesses.

In the last Section 2.8 we shortly report on experiments performed by M. Bourennane and coworkers in the group of H. Weinfurter in Garching, where some witnesses of this chapter have been implemented. In these experiments three- and four-photon entanglement has been detected using entanglement witnesses.

Finally, in the conclusion of this chapter, we summarize our results and discuss further problems and directions for future research which are, in our opinion, worth to address.

## 2.2 Constructing witnesses

Let $|\psi\rangle$ be a pure $n$-partite quantum state which is truly $n$-partite entangled. We propose to take as a witness operator detecting the entanglement of $|\psi\rangle$ an observable of the form

$$\mathcal{W} := \alpha \mathbb{1} - |\psi\rangle\langle\psi|. \tag{2.1}$$

Since $\mathcal{W}$ should be a witness detecting $|\psi\rangle$, we have to guarantee that $Tr(\mathcal{W}|\psi\rangle\langle\psi|) < 0$ as well as $Tr(\mathcal{W}\varrho^{\text{bs}}) < 0$ for all biseparable $\varrho^{\text{bs}}$. This is fulfilled, if $1 > \alpha \geq \sup_{\varrho\in\text{bs}} Tr(|\psi\rangle\langle\psi|\varrho)$. Here, "bs" denoted the set of all biseparable states. Of course, the witness gets better if $\alpha$ is as small as possible. Thus we take $\alpha$ as the supremum from above. Since the pure biseparable states are the extremal points of the set of all biseparable states it suffices to take the supremum over all biseparable pure states:

$$\alpha := \sup_{|\phi\rangle\in\text{bs}} |\langle\phi|\psi\rangle|^2. \tag{2.2}$$

In order to calculate the value of $\alpha$ we first fix a bipartite splitting, and consider only states $|\phi\rangle \in \text{bs}_1$ which are product vectors with respect to this partition. We can write $|\psi\rangle$ in the Schmidt decomposition for this fixed partition, let $\lambda_1$ be the biggest Schmidt coefficient in this decomposition. Then it follows that

$$\lambda_1^2 = \sup_{|\phi\rangle\in\text{bs}_1} |\langle\phi|\psi\rangle|^2. \tag{2.3}$$

The proof is in a more general setting given in Section 2.6. Therefore, $\alpha$ is given by the square of the Schmidt coefficient which is maximal over *all* possible bipartite partitions of $|\psi\rangle$[1].

In a more general case, one might wish to construct a witness which allows to detect states outside a general convex set, which has not to be the set of biseparable states. This can be done in the same manner, although the calculation of the overlap

[1] Note that in the case of several qubits this overlap cannot be smaller than $1/2$ for the following reason: There are bipartite splittings which separate one qubit from the others. In this case only two Schmidt coefficients occur and they cannot be both smaller than $1/\sqrt{2}$.

is in general not as easy as for the biseparable states. We can summarize the results of this section:

**Proposition 2.1.** Let $CS$ be a convex set and let $|\psi\rangle\langle\psi| \notin CS$ be a state outside this set. Then a witness allowing the verification of $|\psi\rangle\langle\psi| \notin CS$ is given by

$$\mathcal{W} = \beta \mathbb{1} - |\psi\rangle\langle\psi| \text{ with } \beta := \sup_{\varrho \in CS} Tr(\varrho |\psi\rangle\langle\psi|). \tag{2.4}$$

If $CS$ is the set of biseparable states, we can compute $\beta$ by computing the Schmidt decompositions of $|\psi\rangle$ for all biseparable partitions and taking $\beta$ as the squared maximal Schmidt coefficient appearing in all these decompositions.

## 2.3 Decomposing witnesses

A witness is always a nonlocal observable, since it has always an entangled eigenstate. This comes from the fact that a witness must have a negative eigenvalue and since the witness should be positive on all product states, the corresponding eigenvector cannot be a product vector. In the most general case we have a witness on a tensor product $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B \otimes ... \otimes \mathcal{H}_Z$ of two or more finite-dimensional Hilbert spaces. In order to slenderize the notation we look here at the case that we have a bipartite $N \times N$-system: $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B$ with $\dim(\mathcal{H}_A) = \dim(\mathcal{H}_B) = N$. But all definitions in this section can be extended to more parties in an obvious manner.

In order to measure the expectation value of the witness locally, we have to decompose it into a sum of tensor products of operators acting on only one system, or, equivalently, we have to decompose it into a sum of projectors onto product vectors:

$$\mathcal{W} = \sum_{i=1}^{m} A_i \otimes B_i = \sum_{i=1}^{n} c_i |e_i\rangle\langle e_i| \otimes |f_i\rangle\langle f_i|. \tag{2.5}$$

By measuring the expectation value of the projectors $|e_i\rangle\langle e_i| \otimes |f_i\rangle\langle f_i|$ and adding the results with the weights $c_i$ this decomposition (2.5) can be measured locally. There are, of course, many possibilities of finding a decomposition like (2.5). So we have to optimize the decomposition in a certain sense. But there are several possibilities of defining an optimized decomposition.

At first sight one might try to minimize the number of product vectors corresponding to minimizing $n$ in (2.5). This optimization procedure is already known from the literature, it was considered in [69]. There it was shown that for a general operator acting on two qubits one needs five product vectors and also a constructive way of computing these product vectors was given.

However, since we want to construct an experimentally accessible scheme for entanglement detection it is natural to look for a decomposition where Alice and Bob have to perform the smallest number of measurements possible. By "measurements" we understand here von Neumann (or projective) measurements, since they can be easily implemented. Such a measurement for Alice corresponds to a choice of an orthonormal basis of $\mathcal{H}_A$, and Bob has to choose an orthonormal basis $\mathcal{H}_B$, too. So any operator of the form

$$\mathcal{M} = \sum_{k,l=1}^{N} c_{kl} |e_k\rangle\langle e_k| \otimes |f_l\rangle\langle f_l| \tag{2.6}$$

with $\langle e_s|e_t\rangle = \langle f_s|f_t\rangle = \delta_{st}$ and real $c_{kl}$ can be measured with only one collective setting of measurement devices of Alice and Bob. Alice and Bob can distinguish the states $|e_k f_l\rangle$, measure the probabilities of these states and add their results with the weights $c_{kl}$ to measure $\mathcal{M}$. We call an operator which can be measured with one measurement setting (like $\mathcal{M}$ in Eq. (2.6)) a *local von Neumann measurement* (LvNM).

Having understood what can be realized with one measurement setting, we can state another optimization strategy. We want to find a decomposition of the form

$$\mathcal{W} = \sum_{i=1}^{K} \sum_{k,l=1}^{N} c_{kl}^{i} |e_k^i\rangle\langle e_k^i| \otimes |f_l^i\rangle\langle f_l^i| \tag{2.7}$$

with $\langle e_s^i|e_t^i\rangle = \langle f_s^i|f_t^i\rangle = \delta_{st}$ and a minimal $K$. This $K$ is the number of collective measurement settings Alice and Bob have to use in order to measure $\mathcal{W}$. This optimization strategy is the aim we are considering in this thesis when we talk about "optimized" decompositions.

The reader should note that minimizing $m$ in (2.5) is not the same as our optimization strategy. Also minimizing the number of product vectors, *i.e.*, minimizing $n$ in (2.5) is not the same. This will become clear already in the next section, when we give a simple two-qubit example.

Also, it is important to note that with one local measurement setting one can never detect entanglement. This is due to the fact that any probability distribution for the measurement of one LvNM in the form of Eq. (2.6) can be caused by mixture of the eigenstates $|e_k f_l\rangle$, which is a separable state.

## 2.4 Examples of witnesses

### 2.4.1 Two qubits

In this section we want to discuss shortly the application of the concept from above to two-qubit systems. It is not our aim to give a complete discussion here, this is done in Ref. [70, I, III]. Let

$$|\psi^-\rangle := \frac{1}{\sqrt{2}} (|01\rangle - |10\rangle) \tag{2.8}$$

be a maximally entangled two-qubit singlet state. Then, following the lines of the previous sections one can easily compute that

$$\mathcal{W}_2 := \frac{1}{2}\mathbb{1} - |\psi^-\rangle\langle\psi^-| \tag{2.9}$$

is an entanglement witness, detecting the entanglement of the state $|\psi\rangle$. In fact, one can show that for this special choice of $|\psi\rangle$ this is even an optimal entanglement witness [70], if $|\psi\rangle$ is not an maximally entangled state, the construction of Section 2.2 does not yield an optimal entanglement witness. Applying now the decomposition into a minimal set of product vectors according to Ref. [69] yields a decomposition of $\mathcal{W}_2$ into a sum of five product vectors [I,III]. These vectors belong to four orthonormal product bases, thus four measurement settings are needed. But one can measure $\mathcal{W}_2$ also with three settings: Using different spin directions defined by $|z^+\rangle = |0\rangle; |z^-\rangle = |1\rangle; |x^\pm\rangle = (|0\rangle \pm |1\rangle)/\sqrt{2}; |y^\pm\rangle = (|0\rangle \pm i|1\rangle)/\sqrt{2}$ we have the decomposition

$$\begin{aligned}\mathcal{W}_2 &= \frac{1}{2}\big(|z^+z^+\rangle\langle z^+z^+| + |z^-z^-\rangle\langle z^-z^-| + |x^+x^+\rangle\langle x^+x^+| + \\ &\quad + |x^-x^-\rangle\langle x^-x^-| - |y^+y^-\rangle\langle y^+y^-| - |y^-y^+\rangle\langle y^-y^+|\big) \\ &= \frac{1}{4}\left(\mathbb{1}\otimes\mathbb{1} + \sigma_z\otimes\sigma_z + \sigma_x\otimes\sigma_x + \sigma_y\otimes\sigma_y\right). \end{aligned} \tag{2.10}$$

In the calculation of this decomposition the formula

$$A\otimes B + C\otimes D = \frac{1}{2}\big((A+C)\otimes(B+D) + (A-C)\otimes(B-D)\big) \tag{2.11}$$

is of great help, with this we can easily see that

$$\begin{aligned}|\psi^-\rangle\langle\psi^-| &= \frac{1}{2}(|01\rangle\langle 01| + |10\rangle\langle 10| - |01\rangle\langle 10| - |10\rangle\langle 01|) \\ &= \frac{1}{2}(|z^+z^-\rangle\langle z^+z^-| + |z^-z^+\rangle\langle z^-z^+| - |0\rangle\langle 1|\otimes|1\rangle\langle 0| - |1\rangle\langle 0|\otimes|0\rangle\langle 1|) \\ &= \frac{1}{2}(|z^+z^-\rangle\langle z^+z^-| + |z^-z^+\rangle\langle z^-z^+|) - \frac{1}{4}(\sigma_x\otimes\sigma_x + \sigma_y\otimes\sigma_y). \end{aligned} \tag{2.12}$$

Using now the fact that $\mathbb{1}\otimes\mathbb{1} = |z^+z^+\rangle\langle z^+z^+| + |z^+z^-\rangle\langle z^+z^-| + |z^-z^+\rangle\langle z^-z^+| + |z^-z^-\rangle\langle z^-z^-|$ we arrive at (2.10).

In Eq. (2.10) the witness is decomposed into six product vectors, but it requires only spin measurements into three directions: Alice and Bob have only to set up their Stern-Gerlach devices in the $x$-, $y$- and $z$-direction to measure $\mathcal{W}_2$. One can show that this decomposition is optimal:

**Proposition 2.2.** In a two-qubit system a decomposition of $\mathcal{W}_2$ of the form (2.7) requires at least three measurements, thus the decomposition (2.10) is optimal.

*Proof.* Since the optimality proofs are technical and similar for all witnesses we have relegated them to Section 2.5. □

### 2.4.2 Three qubits

As we have seen in the introduction there are several types of witnesses which are of interest in the three-qubit case. We will give now examples of them and discuss their properties. We will also compute local decompositions and address the question whether they are optimal or not. The proofs of optimality are relegated to Section 2.5.

Let us start with witnesses for the GHZ class. For this case a witness operator was already constructed in [46]. It is given by

$$\mathcal{W}_{GHZ} = \frac{3}{4}\mathbb{1} - |GHZ_3\rangle\langle GHZ_3|. \tag{2.13}$$

Thus, if $\varrho$ is a mixed state with $Tr(\varrho W_{GHZ}) < 0$ the state $\varrho$ belongs to the GHZ class. The construction of this witness follows Proposition 2.1 in Section 2.2. The value $3/4$ corresponds to the maximal squared overlap between the state $|GHZ_3\rangle$ and all pure W class states. This witness detects states of the type $\varrho(p) = p|GHZ_3\rangle\langle GHZ_3| + (1-p)\mathbb{1}/8$ for $p > 5/7 \approx 0.714$ as belonging to the GHZ class.

A decomposition of $\mathcal{W}_{GHZ}$ can be computed with similar methods as for the two-qubit case, it yields

$$\begin{aligned}\mathcal{W}_{GHZ} \quad = \quad & \frac{1}{8}\Big(5 \cdot \mathbb{1}\otimes\mathbb{1}\otimes\mathbb{1} - \mathbb{1}\otimes\sigma_z\otimes\sigma_z - \sigma_z\otimes\mathbb{1}\otimes\sigma_z - \sigma_z\otimes\sigma_z\otimes\mathbb{1} - \\ & -2\cdot\sigma_x^{\otimes 3} + \frac{1}{2}\cdot(\sigma_x+\sigma_y)^{\otimes 3} + \frac{1}{2}\cdot(\sigma_x-\sigma_y)^{\otimes 3}\Big).\end{aligned} \tag{2.14}$$

This witness can be measured with four collective measurement settings, namely the settings $\sigma_z^{\otimes 3}, \sigma_x^{\otimes 3}$ and $((\sigma_x \pm \sigma_y)/\sqrt{2})^{\otimes 3}$. Note that the last two settings are just measurements of a spin in the direction between the $x$ and $y$ directions. This decomposition is optimal:

**Proposition 2.3.** The witness (2.13) can not be measured with three LvNMs, *i.e.*, the decomposition (2.14) is optimal.

*Proof.* See Section 2.5. □

Now we turn to witnesses for detecting true tripartite entanglement. These witnesses are, of course designed for a special state. We focus on three interesting pure states, which one might want to detect.

Let us first assume that we want to detect tripartite entanglement for a three-qubit GHZ state. In this case, a good choice for a witness would be

$$\mathcal{W}_3^{(\mathrm{ghz})} = \frac{1}{2}\mathbb{1} - |GHZ_3\rangle\langle GHZ_3|. \tag{2.15}$$

The construction of this witness is similar as above, here $1/2$ is the maximal squared overlap between the state $|GHZ_3\rangle$ and all pure biseparable states. With this witness one can prove that the states $\varrho(p) = p|GHZ_3\rangle\langle GHZ_3| + (1-p)\mathbb{1}/8$ are for $p > 3/7 \approx 0.429$ true tripartite entangled. Of course, this witness can be measured with the same measurements from (2.14) and this decomposition is also optimal.

In case one wants to detect states in the vicinity of a W state, a witness would be:

$$\mathcal{W}_3^{(\mathrm{w})} = \frac{2}{3}\mathbb{1} - |W\rangle\langle W|, \tag{2.16}$$

This witness detects tripartite entanglement in states of the form $\varrho(p) = p|W\rangle\langle W| + (1-p)\mathbb{1}/8$ for $p > 13/21 \approx 0.619$. Its local decomposition is given by

$$\begin{aligned}\mathcal{W}_3^{(\mathrm{w})} &= \frac{1}{24}\Big(17 \cdot \mathbb{1} \otimes \mathbb{1} \otimes \mathbb{1} + 7 \cdot \sigma_z^{\otimes 3} \\ &+3 \cdot (\sigma_z \otimes \mathbb{1} \otimes \mathbb{1} + \mathbb{1} \otimes \sigma_z \otimes \mathbb{1} + \mathbb{1} \otimes \mathbb{1} \otimes \sigma_z) + \\ &+5 \cdot (\sigma_z \otimes \sigma_z \otimes \mathbb{1} + \sigma_z \otimes \mathbb{1} \otimes \sigma_z + \mathbb{1} \otimes \sigma_z \otimes \sigma_z) + \\ &-(\mathbb{1} + \sigma_z + \sigma_x)^{\otimes 3} - (\mathbb{1} + \sigma_z - \sigma_x)^{\otimes 3} \\ &-(\mathbb{1} + \sigma_z + \sigma_y)^{\otimes 3} - (\mathbb{1} + \sigma_z - \sigma_y)^{\otimes 3}\Big). \qquad (2.17)\end{aligned}$$

Here, only five correlated measurement settings are necessary, namely $\sigma_z^{\otimes 3}$ and $((\sigma_z \pm \sigma_i)/\sqrt{2})^{\otimes 3}; i = x, y$. This decomposition is also optimal:

**Proposition 2.4.** The witness $\mathcal{W}_3^{(\mathrm{w})}$ can not be measured with four measurement settings, *i.e.,* the decomposition (2.17) is optimal.

*Proof.* See Section 2.5. □

Another interesting three-qubit state is the state

$$|OC\rangle = \sqrt{\frac{2}{3}}|001\rangle - \sqrt{\frac{1}{6}}|010\rangle - \sqrt{\frac{1}{6}}|100\rangle. \qquad (2.18)$$

This state is of interest for the following reason: One can perform universal cloning by teleportation using a tripartite entangled state. The state which is needed for optimal cloning is just the state $|OC\rangle$ [71].

The state $|OC\rangle$ contains true tripartite entanglement and the squared overlap with the biseparable states can be calculated as described above, the result is 5/6. So we arrive at the witness

$$\begin{aligned}\mathcal{W}_3^{(\mathrm{oc})} &= \frac{5}{6}\mathbb{1} - |OC\rangle\langle OC| \\ &= \frac{1}{48}\Big(38 \cdot \mathbb{1} \otimes \mathbb{1} \otimes \mathbb{1} + 10 \cdot \sigma_z \otimes \sigma_z \otimes \sigma_z + 6 \cdot \mathbb{1} \otimes \mathbb{1} \otimes \sigma_z + \\ &+2 \cdot \sigma_z \otimes \sigma_z \otimes \mathbb{1} + 8 \cdot \sigma_z \otimes \mathbb{1} \otimes \sigma_z + 8 \cdot \mathbb{1} \otimes \sigma_z \otimes \sigma_z + \\ &-(\mathbb{1} + \sigma_z + \sigma_x) \otimes (\mathbb{1} + \sigma_z - \sigma_x) \otimes (\mathbb{1} + \sigma_z - 2\sigma_x) \\ &-(\mathbb{1} + \sigma_z - \sigma_x) \otimes (\mathbb{1} + \sigma_z + \sigma_x) \otimes (\mathbb{1} + \sigma_z + 2\sigma_x) \\ &-(\mathbb{1} + \sigma_z + \sigma_y) \otimes (\mathbb{1} + \sigma_z - \sigma_y) \otimes (\mathbb{1} + \sigma_z - 2\sigma_y) \\ &-(\mathbb{1} + \sigma_z - \sigma_y) \otimes (\mathbb{1} + \sigma_z + \sigma_y) \otimes (\mathbb{1} + \sigma_z + 2\sigma_y)\Big). \qquad (2.19)\end{aligned}$$

This witness requires five measurement settings, and again it can be shown to be optimal:

**Proposition 2.5.** The witness $\mathcal{W}_3^{(\mathrm{oc})}$ can not be measured with four measurement settings, thus the decomposition (2.19) is optimal.

*Proof.* See Section 2.5. □

Finally, let us mention that with similar methods as above one can also construct witnesses which detect just some entanglement, which might be only biseparable entanglement. For this purpose, one has to calculate the maximal squared overlap between a given state and the fully separable states. This in general a hard task, but for some states it has been done [72]. For instance, one can show that the maximal squared overlap between the state $|W\rangle$ and the fully separable states is 4/9. Thus, the observable

$$\mathcal{W}_e^{(\mathrm{w})} = \frac{4}{9}\mathbb{1} - |W\rangle\langle W|, \tag{2.20}$$

is a witness, ruling out full separability.

### 2.4.3 Four qubits

Since there is not a classification like the GHZ class and the W class for four qubits, the main task in this case is just to detect true four-partite entanglement. Again, there are several interesting states. We focus here on three states, namely the four-qubit GHZ state, the so-called four-qubit singlet state and the four-qubit cluster state.

The four-qubit GHZ state

$$|GHZ_4\rangle = \frac{1}{\sqrt{2}}(|0000\rangle + |1111\rangle) \tag{2.21}$$

plays a similar distinguished role in the discussion of local hidden variable theories for four parties as the three-qubit GHZ state for three parties. So it natural to start the discussion with this state. A witness for the four-qubit GHZ state is given by

$$\mathcal{W}_4^{(\mathrm{ghz})} = \frac{1}{2}\mathbb{1} - |GHZ_4\rangle\langle GHZ_4|, \tag{2.22}$$

and its local decomposition is

$$\begin{aligned}\mathcal{W}_4^{(\mathrm{ghz})} \quad = \quad & \frac{1}{32}\Big(16\cdot\mathbb{1}\otimes\mathbb{1}\otimes\mathbb{1}\otimes\mathbb{1} - (\mathbb{1}+\sigma_z)^{\otimes 4} - (\mathbb{1}-\sigma_z)^{\otimes 4} \\ & +(\sigma_x+\sigma_y)^{\otimes 4} + (\sigma_x-\sigma_y)^{\otimes 4} - 4\cdot\sigma_x^{\otimes 4} - 4\cdot\sigma_y^{\otimes 4}\Big).\end{aligned} \tag{2.23}$$

This witness only requires five measurement settings. This is only one more than the three qubit GHZ witness, thus it is probably the optimal decomposition. But we have no rigorous proof of it. We will explain the main difficulty in Section 2.5.

Another very interesting four qubit state is the so called *four-qubit singlet state* [73]:

$$|\Psi^{(4)}\rangle = \frac{1}{\sqrt{3}}\Big(|0011\rangle + |1100\rangle - \frac{1}{2}(|0110\rangle + |1001\rangle + |0101\rangle + |1010\rangle)\Big). \tag{2.24}$$

Let us briefly mention some of its remarkable properties: First, note that the state $|\Psi^{(4)}\rangle$ can be viewed as a superposition of a four-qubit GHZ state of the type $|GHZ_4'\rangle = (|0011\rangle + |1100\rangle)/\sqrt{2}$ and a product of two singlet states, one singlet between the first two parties and another between the third and the fourth party. The name "four-qubit singlet state" comes from the fact that $|\Psi^{(4)}\rangle$ is form invariant when the same change of basis is applied to all four qubits simultaneously, *i.e.,* under $U \otimes U \otimes U \otimes U$ transformations. From an experimental point of view, it is also remarkable that the state $|\Psi^{(4)}\rangle$ is relatively easy to produce by a second order process in spontaneous parametric down conversion. We will discuss this later in Section 2.8. Another fact is that $|\Psi^{(4)}\rangle$ violates some Bell inequalities, but up to now no Bell inequality is known, which can prove that the state $|\Psi^{(4)}\rangle$ contains true four-partite entanglement, one can only prove that the state is not fully separable [73, 74].

The last two properties indicate that the state $|\Psi^{(4)}\rangle$ is an ideal test bed for the experimental implementation of witness operators. In fact, the experimental demonstration of four-partite entanglement has been performed, we will shortly report on these experiments in Section 2.8. Here we only want to calculate the witness, it is given by:

$$\begin{aligned}
\mathcal{W}_4^{\Psi^{(4)}} &= \frac{3}{4}\mathbb{1} - |\Psi^{(4)}\rangle\langle\Psi^{(4)}| \\
&= \frac{1}{48}\Big(33 \cdot \mathbb{1}^{\otimes 4} - \sum_{i=x,y,z} \big[\sigma_i \otimes \sigma_i \otimes \mathbb{1} \otimes \mathbb{1} + \mathbb{1} \otimes \mathbb{1} \otimes \sigma_i \otimes \sigma_i - \sigma_i^{\otimes 4} \\
&\quad -2 \cdot (\sigma_i \otimes \mathbb{1} \otimes \sigma_i \otimes \mathbb{1} + \sigma_i \otimes \mathbb{1} \otimes \mathbb{1} \otimes \sigma_i + \\
&\quad +\mathbb{1} \otimes \sigma_i \otimes \sigma_i \otimes \mathbb{1} + \mathbb{1} \otimes \sigma_i \otimes \mathbb{1} \otimes \sigma_i)\big]\Big). \\
&\quad +3 \cdot (\sigma_x \otimes \sigma_x \otimes \sigma_y \otimes \sigma_y + \sigma_y \otimes \sigma_y \otimes \sigma_x \otimes \sigma_x + \sigma_x \otimes \sigma_x \otimes \sigma_z \otimes \sigma_z \\
&\quad +\sigma_z \otimes \sigma_z \otimes \sigma_x \otimes \sigma_x + \sigma_y \otimes \sigma_y \otimes \sigma_z \otimes \sigma_z + \sigma_z \otimes \sigma_z \otimes \sigma_y \otimes \sigma_y) \\
&\quad -(\sigma_x + \sigma_y)^{\otimes 4} - (\sigma_x - \sigma_y)^{\otimes 4} - (\sigma_x + \sigma_z)^{\otimes 4} \\
&\quad -(\sigma_x - \sigma_z)^{\otimes 4} - (\sigma_x + \sigma_z)^{\otimes 4} - (\sigma_y - \sigma_z)^{\otimes 4}.
\end{aligned} \tag{2.25}$$

Although this decomposition is a little bit longer, it requires only fifteen measurement settings.

The last four-qubit state we want to deal with is the so-called four-qubit *cluster state* [75]. In Section 2.7 we will give a brief introduction into the formalism and physical meaning of the cluster states, here we only write down the state. The four-qubit cluster state is given by

$$|\psi_4^{cl}\rangle = \frac{1}{2}\big(|0000\rangle + |0011\rangle + |1100\rangle - |1111\rangle\big). \tag{2.26}$$

One can calculate that the maximized squared overlap between the cluster state and the biseparable states equals 1/2, thus we arrive at the witness

$$\begin{aligned}
\mathcal{W}_4^{\psi_4^{cl}} &= \frac{1}{2}\mathbb{1} - |\psi_4^{cl}\rangle\langle\psi_4^{cl}| \\
&= \frac{1}{16}\big(7 \cdot \mathbb{1}^{\otimes 4} - \sigma_z^{\otimes 4} - \sigma_z \otimes \sigma_z \otimes \mathbb{1} \otimes \mathbb{1} - \mathbb{1} \otimes \mathbb{1} \otimes \sigma_z \otimes \sigma_z \\
&\quad -\mathbb{1} \otimes \sigma_z \otimes \sigma_x \otimes \sigma_x - \sigma_z \otimes \mathbb{1} \otimes \sigma_x \otimes \sigma_x + \mathbb{1} \otimes \sigma_z \otimes \sigma_y \otimes \sigma_y \\
&\quad +\sigma_z \otimes \mathbb{1} \otimes \sigma_y \otimes \sigma_y - \sigma_x \otimes \sigma_x \otimes \mathbb{1} \otimes \sigma_z - \sigma_x \otimes \sigma_x \otimes \sigma_z \otimes \mathbb{1} \\
&\quad +\sigma_y \otimes \sigma_y \otimes \mathbb{1} \otimes \sigma_z + \sigma_y \otimes \sigma_y \otimes \sigma_z \otimes \mathbb{1} - \sigma_x \otimes \sigma_y \otimes \sigma_x \otimes \sigma_y \\
&\quad -\sigma_x \otimes \sigma_y \otimes \sigma_y \otimes \sigma_x - \sigma_y \otimes \sigma_x \otimes \sigma_x \otimes \sigma_y - \sigma_y \otimes \sigma_x \otimes \sigma_y \otimes \sigma_x\big).
\end{aligned} \tag{2.27}$$

This decomposition requires nine correlated measurement settings.

## 2.5 Optimality proofs

In this section we present a method to prove that a given decomposition of a nonlocal observable into local measurements is optimal. Let us start with the simple case of two qubits:

**Proposition 2.2.** In a two-qubit system a decomposition of $\mathcal{W}_2 = 1/2 \cdot \mathbb{1} - |\psi^-\rangle\langle\psi^-|$ into local measurements of the form (2.7) requires at least three measurement settings.

*Proof.* Consider a decomposition requiring two measurements:

$$\mathcal{W}_2 = \sum_{i,j=1}^{2} c_{ij}^1 |A_i^1\rangle\langle A_i^1| \otimes |B_j^1\rangle\langle B_j^1| + \sum_{i,j=1}^{2} c_{ij}^2 |A_i^2\rangle\langle A_i^2| \otimes |B_j^2\rangle\langle B_j^2|. \tag{2.28}$$

We expand the right hand side as well as the left hand side of this equation in a product basis of the operator space. We write $\mathcal{W}_2 = \sum_{i,j=0}^3 \lambda_{ij}\ \sigma_i \otimes \sigma_j$ and have, according to (2.10), $(\lambda_{ij}) = \mathbb{1}/4$.

On the right hand side, we expand the first sum as $\sum_{i,j=1}^2 c_{ij}^1 |A_i^1\rangle\langle A_i^1| \otimes |B_j^1\rangle\langle B_j^1| = \sum_{i,j=0}^3 \mu_{ij}\ \ \sigma_i \otimes \sigma_j$. To do this, we write any projector 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 $\vec{s}_{A_1^1} = (1/2, s_1^A, s_2^A, s_3^A) = (1/2, \vec{\mathfrak{s}}_A^1)$ and $|A_2^1\rangle\langle A_2^1|$, since $|A_1^1\rangle$ and $|A_2^1\rangle$ are orthogonal, by $\vec{s}_{A_2^1} = (1/2, -s_1^A, -s_2^A, -s_3^A) = (1/2, -\vec{\mathfrak{s}}_A^1)$; $|B_1^1\rangle\langle B_1^1|$ can be written similarly. From this it follows that if we look at the $4 \times 3$ sub-matrix $\mu^{red}$ we have

$$\begin{aligned} \mu^{red} &:= \mu_{i=0,\dots,3;j=1,\dots,3} \\ &= c_{11}^1 (\vec{s}_{A_1^1})^T \cdot \vec{\mathfrak{s}}_B^1 - c_{12}^1 (\vec{s}_{A_1^1})^T \cdot \vec{\mathfrak{s}}_B^1 + c_{21}^1 (\vec{s}_{A_2^1})^T \cdot \vec{\mathfrak{s}}_B^1 - c_{22}^1 (\vec{s}_{A_2^1})^T \cdot \vec{\mathfrak{s}}_B^1 \\ &= (c_{11}^1 \cdot \vec{s}_{A_1^1} - c_{12}^1 \cdot \vec{s}_{A_1^1} + c_{21}^1 \cdot \vec{s}_{A_2^1} - c_{22}^1 \cdot \vec{s}_{A_2^1})^T \cdot \vec{\mathfrak{s}}_B^1. \end{aligned} \tag{2.29}$$

Thus, the reduced matrix can be written as $\mu_{ij}^{red} = \mathfrak{a}_i \mathfrak{b}_j$. In other words, it is of rank one. The corresponding sub-matrix from the second sum on the right hand side of (2.28) is for the same reasons also of rank one. But then we arrive at a contradiction: The corresponding sub-matrix from $(\lambda_{ij}) = \mathbb{1}/4$ is of rank three and no matrix of rank three can be written as a sum of two matrices of rank one. □

Having realized that the proof relied on the special structure of a sub-matrix for one LvNM, we can generalize it to multi-qubit systems. This leads to the notion of the *tensor rank* which we will introduce now.

**Definition 2.6.** Let $\tau = (\tau_{ij\dots n})$ be an $n$-dimensional tensor over some field $F$. If we can find vectors $\mathfrak{a}, \mathfrak{b}, \dots, \mathfrak{n}$ such that

$$\tau_{ij\dots n} = \mathfrak{a}_i \cdot \mathfrak{b}_j \cdot \dots \cdot \mathfrak{n}_n \tag{2.30}$$

holds, we call $\tau$ a tensor of rank one. If $\tau$ is not of rank one, we write $\tau$ as a sum of rank one tensors $\eta_r$:

$$\tau = \sum_{r=1}^{R} \eta_r, \tag{2.31}$$

with a minimal $R$. Then $R$ is called the tensor rank of $\tau$, we write $\mathfrak{R}_F(\tau) := R$.

Note that this definition coincides with the rank of a matrix, if we consider matrices as two-dimensional tensors. We will discuss more properties of the tensor rank later. Then we will also argue, why the field $F$ is important in this definition. First, we can state:

**Theorem 2.7.** Let $\mathcal{W}$ be an $n$-qubit witness, expressed in the basis of the Pauli matrices as

$$\mathcal{W} = \sum_{i,j,\ldots,n=0}^{3} \lambda_{ij\ldots n}\, \sigma_i \otimes \sigma_j \otimes \ldots \otimes \sigma_n. \tag{2.32}$$

Define the $4 \times 3 \times \ldots \times 3$ tensor $\lambda^{red}$ by

$$\lambda^{red} := \lambda_{i=0,\ldots,3;j,\ldots,n=1,\ldots,3}. \tag{2.33}$$

Then a local measurement of $\mathcal{W}$ requires at least $\mathfrak{R}_{\mathbb{R}}(\lambda^{red})$ LvNMs[2].

*Proof.* We only have to show that the reduced tensor coming from a single LvNM has rank one. This follows directly from a calculation as in Eq. (2.29). □

Please note that this theorem only gives a lower bound of the number of LvNMs. It is not clear whether we can always find a decomposition into $\mathfrak{R}_{\mathbb{R}}(\lambda^{red})$ LvNMs. Now we can state some general properties of the tensor rank. For the proofs we have to refer to the literature.

**Proposition 2.8.** The tensor rank has the following properties:
(a) It depends on the underlying field, especially we have for a general real tensor:

$$\mathfrak{R}_{\mathbb{R}}(\tau) > \mathfrak{R}_{\mathbb{C}}(\tau). \tag{2.34}$$

(b) Up to now there is no canonical way to compute the tensor rank for a given tensor. In fact, one can prove that over finite fields the computation of the rank of a general three-dimensional tensor is NP complete, over $\mathbb{Q}$ it is NP hard.
(c) Define $r(k, l, \ldots, n)$ as the maximal rank over $\mathbb{R}$ which can appear for real $k \times l \times \ldots \times n$ tensors. The following facts are known:

$$\begin{aligned} r(2,2,2) &= 3, & (2.35)\\ r(3,3,3) &= 5, & (2.36)\\ r(4,3,3) &= 6, & (2.37)\\ r(k_1,\ldots,k_n) &\geq \frac{\prod_{i=1}^{n} k_i}{\sum_{i=1}^{n} k_i - n + 1}. & (2.38) \end{aligned}$$

[2] Of course, one can also define $\lambda^{red}$ with respect to another than the first subsystem, e.g. by taking the $3 \times 4 \times \ldots \times 3$ tensor.

*Proof.* For a discussion of property (a) see Ref. [76]. Property (b) was shown in [77]. Eqs. (2.35, 2.36, 2.37) were proved in [78], and Eq. (2.38) in [79]. □

In property (a) an important difference between matrices and tensors becomes manifest: For a given real matrix the rank does not change, if it is considered as a complex matrix. This is not the case for tensors. Although property (b) is in a certain sense disappointing, we will now prove the optimality of our three-qubit decompositions. Later we will discuss the problems in the four-qubit and in the general case.

**Proposition 2.3.** The witness $\mathcal{W}_{GHZ} = 3/4 \cdot \mathbb{1} - |GHZ_3\rangle\langle GHZ_3|$ can not be measured with three LvNMs, *i.e.*, the decomposition (2.14) is optimal.

*Proof.* We write the witness as: $\mathcal{W}_{GHZ} = 1/8 \cdot \sum_{i,j,k=0}^{3} \lambda_{ijk}\sigma_i \otimes \sigma_j \otimes \sigma_k$, and from (2.14) we can compute $\lambda_{ijk}$. Then we look as proposed at the $4 \times 3 \times 3$ tensor $\lambda^{red}$. Its components are given by

$$\begin{aligned}
\lambda_{0ij} &= \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix} := A^{(0)}, \quad \lambda_{1ij} = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} := A^{(1)}, \\
\lambda_{2ij} &= \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} := A^{(2)}, \quad \lambda_{3ij} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} := A^{(3)}.
\end{aligned}$$

To show that $\mathcal{W}_{GHZ}$ can not be measured with three settings we have to show that the tensor rank of $\lambda$ is bigger than three. Let us assume the contrary, *i.e.*, we have

$$\begin{aligned}
\lambda_{ijk}^{red} &= \sum_{r=1}^{3} \mathfrak{a}_i^{(r)} \mathfrak{b}_j^{(r)} \mathfrak{c}_k^{(r)} \\
&= \mathfrak{a}_i^{(1)} B_{jk}^{(1)} + \mathfrak{a}_i^{(2)} B_{jk}^{(2)} + \mathfrak{a}_i^{(3)} B_{jk}^{(3)}. \qquad (2.39)
\end{aligned}$$

This equation implies that the three $3 \times 3$ matrices $B^{(r)}, r \in \{1...3\}$ are of rank one and that $A^{(0)}, A^{(1)}$ and $A^{(2)}$ can be represented as linear combinations of the $B^{(r)}$. ($A^{(3)}$ does not matter here, since it is the null matrix.) Since the $A^{(i)}$ span a three-dimensional subspace in the space of all $3 \times 3$ matrices, the $B^{(r)}$ have to be linearly independent (as matrices) and have to span the same space. If this is true, any of the $B^{(r)}$ can be written as a linear combination of the $A^{(i)}$. But a general linear combination of the $A^{(i)}$ is of the form

$$\mathcal{A} = \begin{pmatrix} -\alpha & \beta & 0 \\ \beta & \alpha & 0 \\ 0 & 0 & \gamma \end{pmatrix}. \qquad (2.40)$$

This is of rank one if and only if $\alpha = \beta = 0$. Thus, we arrive at a contradiction, the $B^{(r)}$ cannot be linear independent. □

Similarly, we can prove the other cases:

**Proposition 2.4.** The witness $\mathcal{W}_3^{(\mathrm{w})}$ can not be measured with four measurement settings, *i.e.*, the decomposition (2.17) is optimal.

*Proof.* The proof is similar to the proof from above. We compute $\mathcal{W}_3^{(\mathrm{w})} = 1/24 \cdot \sum_{i,j,k=0}^{3} \lambda_{ijk}\sigma_i \otimes \sigma_j \otimes \sigma_k$, take the $4 \times 3 \times 3$ tensor $\lambda^{red}$ and arrive at

$$\begin{aligned}
\lambda_{0ij} &= \begin{pmatrix} -2 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 1 \end{pmatrix} := A^{(0)}, \quad \lambda_{1ij} = \begin{pmatrix} 0 & 0 & -2 \\ 0 & 0 & 0 \\ -2 & 0 & 0 \end{pmatrix} := A^{(1)}, \\
\lambda_{2ij} &= \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -2 \\ 0 & -2 & 0 \end{pmatrix} := A^{(2)}, \quad \lambda_{3ij} = \begin{pmatrix} -2 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 3 \end{pmatrix} := A^{(3)}.
\end{aligned}$$

Again, it suffices to show that we cannot find four matrices $B^{(r)}, r \in \{1, ..., 4\}$ of rank one such that $A^{(0)}, A^{(1)}, A^{(2)}$ and $A^{(3)}$ can be represented as linear combinations of the $B^{(r)}$. Here, the assumption that we have four $B^{(r)}$ fails due to similar reasons as above: As above, the $B^{(r)}$ have to be linearly independent and is must be possible to write any of the $B^{(r)}$ as a linear combination of the $A^{(i)}$. And a general linear combination of the $A^{(i)}$ is of the form

$$\mathcal{A} = \begin{pmatrix} \alpha & 0 & \beta \\ 0 & \alpha & \gamma \\ \beta & \gamma & \delta \end{pmatrix}, \tag{2.41}$$

and this is of rank one if and only if $\alpha = \beta = \gamma = 0$. Thus, we arrive at a contradiction. □

**Proposition 2.5.** The witness $\mathcal{W}_3^{(\mathrm{oc})}$ can not be measured with four measurement settings, *i.e.* the decomposition (2.19) is optimal.

*Proof.* The proof is more ore less the same as for the witness $\mathcal{W}_3^{(\mathrm{w})}$. One computes the four $A^{(i)}$, they span again a four-dimensional space. A linear combination of them is again of the form (2.41). So one arrives at the same contradiction. □

Note that the proofs given here are quite similar to a way of determining the tensor rank given in [80]. Note also, that the scheme presented here, can due to Eq. (2.37) only prove the optimality for less that seven LvNMs.

Now we turn to the discussion of the case of four or more qubits. The main difficulty for this case lies in the fact that for the corresponding tensors the rank over $\mathbb{C}$ and $\mathbb{R}$ seem to be different. To give an example, let us look at the four-qubit GHZ witness $\mathcal{W}_4^{(ghz)} = 1/2 \cdot \mathbb{1} - |GHZ_4\rangle\langle GHZ_4|$. This witness can be decomposed into five LvNMs and this decomposition is probably optimal. Thus we have to prove that for the corresponding reduced tensor $\lambda^{red}$ the relation $\mathfrak{R}_{\mathbb{R}}(\lambda^{red}) = 5$ holds. But over the complex numbers we have $\mathfrak{R}_{\mathbb{C}}(\lambda^{red}) \leq 4$. This can be seen as follows: We have $|GHZ_4\rangle\langle GHZ_4| = (|0\rangle\langle 0|^{\otimes 4} + |1\rangle\langle 0|^{\otimes 4} + |0\rangle\langle 1|^{\otimes 4} + |0\rangle\langle 0|^{\otimes 4})/2$. This is a decomposition into four "local measurements with non-Hermitean operators." Writing the operators $|0\rangle\langle 1|$ and $|1\rangle\langle 0|$ in the basis of the Pauli matrices (with complex coefficients) yields a decomposition of $\lambda^{red}$ as a sum of four complex tensors of rank one, thus $\mathfrak{R}_{\mathbb{C}}(\lambda^{red}) \leq 4$. From this it also follows that the strategy of our proofs for the three-qubit case cannot be applied to this witness: In our proofs, we did not take into account that we are determining the tensor rank over $\mathbb{R}$, thus we can never prove $\mathfrak{R}_{\mathbb{R}}(\lambda^{red}) = 5$ with this method.

There are, however, some general conclusions which can be drawn from the general results on the tensor rank:

**Corollary 2.9.** (a) To measure a general observable of an $n$-qubit system one needs at least $2 \cdot 3^{n-1}/(n+1)$ LvNMs.
(b) To decompose a general observable into a sum of projectors onto product vectors, one needs at least $4^n/(3n+1)$ product vectors.

*Proof.* Part (a) follows directly from Theorem 2.7 and equation (2.38), since $r(4,3,...,3) = 2 \cdot 3^{n-1}/(n+1)$. Part (b) follows from a reformulation of Theorem 2.7: When trying to decompose into product vectors, the whole $4 \times 4 \times ... \times 4$ tensor $\lambda$ has to be taken into account not only $\lambda^{red}$. Then again Eq. (2.38) can be applied. □

Note that for state tomography $3^n$ measurement settings are required. This is only to a factor of $3(n+1)/2$ more than the method of decomposing an general observable into LvNMs. Thus, the statement (a) ensures us that there are observables, which are difficult to measure by local measurements. This means that to measure them nearly as much LvNMs as for state tomography are required. In this sense they exhibit a strong non-locality. Identifying these observables and clarifying their physical properties would be a nice field for further study.

## 2.6 Calculation of overlaps

In this more technical section we first want to recall some facts about the singular value decomposition of matrices and the Schmidt decomposition. Then we show how one can use these insights to calculate easily the maximal overlap between an $n$-partite pure quantum state and the biseparable states. Our results allow also to calculate the maximal overlap between a pure state and pure states of lower Schmidt rank. Let us start by recalling some facts about the singular value decomposition.

**Theorem 2.10** (Singular value decomposition). Let $A$ be a complex $N \times M$ matrix of rank $k$. Then there exist an $N \times M$ matrix $D$, a unitary $N \times N$ matrix $U$ and a unitary $M \times M$ matrix $V$ such that

$$A = UDV \tag{2.42}$$

holds, where $D$ has the following properties: (a) The off-diagonal terms of $D$ vanish, *i.e.*, $D_{ij} = 0$ for $i \neq j$. (b) The diagonal terms obey $D_{11} \geq D_{22} \geq ... \geq D_{kk} > D_{k+1,k+1} = D_{k+2,k+2} = ... = 0$. (c) The $D_{ii}$ are for $1 \leq i \leq k$ the positive square roots of the eigenvalues of $AA^\dagger$.
The numbers $D_{ii}$ are thus uniquely determined and are called the singular values of $A$. We often write them decreasingly ordered as $\sigma_i(A) = D_{ii}, i = 1,...,k$. If $A$ is real, then $U$ and $V$ can also be chosen to be real.

*Proof.* This statement is included in many textbooks, for a discussion of it and its implications see [81] and especially [82, Chapter 3]. □

The given formulation of the singular value decomposition may sound quite technical, however, it has a clear interpretation: The singular value decomposition is a generalization of the spectral decomposition to non-Hermitean matrices. For instance, if $A$ is a quadratic matrix, $D$ is a diagonal matrices, and if $A$ is Hermitean and has positive eigenvalues, then we can choose $U$ and $V$ such that $U = V^\dagger$. From the singular value decomposition one can directly derive the Schmidt decomposition:

**Theorem 2.11.** Let $|\psi\rangle = \sum_{i,j=1}^{N,M} c_{ij}|a_i b_j\rangle \in \mathcal{H}_A \otimes \mathcal{H}_B$ be a pure state. Then there exist an orthonormal basis $|\alpha_i\rangle$ of $\mathcal{H}_A$ and an orthonormal basis $|\beta_j\rangle$ of $\mathcal{H}_B$ such that

$$|\psi\rangle = \sum_{k=1}^{R} \lambda_k |\alpha_k \beta_k\rangle \tag{2.43}$$

holds, with positive real coefficients $\lambda_k$. The $\lambda_k$ are uniquely determined by the coefficient matrix $C = (c_{ij})$. They are the square roots of the eigenvalues of the matrix $CC^\dagger$, *i.e.*, $\lambda_k = \sigma_k(C)$.

*Proof.* This follows directly from the singular value decomposition of the matrix $C$ :

$$\begin{aligned} |\psi\rangle &= \sum_{ij} c_{ij}|a_i b_j\rangle = \sum_{ijkl} U_{ik} D_{kl} V_{lj} |a_i b_j\rangle \\ &= \sum_{k} D_{kk} (\sum_{i} U_{ki}^{*}|a_i\rangle)(\sum_{j} V_{kj}|b_j\rangle) \end{aligned} \tag{2.44}$$

which proves the claim, if we set $\lambda_k = D_{kk}$. The vectors $|\alpha_k\rangle = (\sum_i U_{ki}^{*}|a_i\rangle)$ are orthogonal, since $U$ is unitary and the $|a_i\rangle$ are orthogonal. Also the vectors $|\beta_k\rangle = (\sum_j V_{kj}|b_j\rangle)$ are orthogonal. □

Now we can formulate the main result of this section. We can derive a formula for the calculation of the maximal overlap between a given pure state and the states of a fixed lower Schmidt rank. Especially, if we look at states of Schmidt rank one, we can calculate the maximal overlap between a given state and the product states. Let us first formulate our main result:

**Theorem 2.12.** Let $|\psi\rangle$ be a pure bipartite state with Schmidt rank $n$ and decreasingly ordered Schmidt coefficients $\sigma_i$. Then the maximal overlap between $|\psi\rangle$ and the states with Schmidt rank $k \leq n$ is given by

$$\sup_{|\phi\rangle \in S_k} |\langle\phi|\psi\rangle|^2 = \sum_{i=1}^{k} \sigma_i^2. \tag{2.45}$$

where we denoted by $S_k$ the set of pure states with Schmidt rank $k$.

To prove this result, we need two lemmata:

**Lemma 2.13.** Let $W$ be an $n$-dimensional complex vector space, $A$ be an $n \times n$-matrix. Fix a $k \leq n$ and let $U, V$ be $k$-dimensional subspaces of $W$. Denote by $|\phi_i\rangle$ (respectively, $|\psi_i\rangle$) any orthonormal basis of $U$ (respectively, $V$). Then

$$\max_{U,|\phi_i\rangle} \max_{V,|\psi_i\rangle} |\sum_{i=1}^{k} \langle\phi_i|A|\psi_i\rangle| = \sum_{i=1}^{k} \sigma_i(A) \tag{2.46}$$

holds. The maxima in (2.46) are taken over all possible subspaces and over all possible orthonormal bases of these subspaces.

*Proof.* For a proof of this fact see [82, Theorem 3.4.1]. Note that this is a generalization of the Rayleigh-Ritz theorem from the spectral decomposition to the singular value decomposition. Note also, that the estimation "$\geq$" in (2.46) is clear, since we can chose the $|\phi_i\rangle$ and $|\psi_i\rangle$ to be the vectors corresponding to the biggest singular values. □

**Lemma 2.14.** Let $s_i, i = 1, ..., k$ be some positive, decreasingly ordered coefficients. Then in the situation of Lemma 2.13

$$\max_{U,|\phi_i\rangle} \max_{V,|\psi_i\rangle} |\sum_{i=1}^{k} s_i \langle\phi_i|A|\psi_i\rangle| = \sum_{i=1}^{k} s_i \sigma_i(A) \tag{2.47}$$

holds.

*Proof.* The estimation "$\geq$" is clear, as in Lemma 2.13. In order to show the "$\leq$"-part we can choose $t_i \geq 0, i = 1, ..., k$ such that $s_i = \sum_{l=i}^{k} t_l$, *i.e.*, we have $t_k = s_k, t_{k-1} = s_{k-1} - s_k$, etc. Then we have, with the help of Lemma 2.13:

$$\begin{aligned} |\sum_{i=1}^{k} s_i \langle\phi_i|A|\psi_i\rangle| &= |\sum_{i=1}^{k}\sum_{l=i}^{k} t_l \langle\phi_i|A|\psi_i\rangle| = |\sum_{l=1}^{k} t_l \sum_{i=1}^{l} \langle\phi_i|A|\psi_i\rangle| \\ &\leq \sum_{l=1}^{k} t_l |\sum_{i=1}^{l} \langle\phi_i|A|\psi_i\rangle| \leq \sum_{l=1}^{k} t_l \sum_{i=1}^{l} \sigma_i(A) \\ &= \sum_{i=1}^{k}\sum_{l=i}^{k} t_l \sigma_i(A) = \sum_{i=1}^{k} s_i \sigma_i(A). \end{aligned} \tag{2.48}$$

This proves the Lemma 2.14. □

After these preliminaries we can prove the theorem:

*Proof of the Theorem 2.12.* We can express $|\psi\rangle = \sum_{ij} A_{ij} |ij\rangle$ and similarly $|\phi\rangle = \sum_{ij} C_{ij} |ij\rangle$. Then we have $\sigma_i = \sigma_i(A)$. We can further perform a Schmidt decomposition of $|\phi\rangle$ as a singular value decomposition of $C$ which can be written as

$$C_{ij} = \sum_{l=1}^{k} \tau_l \mathfrak{a}_i^l \mathfrak{b}_j^l, \tag{2.49}$$

where $\tau_l = \sigma_l(C)$ are the singular values, $\mathfrak{a}^l$ is the $l$-th column of $U$ in Eq. 2.42, and $\mathfrak{b}^l$ is the $l$-th row of $V$ in Eq. 2.42. The vectors $\mathfrak{a}^l$ and $\mathfrak{b}^l$ are pairwise orthogonal: $\langle\mathfrak{a}^l|\mathfrak{a}^m\rangle = \delta_{lm}$ and $\langle\mathfrak{b}^l|\mathfrak{b}^m\rangle = \delta_{lm}$. So it suffices to maximize

$$\begin{aligned} \sup_{|\phi\rangle \in S_k} |\langle\phi|\psi\rangle| &= \sup_{C} |\sum_{ij} C_{ij}^* A_{ij}| \\ &= \sup_{\tau_l} \sup_{\mathfrak{a}^l} \sup_{\mathfrak{b}^l} |\sum_{l=1}^{k} \tau_l \langle\mathfrak{a}^l|A|\mathfrak{b}^{l*}\rangle| \\ &= \sup_{\tau_i} \sum_{i=1}^{k} \tau_i \sigma_i(A), \end{aligned} \tag{2.50}$$

where we have used Lemma 2.14. So all in all we are left with a linear optimization problem with the constraint $\sum_i \tau_i^2 = 1$, which has to hold since $|\phi\rangle$ should be normalized. This can be solved with Lagrange multipliers, and the result is (2.45). □

**Corollary 2.15.** Let $|\psi\rangle$ be a pure state. Then the maximal squared overlap between $|\psi\rangle$ and the product states is given by the biggest squared Schmidt coefficient $\sigma_1^2$,

$$\sup_{|a\rangle,|b\rangle} |\langle a,b|\psi\rangle|^2 = \sigma_1^2. \tag{2.51}$$

This is just a simple consequence of the fact that the pure product states are just the states of Schmidt rank one. This is also the fact that has been used to calculate the constants for the witnesses. It will again be used in the next chapter.

Finally, it is important to note that due to Theorem 2.12 many of the criteria in this thesis can be straightforwardly extended to the detection of states of a given high Schmidt rank. By this we mean that one may define other convex sets besides the set of separable states by convex combinations: A mixed state is of Schmidt number $k$ iff it can be written as a convex combination of pure states of Schmidt rank $k$ or smaller. This is a natural generalization of the concept of the separable states, which are in this scheme the mixed states of Schmidt number one [83, 84]. One can define witnesses in a similar way, as observables which can detect that the Schmidt number of some state exceeds some number [85]. Thus, the bound on the overlap in Eq. (2.45) enables us to derive criteria for the detection of high Schmidt number in the same manner as for the detection of entanglement.

## 2.7 Stabilizer witnesses

In the preceding sections we dealt with witnesses for a given state $|\psi\rangle$ of the form

$$\mathcal{W} = C \cdot \mathbb{1} - |\psi\rangle\langle\psi|, \tag{2.52}$$

where $C$ is a constant chosen in such a way that the witness is positive on some set of states. Here, we want to introduce a new way of of construction witnesses, namely of the form

$$\widetilde{\mathcal{W}} = C' \cdot \mathbb{1} - \sum_i S_i \tag{2.53}$$

where $C'$ is again a constant and the $S_i$ are the so-called *stabilizing operators* of the state, *i.e.*, they obey

$$S_i|\psi\rangle = |\psi\rangle. \tag{2.54}$$

Note that $|\psi\rangle\langle\psi|$ is also a stabilizing operator of the state $|\psi\rangle$. But the main idea is to take *stabilizers which are local observables.* The usefulness of the stabilizer formalism for the detection of entanglement has been first demonstrated in [86]. In this way we can construct witnesses which are by definition measurable with local measurements. As we will see, for many states this idea turns out to be useful. We will discuss three classes of states, for which witnesses based on the stabilizer formalism are easy to construct.

The first class we want to discuss are the so-called cluster states [75, 87]. They are good candidates for our approach since their definition relies on stabilizer equations of the type (2.54): The $n$-qubit cluster state, $|CL_n\rangle$, is uniquely defined by the following equations[3]:

$$\begin{aligned} \sigma_x^{(1)}\sigma_z^{(2)}|CL_n\rangle &= |CL_n\rangle, \\ \sigma_z^{(k-1)}\sigma_x^{(k)}\sigma_z^{(k+1)}|CL_n\rangle &= |CL_n\rangle; \text{ for } k=2,3,...,n-1, \\ \sigma_z^{(n-1)}\sigma_x^{(n)}|CL_n\rangle &= |CL_n\rangle. \end{aligned} \tag{2.55}$$

In the more general definition, also eigenvalues $-1$ are allowed in Eq. (2.55). For simplicity, however, we will concentrate on cluster states with the eigenvalue $+1$. In the following we denote the stabilizing operators from Eq. (2.55) as $S_i^{CL_n}, i=1,...,n$.

Let us not some properties of these states: First, we note that one can prove that the state $|CL_n\rangle$ is uniquely determined by the equations (2.55). Nevertheless, the class of all cluster states is big enough to contain many interesting states: For two qubits the cluster state is a Bell state, and for three qubits the cluster state is up to a local change of the basis a GHZ state. There is also an explicit formula for the cluster state:

$$|CL_n\rangle = \frac{1}{\sqrt{2^n}}\bigotimes_{i=1}^{n}(|0\rangle_i\sigma_z^{(i+1)}+|1\rangle_i). \tag{2.56}$$

From this one can calculate $|CL_n\rangle$ by multiplying out the right hand side and applying $\sigma_z$ to the consecutive qubit. Two further properties of the cluster states are important: An $n$-qubit cluster state is alway truly $n$-partite entangled. Furthermore, this entanglement is robust: $n/2$ measurements on different qubits are required to disentangle the state [75]. Physically, a cluster state can be created by applying the Ising-type time evolution $U = \exp\left(i\pi/4\sum_k \sigma_z^{(k)}\sigma_z^{(k+1)}\right)$ to the fully separable state $|x^+, x^+, ..., x^+\rangle$. Now we can formulate our main result concerning cluster states:

**Theorem 2.16.** Let $S_i^{CL_n}$ be the stabilizing operators of an $n$ qubit cluster state, defined in (2.55). Then

$$\widetilde{\mathcal{W}}_n^{(\mathrm{cl})} = (n-1)\mathbb{1} - \sum_{i=1}^{n} S_i^{CL_n} \tag{2.57}$$

is an entanglement witness detecting true $n$-qubit entanglement. $\widetilde{\mathcal{W}}_n^{(\mathrm{cl})}$ detects the state $|CL_n\rangle$ and, more general, it detects states of the form $\varrho(p) = p|CL_n\rangle\langle CL_n| + (1-p)\mathbb{1}/2^n$ for $p > (n-1)/n$.

Before proving this theorem let us discuss its implications. First, note that the witness (2.57) cannot be improved by choosing a smaller constant than $(n-1)$. This is due to the fact that for the biseparable state $|\phi\rangle = |CL_{n-1}\rangle|0\rangle$ fulfills already $Tr(\mathcal{W}_n^{(\mathrm{cl})}|\phi\rangle\langle\phi|) = 0$.

The remarkable fact is that a measurement of the witness $\mathcal{W}_n^{(\mathrm{cl})}$ requires only two measurement settings: From the coincidence probabilities of $\sigma_x^{(1)}\sigma_z^{(2)}\sigma_x^{(3)}...$ and

[3] $\sigma_x^{(i)}$ denotes a measurement of $\sigma_x$ on the $i$-th qubit. To keep the notation short, throughout this section we omit the tensor product symbol "⊗".

$\sigma_z^{(1)}\sigma_x^{(2)}\sigma_z^{(3)}$... on can determine the expectation values of all stabilizing operators, and thus the expectation value of the witness. This shows that in the case of cluster states only two measurements are needed to detect the entanglement. On the other hand, one has to admit that the witness $\mathcal{W}_n^{(\mathrm{cl})}$ is not very robust against noise. The value $p > (n-1)/n$ needed for a detection is quite high. To prove Theorem 2.16 we first calculate another witness:

**Proposition 2.17.** Let $|CL_n\rangle$ be the $n$-qubit cluster state. Then

$$\mathcal{W}_n^{(\mathrm{cl})} = \frac{1}{2}\mathbb{1} - |CL_n\rangle\langle CL_n| \tag{2.58}$$

is an entanglement witness detecting true $n$-qubit entanglement. $\mathcal{W}_n^{(\mathrm{cl})}$ detects the state $|CL_n\rangle$ and, more general, it detects states of the form $\varrho(p) = p|CL_n\rangle\langle CL_n| + (1-p)\mathbb{1}/2^n$ for $p > (2^{n-1}-1)/(2^n-1)$.

This witness is not so sensitive to noise, but here one has to decompose $|CL_n\rangle\langle CL_n|$ into local measurements. This leads to a decomposition with many settings, for $n = 3$ one needs four settings[4], for $n = 4$ it seems that one needs nine settings[5]. Now we can go to the proofs:

*Proof of Proposition 2.17.* In order to prove Proposition 2.17 we have to show that the maximal Schmidt coefficient which appears if we expand $|CL_n\rangle$ in the Schmidt decomposition for an arbitrary, but fixed bipartite splitting is not bigger than $1/\sqrt{2}$. In some cases this can be verified directly via Eq. (2.56), but for the general case this is not straightforward. But we can prove it the following way: It has been shown that from a given cluster state one can prepare a singlet state for an arbitrary pair of qubits by local measurements only [75]. This implies that if we fix a bipartite splitting and choose two qubits on each party, we can generate via LOCC a singlet state for this two qubits from the cluster state. Thus we can generate a state between the splittings for which the Schmidt coefficients are not bigger than $1/\sqrt{2}$. At this point we use the so-called Nielsen theorem on the behavior of Schmidt coefficients under LOCC: A transformation $|\psi\rangle \rightarrow |\phi\rangle$ is possible if and only if the vector of Schmidt coefficients of $|\phi\rangle$ majorizes the vector of Schmidt coefficients of $|\psi\rangle$ [88]. But this implies here, that the biggest squared Schmidt coefficient of the original state $|CL_n\rangle$ cannot exceed $1/\sqrt{2}$. □

*Proof of Theorem 2.16.* In order to prove that $\widetilde{\mathcal{W}}_n^{(\mathrm{cl})}$ is an entanglement witness, we prove that detection via $\widetilde{\mathcal{W}}_n^{(\mathrm{cl})}$ implies detection via $\mathcal{W}_n^{(\mathrm{cl})}$, *i.e.*, $Tr(\varrho\widetilde{\mathcal{W}}_n^{(\mathrm{cl})}) < 0 \Rightarrow Tr(\varrho\mathcal{W}_n^{(\mathrm{cl})}) < 0$. To show this let us look at the operator

$$X = \widetilde{\mathcal{W}}_n^{(\mathrm{cl})} - 2\mathcal{W}_n^{(\mathrm{cl})} = (n-2)\mathbb{1} - \sum_{i=1}^{n} S_i^{CL_n} + 2|CL_n\rangle\langle CL_n|. \tag{2.59}$$

We have to show $X \geq 0$ to complete the proof. To do this, let us look at the total set of cluster states with eigenvalues $\pm 1$ in Eq. (2.55). There are $2^n$ different of

[4] We have computed this decomposition and proved its optimality, since the three-qubit cluster state is up to local unitaries the GHZ state.

[5] This witness was presented in Eq. (2.27). Note that the state $|\psi_4^{cl}\rangle$ given in Eq. (2.26) is the cluster state $|CL_n\rangle$ up to local unitaries.

them, let us denote them by $|CL_n^i\rangle, i = 1, ..., 2^n$ and $|CL_n^1\rangle = |CL_n\rangle$. The vectors $|CL_n^i\rangle$ are pairwise orthogonal since they are eigenvectors of the same observables with different eigenvalues. Thus, they form an orthonormal basis of the total Hilbert space. Due to the eigenvalue properties of the $|CL_n^i\rangle$ the relations $\langle CL_n^i|X|CL_n^i\rangle \geq 0$ and $\langle CL_n^i|X|CL_n^j\rangle = 0$ for $i \neq j$ hold. In other words, $X$ is in the $|CL_n^i\rangle$ basis diagonal with positive entries on the diagonal. This implies $X \geq 0$. □

Sometimes the noise might be too large, so that the state under consideration does not contain any true multipartite entanglement anymore. However, one might be interested in detecting some entanglement, even if it is not multipartite entanglement. This can also be done, using the same measurements.

**Theorem 2.18.** The following entanglement witness detects entanglement which might be only biseparable entanglement for states close to an $n$-qubit cluster state

$$\mathcal{Q}_{CL_n} := (n-1) \cdot \mathbb{1} - S_1^{CL_n} - 2\sum_{k=2}^{n-1} S_k^{CL_n} - S_n^{CL_n}. \tag{2.60}$$

*Proof.* We need the Cauchy-Schwarz inequality $\vec{a} \cdot \vec{b} \leq |\vec{a}||\vec{b}|$ and the fact that $\langle\sigma_x^{(k)}\rangle^2 + \langle\sigma_z^{(k)}\rangle^2 \leq 1$. This bound follows directly from the geometry of the Bloch sphere, we will give another proof in a more general setting in the next chapter. Then we can estimate for product states $q_k := \langle S_k^{CL_n} + S_{k+1}^{CL_n}\rangle = \langle\sigma_z^{(k-1)}\rangle\langle\sigma_x^{(k)}\rangle\langle\sigma_z^{(k+1)}\rangle + \langle\sigma_z^{(k)}\rangle\langle\sigma_x^{(k+1)}\rangle\langle\sigma_z^{(k+2)}\rangle \leq |\langle\sigma_x^{(k)}\rangle\langle\sigma_z^{(k+1)}\rangle + \langle\sigma_z^{(k)}\rangle\langle\sigma_x^{(k+1)}\rangle| \leq 1$. To keep the notation short, we denoted here for the end of the chain $\sigma_z^{(0)} = \sigma_z^{(n+1)} = \mathbb{1}$. Thus we have $\langle\mathcal{Q}_{CL_n}\rangle = (n-1) - \sum_{k=1}^{n-1} q_k \geq 0$. Due to the convexity this holds also all fully separable states. The constant term in Eq. (2.60) is the smallest possible, since the separable state $|x^+\rangle|z^+\rangle|x^+\rangle|z^+\rangle...$ gives $\langle\mathcal{Q}_{CL_n}\rangle = 0$. □

Another class where the stabilizer formalism turns out to be useful is the class of GHZ states. It is easy to see that $\mathcal{W}_n^{(ghz)} := \mathbb{1}/2 - |GHZ_n\rangle\langle GHZ_n|$ detects genuine $n$-qubit entanglement in the vicinity of the $n$-qubit state $|GHZ_n\rangle = (|0...0\rangle + |1...1\rangle)/\sqrt{2}$.

Again, for the state $|GHZ_n\rangle$ a witness for the detection of genuine $n$-qubit entanglement can be constructed based on the stabilizer formalism. It reads:

$$\widetilde{\mathcal{W}}_n^{(\mathrm{ghz})} := (n-1) \cdot \mathbb{1} - \prod_{k=1}^{n} \sigma_x^{(k)} - \sum_{k=1}^{n-1} \sigma_z^{(k)}\sigma_z^{(k+1)}. \tag{2.61}$$

The proof is essentially the same as the one for the cluster state: one looks at $X := \widetilde{\mathcal{W}}_n^{(\mathrm{ghz})} - 2\mathcal{W}_n^{(\mathrm{ghz})}$ and shows that $X \geq 0$. This can be easily done in a basis of GHZ states.

The observables appearing in the right hand side of Eq. (2.61) are again stabilizing operators of the GHZ state. Also, only two measurement settings are necessary. For the detection of biseparable entanglement a similar witness can be constructed.

**Theorem 2.19.** The following entanglement witness detects biseparable entanglement for states close to the $n$-qubit GHZ state

$$\mathcal{Q}_{GHZ_n} := (n-1)\Big(\mathbb{1} - \prod_{k=1}^{n} \sigma_x^{(k)}\Big) - \sum_{k=1}^{n-1} \sigma_z^{(k)} \sigma_z^{(k+1)}. \tag{2.62}$$

*Proof.* For product states we can estimate as before with the Cauchy-Schwarz inequality $\langle \prod_k \sigma_x^{(k)} \rangle + \langle \sigma_z^{(m)} \sigma_z^{(m+1)} \rangle \leq |\langle \sigma_x^{(m)} \rangle| \cdot |\langle \sigma_x^{(m+1)} \rangle| + |\langle \sigma_z^{(m)} \rangle| \cdot |\langle \sigma_z^{(m+1)} \rangle| \leq 1$. The constant in Eq. (2.62) is the smallest possible since the separable state $|z^- z^- z^- ...\rangle$ gives $\langle \mathcal{Q}_{GHZ_N} \rangle = 0$. □

Finally, we want to show that the previous results can straightforwardly be generalized for graph states [89, 90]. These states are defined in the following way: One takes a graph, *i.e.*, a set of $n$ vertices, where some of the vertices are connected by edges. These connections can be described by an adjacency matrix $\Gamma$ of the graph: $\Gamma_{kl} = 1$ (0) if the vertices $k$ and $l$ are connected (not connected). From this one defines an $n$-qubit state as an eigenstate with eigenvalue 1 of the stabilizing operators

$$S_k^{G_n} := \sigma_x^{(k)} \prod_{l \neq k} (\sigma_z^{(l)})^{\Gamma_{kl}}. \tag{2.63}$$

Physically, $\Gamma_{kl} = 1$ (0) means that the spins $k$ and $l$ interact (or not) by an Ising-type interaction. We can assume, that the graph does not decay into two separate subgraphs, otherwise the graph state would be biseparable. The following entanglement witness detects biseparable entanglement for states close to an $n$-qubit graph state

$$\mathcal{Q}_{G_n} := \frac{1}{2}\Big(\sum_k n_k\Big) \cdot \mathbb{1} - \sum_k n_k S_k^{G_n}, \tag{2.64}$$

where $n_k = \sum_l \Gamma_{k,l}$ is the number of neighbors of spin $k$. A witness detecting genuine $n$-party entanglement can be defined as

$$\widetilde{\mathcal{W}}_n^{(\mathrm{g})} := (n-1) \cdot \mathbb{1} - \sum_k S_k^{G_n}. \tag{2.65}$$

The proofs are essentially the same as before: For $\mathcal{Q}_{G_N}$ one estimates each connection using the Cauchy-Schwarz inequality and for $\widetilde{\mathcal{W}}_n^{(\mathrm{g})}$ one uses the fact that one can produce from a graph state by local means a singlet between an arbitrary pair of qubits [90].

## 2.8 Experiments

Now we turn to the description of recent experiments, performed by Mohamed Bourennane, Manfred Eibl, Christian Kurtsiefer, Sascha Gaertner, and Harald Weinfurter in Munich and Garching. These were the first experiments where witnesses

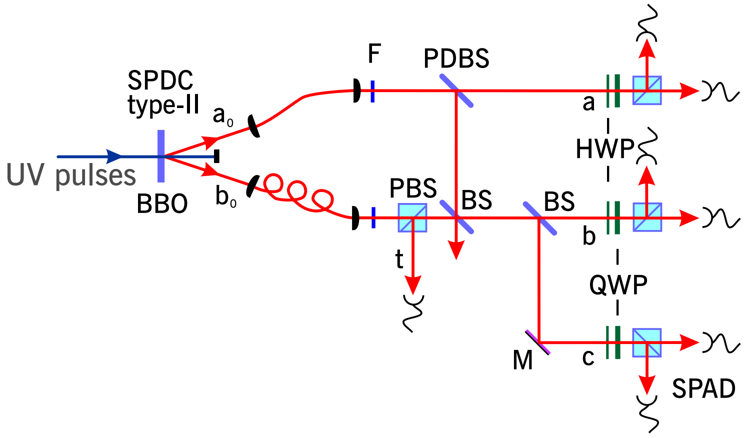

Figure 2.1. *Setup for the generation of the three-photon W state. See text for details.*

were used for the detection of genuine multipartite entanglement[6]. To be precise, the witnesses for the three-qubit W state and the four-qubit $\Psi^{(4)}$ state were measured. These experiments used polarized photons as qubits. So let us start our section by discussing the generation of entangled states.

One possibility to generate entangled states of polarized photons in a laboratory uses the so called type-II spontaneous parametric down conversion (SPDC) [92]. In this process, an ultraviolet (UV) laser pumps a nonlinear beta-Barium-borate (BBO) crystal. Multiple emission events during one pump pulse lead to the emission of the state

$$|\psi\rangle \sim \exp(-i\alpha(a_{0,h}^{+}b_{0,v}^{+} + a_{0,v}^{+}b_{0,h}^{+}))|0\rangle \tag{2.66}$$

distributed onto two modes, $a_0$ and $b_0$. Here, $a_{0,h}^{+}$ denotes a creation operator, creating a horizontal polarized photon in mode $a_0$, and $b_{0,v}^{+}$ creates a vertical polarized photon in mode $b_0$. In the second order of this process only the quadratic terms are relevant, leading to

$$|\psi\rangle \sim (a_{0,h}^{+}b_{0,v}^{+} + a_{0,v}^{+}b_{0,h}^{+})^2|0\rangle. \tag{2.67}$$

This results in the state

$$|\psi\rangle \sim |2h\rangle_{a_0}|2v\rangle_{b_0} + |2v\rangle_{a_0}|2h\rangle_{b_0} + |1h,1v\rangle_{a_0}|1h,1v\rangle_{b_0}, \tag{2.68}$$

where $|2h\rangle_{a_0}$ denotes a state of two horizontal polarized photons in mode $a_0$, and $|1h,1v\rangle_{b_0}$ denotes one horizontal and one vertical polarized photon in mode $b_0$, etc.[7]

The production of a W state from the four-photon state (2.68) is based on the idea to take one photon as a trigger [93, 94] (see Figure 2.1)[8]. Let us assume that in all the modes $a, b, c$ as well as in the trigger mode $t$ a photon is are detected. Then

[6] The experimental detection of two-qubit entanglement using witnesses has been first reported in Ref. [91]

[7] Note that in first order the two-qubit Bell state $|h\rangle_{a_0}|v\rangle_{b_0} + |v\rangle_{a_0}|h\rangle_{b_0}$ is produced.

[8] We thank M. Bourennane for the abandonment of the Figures 2.1 – 2.4.

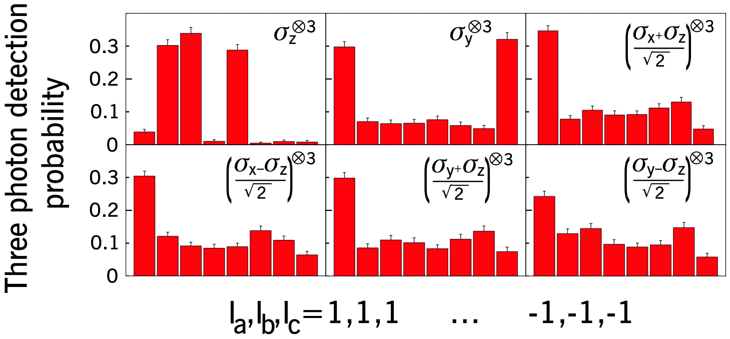

Figure 2.2. *Threefold coincidence probabilities for the six measurement settings, needed for the detection of the three-qubit W state. See text for further details.*

the photon in the mode $t$ must be vertically polarized, due to the polarizing beam splitter PBS. Thus two of the three photons in the modes $a, b, c$ must be horizontally polarized, and one vertically polarized. To distribute these photons with the correct probability over the three modes, a polarization dependent beam splitter (PDBS) in arm $a$ (with transmissions $t_H = 2t_V$) is used. The remaining two photons are distributed into the two arms by two other 50-50 beamsplitters (BS).

For the detection of the tripartite entanglement of the W state two different witnesses were used. The first one is

$$\mathcal{W}_3^{(\mathrm{w},1)} = \frac{2}{3}\mathbb{1} - |W\rangle\langle W| \tag{2.69}$$

which is known from Eqs. (2.16) and (2.17). The second one is

$$\begin{aligned}\mathcal{W}_3^{(\mathrm{w},2)} &= \frac{1}{2}\mathbb{1} - |\overline{GHZ}\rangle\langle\overline{GHZ}| \\ &= \frac{1}{8}\Big(3\cdot\mathbb{1}\otimes\mathbb{1}\otimes\mathbb{1} - \mathbb{1}\otimes\sigma_y\otimes\sigma_y - \sigma_y\otimes\mathbb{1}\otimes\sigma_y - \sigma_y\otimes\sigma_y\otimes\mathbb{1} - \\ &\quad +2\cdot\sigma_z^{\otimes 3} - \frac{1}{2}\cdot(\sigma_z+\sigma_x)^{\otimes 3} - \frac{1}{2}\cdot(\sigma_z-\sigma_x)^{\otimes 3}\Big).\end{aligned} \tag{2.70}$$

where $|\overline{GHZ}\rangle = (|y^+y^+y^+\rangle - |y^-y^-y^-\rangle)/\sqrt{2}$. This witness is also known from Section 2.4.2, here it is only written in a different basis. Note that $\mathcal{W}_3^{(w,2)}$ can also serve for a detection of GHZ-type entanglement, when $Tr(\mathcal{W}_3^{(w,2)}\varrho) < -1/4$ the state $\varrho$ belongs to the GHZ class. Only six local measurement settings are required, to measure the expectation value of the two witnesses. Theoretically, if the W state were produced perfectly, one would expect

$$\begin{aligned} Tr(\mathcal{W}_3^{(\mathrm{w},1)}|W\rangle\langle W|) &= -\frac{1}{3}, \\ Tr(\mathcal{W}_3^{(\mathrm{w},2)}|W\rangle\langle W|) &= -\frac{1}{4}. \end{aligned} \tag{2.71}$$

In the experiments, each measurement setting corresponds to a special setting of half- and quarter- wave plates (HWP and QWP) and the beam splitter in the detector of

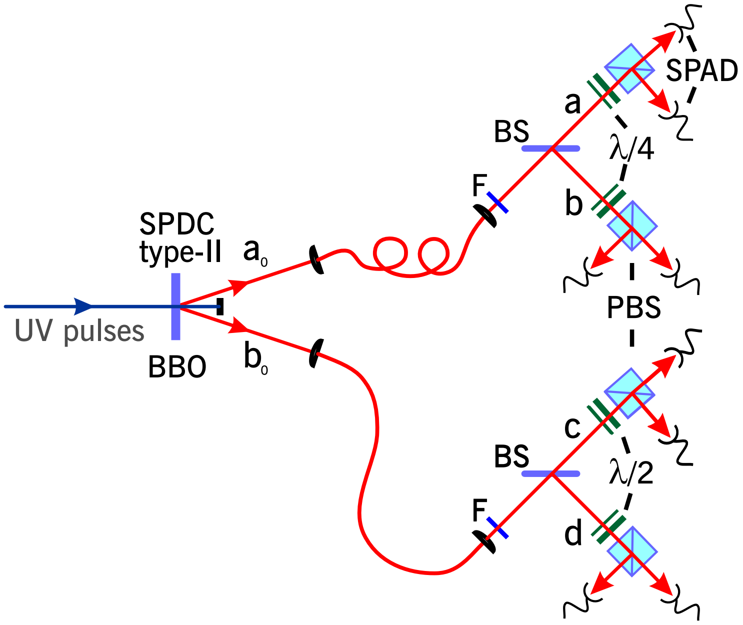

Figure 2.3. *Setup for the generation of the four-photon $\Psi^{(4)}$ state. See text for details.*

each of the three modes. This results in coincidence probabilities for a given setting for the eight possible outcomes. These coincidence probabilities for the six required measurements are shown in Figure 2.2. From these probabilities, the mean values of the six settings and thus the expectation value of the witnesses can be calculated. The result is:

$$\begin{aligned} Tr(\mathcal{W}_3^{(\mathrm{w},1)}\rho_W)_{exp} &= -0.197 \pm 0.018, \\ Tr(\mathcal{W}_3^{(\mathrm{w},2)}\rho_W)_{exp} &= -0.139 \pm 0.030. \end{aligned} \tag{2.72}$$

which proves that the state produced in the experiment was truly tripartite entangled.

It is remarkable that the mean value of the witness $\mathcal{W}_3^{(\mathrm{w},2)}$ is not smaller than $-1/4$. As already mentioned, a mean value smaller than that would indicate GHZ entanglement. Thus the result is in agreement with the assumption that the produced state is indeed in the W class of mixed states.

The question whether Bell inequalities could also serve for a detection of tripartite entanglement in this case is not easy to decide. In the first experiments to produce W states, Bell inequalities failed to detect the tripartite entanglement [94]. However, with a better choice of the settings this might be possible [95].

The production of the four-qubit singlet state is conceptually simpler (see Figure 2.3) than the production of the W state [73, 74]. After preparing the state $|\psi\rangle \sim (a_{0,h}^+ b_{0,v}^+ + a_{0,v}^+ b_{0,h}^+)^2|0\rangle$ in the modes $a_0$ and $b_0$ two beam-splitters are used to distribute the photons in the four modes $a, b, c$ and $d$. Using the fact that

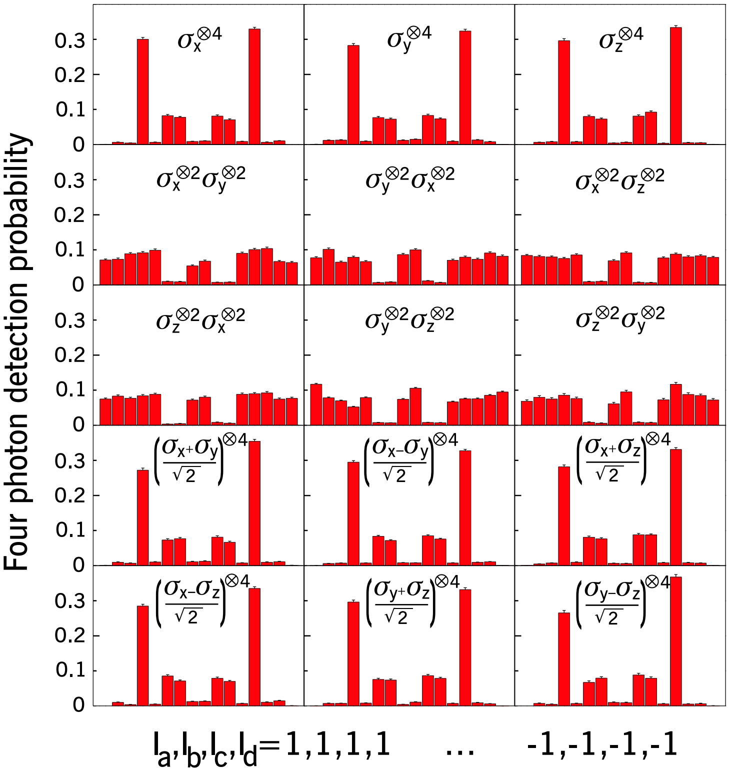

Figure 2.4. *Fourfold coincidence probabilities for the fifteen measurement settings, needed for the detection of the four-qubit $\Psi^{(4)}$ state. See text for further details.*

these beam-splitters transform the creation operators as $a_{0,h}^{+} = (a_h^{+} + b_h^{+})/\sqrt{2}$ and $b_{0,h}^{+} = (c_h^{+} + d_h^{+})/\sqrt{2}$, etc. one can directly calculate that fourfold coincidences in all the four modes $a, b, c$, and $d$ can only occur when the state

$$\begin{aligned}|\psi\rangle \quad \sim \quad & |h\rangle_a|h\rangle_b|v\rangle_c|v\rangle_d + |v\rangle_a|v\rangle_b|h\rangle_c|h\rangle_d + \\ & + \frac{1}{2}\big(|v\rangle_a|h\rangle_b + |h\rangle_a|v\rangle_b\big)\big(|v\rangle_c|h\rangle_d + |h\rangle_c|v\rangle_d\big) \end{aligned} \tag{2.73}$$

is produced. This state is, up to a sign which can be removed by a local change of the basis the four qubit singlet state $|\Psi^{(4)}\rangle$, introduced in Eq. (2.24). The witness for this state, calculated in Section 2.4.3, is given by $\mathcal{W}_4^{\Psi^{(4)}} = 3/4 \cdot \mathbb{1} - |\Psi^{(4)}\rangle\langle\Psi^{(4)}|$. Theoretically one would expect

$$Tr(\mathcal{W}_4^{\Psi^{(4)}}|\Psi^{(4)}\rangle\langle\Psi^{(4)}|) = -\frac{1}{4}. \tag{2.74}$$

The measurement of the witness requires fifteen measurement settings. These measurement data for these are shown in Figure 2.4. From these data one can conclude that the expectation value of the witness is given by

$$Tr(\mathcal{W}_4^{\Psi^{(4)}}\rho_{\Psi^{(4)}})_{exp} = -0.151 \pm 0.01, \tag{2.75}$$

which confirms the genuine four-partite entanglement beyond any doubt.

For the $\Psi^{(4)}$ state the possible detection with a Bell inequality is even more problematic: Up to now there is no Bell inequality known which could prove that the state $\Psi^{(4)}$ is truly four-partite entangled. Thus witnesses are the only possibility nowadays to prove the entanglement in this case.

Finally, note that in the reported experiments the production of the entangled states used post selection: One can conclude that the desired state was produced, when coincidences on all modes occur, but in many cases only two photons or different four photon states are produced. This is a typical drawback of multi-photon experiments. A different approach which allows to produce entangled states directly without post selection uses ions in ion traps. Also in these systems, W states have been produced recently [96].

## 2.9 Conclusion

In conclusion, we have demonstrated in this part of the thesis that witnesses are useful tools for the detection of multipartite entanglement.

First, we have provided a simple method to construct witnesses for the detection of true multipartite entanglement. This method can be applied to all pure multipartite entangled states and is based on the calculation of the maximal overlap. Then, we have introduced the method of local decompositions of witnesses which makes the measurement of witnesses and other nonlocal observables feasible in a laboratory. We calculated local decompositions for several interesting three- and four-qubit witnesses. We developed a method based on the formalism of the tensor rank to show that a given local decomposition is optimal. With this method we were able to show that our decompositions are in many cases optimal. We also developed a different method of construction witnesses based on the stabilizer formalism. This offers the possibility to detect cluster and GHZ states of an arbitrary number of qubits with only two measurements settings. Finally, we reported about experiments, where some of the witnesses presented in this chapter have been implemented to detect three- and four-photon entanglement.

There are several interesting topics which should be addressed further. One challenging task, is of course, to develop other techniques for the construction of witnesses. The two methods presented here, the one based on the overlap and the one based on the stabilizer formalism, are quite simple, however it is unlikely that they are able to detect all multipartite entangled states. Constructing new witnesses with a high robustness against noise would be of great interest.

Another interesting question lies in the investigation of the payoff between the number of measurements needed and the robustness against noise of the witness. In calculations for the stabilizer witnesses, it turned out that they are easy to measure, but fragile against disturbance of the state to be detected. This can be viewed as a property of the states: Highly entangled pure states can be detected with only few measurement settings. If noise is added, the entanglement in the state decreases and more settings are necessary. Thus the question arises, whether the number of measurement settings needed for the detection of a state can be used as an entanglement measure to quantify the entanglement.

The last interesting topic we want to mention is the development of new separability criteria, which are easy to implement. Witnesses are always linear in the state, but maybe the convex set of the separable states can be better approximated with nonlinear criteria. To make these criteria applicable in experiments, they should rely on directly measurable quantities, state reconstruction should be avoided. In the next chapter, we present some steps towards nonlinear entanglement witnesses. There we use variances and entropies to detect entanglement. These criteria also exhibit some relations between uncertainty relations and entanglement.

# Chapter 3

# Uncertainty relations and entanglement

## 3.1 Overview

The uncertainty principle is, besides entanglement, another aspect of quantum theory, which departs from the classical intuition. The fact that for certain pairs of observables the outcomes of a measurement cannot both be fixed with an arbitrary precision has led to many physical and philosophical discussions. However, the uncertainty principle is, like entanglement, still not fully understood.

In this part of the thesis, we want to address the natural question whether there are any relationships between the uncertainty principle and entanglement. This question has been posed before mainly for infinite-dimensional composite systems. In these cases it turned out that one can formulate many separability criteria in terms of uncertainty relations [97–101]. For finite-dimensional systems this was not so clear, in, fact the first separability criteria in terms of uncertainty relations have been established in Refs. [102, 103].

When talking about uncertainty relations, one has to keep in mind that there are different mathematical formulations for the same physical fact. Besides the standard formulation in terms of variances [104–106] there is another formulation in terms of entropies, the so-called entropic uncertainty relations [107, 108]. The main difference between these formulations lies in the fact that entropic uncertainty relations only take the probabilities of the different outcomes of a measurement into account. Variance based uncertainty relations depend also on the measured values (*i.e.*, the eigenvalues of the observable) itself.

As already mentioned, we establish connections between uncertainty relations and entanglement, but for finite-dimensional systems. For this we will use variance based uncertainty relations as well as entropic uncertainty relations. Thus, this chapter is divided into two parts: In the Sections 3.3 - 3.7 we explore the connections between variance based uncertainty relations and entanglement. In the Sections 3.8 - 3.10 we consider the entropic formulation of the uncertainty principle. In detail, we proceed as follows:

In the first section we recall some facts about the uncertainty principle in general. In the following sections we want to derive separability criteria in terms of variances.

After explaining the main idea (Section 3.3) we start with the discussion of the so-called local uncertainty relations (LURs) [102]. We recall the main idea of the LURs and discuss their drawbacks, e.g. the problems to extend them to multipartite systems. These problems can be solved by considering nonlocal observables, as we will show then. We derive a collection of nonlocal uncertainty relations which can serve for the detection of entanglement in various situations.

In Section 3.6 we discuss the relationship between our criteria and entanglement witnesses. As it will turn out, the variance based criteria are not stronger than witnesses, there are, however, sometimes easier to construct.

A very interesting feature of the variance criteria is that they allow to connect finite-dimensional systems with infinite-dimensional systems. We will recall some facts about infinite-dimensional systems in Section 3.7 and explain this connection. This will enable us to translate separability criteria known from infinite-dimensional systems to finite-dimensional systems.

In Section 3.8 we start our discussion of the relationship between entropic uncertainty relations and entanglement by recalling some facts about entropies, entropic uncertainty relations and related topics.

Finally, we present two different ways to obtain separability criteria in terms of entropic uncertainty relations. In Section 3.9 we develop a method based on the calculation of overlaps. In Section 3.10 we derive another method, which shows how any entropic uncertainty relation on one part of the system gives rise to a separability criterion on the composite quantum system.

## 3.2 The uncertainty principle

Let us start by recalling some facts about the uncertainty principle in quantum mechanics. The uncertainty principle, in its verbal formulation states the following. For certain pairs of observables the measurement results, when both observables are measured, can not both be fixed with an arbitrary precision. There are states, where the results for the measurement of one observable are certain with arbitrary precision, but then the other observable is inevitable afflicted with a nonzero uncertainty.

For a mathematical formulation of this physical fact one has to define what the uncertainty of an observable should be in mathematical terms. There are mainly two ways of measuring the uncertainty of an observable, thus, there are two different approaches to derive uncertainty relations.

The first of these ways uses the variance of an observable as a measure of the uncertainty. The variance of an observable $M$ is defined by

$$\delta^2(M)_\varrho := \langle (M - \langle M \rangle_\varrho)^2 \rangle_\varrho = \langle M^2 \rangle_\varrho - \langle M \rangle_\varrho^2. \tag{3.1}$$

and the standard deviation $\delta = \sqrt{\delta^2}$ by its positive square root. We often suppress the dependence on $\varrho$ in our notation, when there is no risk of confusion. Note that if $\varrho$ is a pure state the variance is zero iff $\varrho$ is an eigenstate of $M$. If we have two observables which do not share a common eigenstate, the variances of both observables cannot both equal zero. Historically, the first uncertainty relation was derived for the observables position ($X$) and momentum ($P$) of a particle [104]. It reads[1]:

$$\delta(P)\delta(X) \geq \frac{\hbar}{2}. \tag{3.2}$$

This relation has been later generalized to arbitrary observables [105], then it reads $\delta(A)\delta(B) \geq |\langle[A,B]\rangle|/2$. Also a generalization to more than two observables has been provided [106]. One may express the fact that some observables $A_i$ do not share a common eigenstate also as

$$\sum_i \delta^2(A_i) \geq C > 0. \tag{3.3}$$

Here, the calculation of the constant $C$ is in general not straightforward. However, as we will show in the later sections of this chapter, this formulation will be useful to derive separability criteria in terms of uncertainty relations.

The second formulation of the uncertainty principle uses entropies as measures of the uncertainty [107]. Measuring an observable $M$ yields a probability distribution of the outcomes. Now one can measure the uncertainty of the observable by the entropy $S(M)$ of this probability distribution. If two observables $M_1$ and $M_2$ do not share common eigenstate, there must be a constant such that

$$S(M_1) + S(M_2) \geq C > 0 \tag{3.4}$$

holds. Equations of this type can also serve as a formulation of the uncertainty principle. Since their discussion presupposes a discussion of possible entropies, we will discuss these so called entropic uncertainty relations later in Section 3.8.

Finally, we want to stress that the validity of uncertainty relations does not rely on the disturbance of the system during the measurement process. In their derivation no reference is made to the collapse of the wave function or similar concepts. Uncertainty relations express simply the mathematical fact that some observables do not share a common eigenstate. Thus, their validity is not restricted to single systems. If many copies of a particle in the same quantum state are given, we may measure on some of them the position and on the others the momentum. However, for the variances of these measurements still Eq. (3.2) is valid.

## 3.3 Variance based criteria

The idea to detect entanglement with variances is based on the following concavity property of the variance.

[1] Note that in the older literature the standard deviation is often defined with a different factor $\delta_{old} := \sqrt{2}\delta$.

**Lemma 3.1.** The variance of an observable is concave in the state. Technically speaking: Let $M_i$ be some observables and $\varrho = \sum_k p_k \varrho_k$ be a convex combination of some states $\varrho_k$. Then

$$\sum_i \delta^2(M_i)_\varrho \geq \sum_k p_k \sum_i \delta^2(M_i)_{\varrho_k} \tag{3.5}$$

holds.

*Proof.* This fact can be proven by a simple calculation. The inequality holds for each $M_i$: $\delta^2(M_i)_\varrho = \sum_k p_k \langle (M_i - \langle M_i \rangle_\varrho)^2 \rangle_{\varrho_k} = \sum_k p_k (\delta^2(M_i)_{\varrho_k} + (\langle M_i \rangle_{\varrho_k} - \langle M_i \rangle_\varrho)^2) \geq \sum_k p_k \delta^2(M_i)_{\varrho_k}$. □

Note that this lemma has a clear physical meaning: One cannot decrease the uncertainty of an observable by mixing several states.

To detect entanglement with this property, we first have to chose some observables $M_i$ and look at the sum of the variances of these observables. Then, we compute a lower bound of this quantity for the pure product states. Due to the lemma from above, the bound will also hold for all separable states. If we find a state which violates this bound, the state must be entangled. For simplicity, we introduce the following definition for these entangled states:

**Definition 3.2.** We call a state $\varrho$ *violating the variance criterion* if there exist $M_i$ and a constant $C > 0$ such that

$$\sum_i \delta^2(M_i)_\varrho < C, \tag{3.6}$$

while for all separable states $\varrho_s$

$$\sum_i \delta^2(M_i)_{\varrho_s} > C \tag{3.7}$$

holds.

A state which violates the variance criterion has to be entangled, and the entanglement of this state can be detected by measuring the variances of the observables $M_i$. The violation of this criterion can be viewed as a violation of some uncertainty relation of the type (3.3), since the condition (3.7) is nothing but such a uncertainty relation with a special bound for the separable states. This bound on the variances of observables can also be interpreted as a nonlinear entanglement witness.

The problem in this approach is twofold: First, one has to identify the appropriate $M_i$ for a given state which one want to detect. Second, one has to compute the lower bound for the product states.

## 3.4 Local uncertainty relations

The first work in which variances were used to detect entanglement in finite-dimensional systems was, to our knowledge, done by Holger Hofmann and Shigeki

Takeuchi [102]. They called their approach the *local uncertainty relations* (LURs). We will discuss it now.

Let $A_i$ be observables on Alice's space of a bipartite system. If the $A_i$ do not share a common eigenstate, there exists, as already mentioned, a number $U_A > 0$ such that

$$\sum_i \delta^2(A_i)_{\varrho_A} \geq U_A \tag{3.8}$$

holds for all states $\varrho_A$ on Alice's space. Hofmann and Takeuchi showed that uncertainty relations of this type are useful tools for the detection of entanglement:

**Proposition 3.3.** Let $\varrho$ be separable and let $A_i, B_i, i = 1, ..., n$ be operators on Alice's (respectively, Bob's) space, fulfilling $\sum_{i=1}^n \delta^2(A_i)_{\varrho_A} \geq U_A$ and $\sum_{i=1}^n \delta^2(B_i)_{\varrho_B} \geq U_B$. We define $M_i := A_i \otimes \mathbb{1} + \mathbb{1} \otimes B_i$. Then

$$\sum_{i=1}^n \delta^2(M_i)_\varrho \geq U_A + U_B \tag{3.9}$$

holds.

*Proof.* Their proof used the fact that for all product states $\delta^2(M_i)_{\varrho_A \otimes \varrho_B} = \delta^2(A_i)_{\varrho_A} + \delta^2(B_i)_{\varrho_B}$ holds. Then the concavity property of the variance proves the claim. □

Let us discuss now the applications of this method. We start with bipartite system, for simplicity we consider only two qubits. Then we investigate whether it is possible to extend the LURs to multipartite systems.

For a single-qubit system is has been shown in [102] that for the Pauli matrices the uncertainty relation

$$\sum_{i=x,y,z} \delta^2(\sigma_i) \geq 2 \tag{3.10}$$

holds. We give a proof differing from the one in Ref. [102] later in this section. Defining $M_i = \sigma_i \otimes \mathbb{1} + \mathbb{1} \otimes \sigma_i$ this yields the LUR

$$\sum_{i=x,y,z} \delta^2(M_i) \geq 4. \tag{3.11}$$

A short calculation shows that from this equation it follows that for all separable states

$$\langle \mathbb{1} \otimes \mathbb{1} + \sigma_x \otimes \sigma_x + \sigma_y \otimes \sigma_y + \sigma_z \otimes \sigma_z \rangle - \frac{1}{2} \sum_{i=x,y,z} \langle \sigma_i \otimes \mathbb{1} + \mathbb{1} \otimes \sigma_i \rangle^2 \geq 0 \tag{3.12}$$

has to hold. This is a quite remarkable equation for the following reason: The first part which is linear in the expectation values is known to be an optimal entanglement witness (see Section 2.4.1 and Eq. (2.10)). From this witness some quadratic terms are subtracted. Thus, in this case, the LUR can be viewed as a nonlinear entanglement witness which improves the linear entanglement witness.

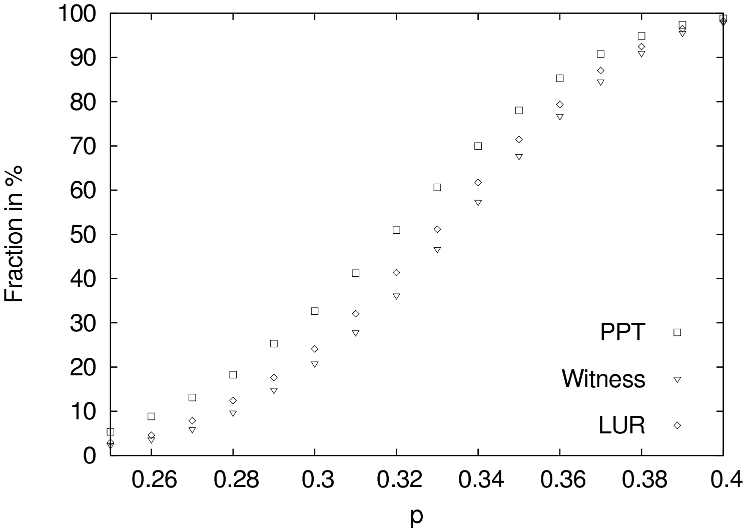

Figure 3.1. *A comparison between the LUR in Eq. (3.12), the witness from Eq. (2.10), and the PPT criterion for states of the type (3.13) with $d = 0.2$. In dependence on $p$ the fraction of states which are detected via the different criteria is shown.*

It is natural to ask at this point how big the improvement by the subtraction is. Let us investigate this issue numerically. To this aim, we look at states of the form

$$\varrho(p,d) := p|\psi^-\rangle\langle\psi^-| + (1-p)\sigma, \tag{3.13}$$

where[2]

$$\|\sigma - \frac{1}{4}\mathbb{1}\| \leq d. \tag{3.14}$$

Physically, one can view these states as mixture of a singlet state and some separable noise, the parameter $p$ determines the fidelity of the singlet state, and the parameter $d$ the properties of the noise: For $d = 0$ the noise consists of white noise. The set of matrices $\varrho(p,d)$ governs a ball in the space of all matrices. We take the value $d = 0.2$ and generate matrices of the form $\varrho(p, 0.2)$ randomly distributed in this ball, using the methods of Ref. [109]. Then we investigate the separability properties of these matrices: We determine the fraction of matrices which are detected by the witness, the LUR, and the PPT criterion, which is necessary and sufficient for entanglement in this case. Of course, these fractions depend on the value of $p$. The results are shown in Figure 3.1. On can clearly see, that the LUR improves the witness significantly, although it is not capable of detecting all states.

To discuss the possible extensions of LURs to multipartite systems, we will focus on three qubits. As it will turn out, already there exist many problems in extending them. We can first state a simple extension:

**Proposition 3.4.** Let $A_i$, $B_i$ and $C_i$ be some operators on Alice's, respectively,

[2] We use the Hilbert Schmidt norm in the space of all operators here, *i.e.*, $\|A\| = \sqrt{Tr(AA^\dagger)}$.

Bob's or Charlie's Hilbert space fulfilling

$$\sum_i \delta^2(A_i) \geq U_A, \quad \sum_i \delta^2(B_i) \geq U_B, \quad \sum_i \delta^2(C_i) \geq U_C, \tag{3.15}$$

and define $M_i := A_i \otimes \mathbb{1} \otimes \mathbb{1} + \mathbb{1} \otimes B_i \otimes \mathbb{1} + \mathbb{1} \otimes \mathbb{1} \otimes C_i$. Then for a fully separable state $\rho = \sum_k p_k \rho_k^A \otimes \rho_k^B \otimes \rho_k^C$ the LUR

$$\sum_i \delta^2(M_i) \geq U_A + U_B + U_C. \tag{3.16}$$

holds. For any biseparable state the LUR

$$\sum_i \delta^2(M_i) \geq \min\{U_A, U_B, U_C\} \tag{3.17}$$

is valid.

*Proof.* These statements can be derived by writing $\rho$ as a convex combination of the appropriate (bi)separable states. Then the calculation is similar the the proof of the LURs. □

Although these criteria are a straightforward extension of the LURs to multipartite systems, they have a big disadvantage: They seem to be quite weak. To give an example, let us look at $A_i = B_i = C_i = \sigma_i, i = x, y, z$. Then it is easy to calculate that the states $|W\rangle$ and $|GHZ\rangle$ violate neither Eq. (3.17), nor Eq. (3.16). The bipartite entanglement in the biseparable state $|\phi\rangle = (|101\rangle - |011\rangle)/\sqrt{2}$ is detected via Eq. (3.16) and for this state also equality holds in (3.17). However, we were unable to find any state which violates Eq. (3.17)[3]. Thus, the LURs in the formulation of Proposition 3.4 seem to fail to detect true tripartite entanglement.

Another idea to investigate multipartite entanglement with LURs which was already proposed in [102], is to investigate the bipartite splittings separately. One has to stress that with this strategy one cannot prove true multipartite entanglement. This is due to the fact that there are biseparable states which are not separable if one looks at a fixed bipartite splitting.

The investigation of this bipartite splittings is, however, interesting, since it leads again to the improvement of already known witnesses by subtracting quadratic terms. We give here a simple three-qubit example, designed for the investigation of a GHZ state.

**Lemma 3.5.** The following uncertainty relations hold for one-, respectively, two-qubit systems:

$$\delta^2(-\sigma_x) + \delta^2(-\sigma_z) + \delta^2(-\mathbb{1}) \geq 1, \tag{3.18}$$

$$\delta^2(-\sigma_x) + \delta^2(-\sigma_z) + \delta^2(-\sigma_z) \geq 1, \tag{3.19}$$

$$\delta^2(\sigma_x \otimes \sigma_x) + \delta^2(\sigma_z \otimes \mathbb{1}) + \delta^2(\sigma_z \otimes \sigma_z) \geq 1, \tag{3.20}$$

$$\delta^2(\sigma_x \otimes \sigma_x) + \delta^2(\sigma_z \otimes \mathbb{1}) + \delta^2(\mathbb{1} \otimes \sigma_z) \geq 1. \tag{3.21}$$

[3] To this aim we generated numerically randomly distributed mixed states as described in [109] and we also generated randomly pure states, but no violation of Eq. (3.17) was found.

Since the proof is a little bit technical, we will present it later. First, we can state:

**Proposition 3.6.** Let $\varrho$ be a three-qubit state. If $\varrho$ is biseparable with respect to the $A|BC$ partition, then

$$\begin{aligned}
&2 - \langle \sigma_x \otimes \sigma_x \otimes \sigma_x \rangle - \langle \sigma_z \otimes \sigma_z \otimes \mathbb{1} \rangle - \langle \mathbb{1} \otimes \sigma_z \otimes \sigma_z \rangle - \\
&-\frac{1}{2}\big( \langle \sigma_x \otimes \mathbb{1} \otimes \mathbb{1} - \mathbb{1} \otimes \sigma_x \otimes \sigma_x \rangle^2 + \langle \sigma_z \otimes \mathbb{1} \otimes \mathbb{1} - \mathbb{1} \otimes \sigma_z \otimes \mathbb{1} \rangle^2 + \\
&+ \langle \mathbb{1} \otimes \mathbb{1} \otimes \mathbb{1} - \mathbb{1} \otimes \sigma_z \otimes \sigma_z \rangle^2 \big) \geq 0
\end{aligned} \tag{3.22}$$

holds. If $\varrho$ is biseparable with respect to the $B|AC$ partition, then

$$\begin{aligned}
&2 - \langle \sigma_x \otimes \sigma_x \otimes \sigma_x \rangle - \langle \sigma_z \otimes \sigma_z \otimes \mathbb{1} \rangle - \langle \mathbb{1} \otimes \sigma_z \otimes \sigma_z \rangle - \\
&-\frac{1}{2}\big( \langle \mathbb{1} \otimes \sigma_x \otimes \mathbb{1} - \sigma_x \otimes \mathbb{1} \otimes \sigma_x \rangle^2 + \langle \sigma_z \otimes \mathbb{1} \otimes \mathbb{1} - \mathbb{1} \otimes \sigma_z \otimes \mathbb{1} \rangle^2 + \\
&+ \langle \mathbb{1} \otimes \mathbb{1} \otimes \sigma_z - \mathbb{1} \otimes \sigma_z \otimes \mathbb{1} \rangle^2 \big) \geq 0
\end{aligned} \tag{3.23}$$

is valid. Finally, if $\varrho$ is biseparable with respect to the $C|AB$ partition, then

$$\begin{aligned}
&2 - \langle \sigma_x \otimes \sigma_x \otimes \sigma_x \rangle - \langle \sigma_z \otimes \sigma_z \otimes \mathbb{1} \rangle - \langle \mathbb{1} \otimes \sigma_z \otimes \sigma_z \rangle - \\
&-\frac{1}{2}\big( \langle \mathbb{1} \otimes \mathbb{1} \otimes \sigma_x - \sigma_x \otimes \sigma_x \otimes \mathbb{1} \rangle^2 + \langle \mathbb{1} \otimes \mathbb{1} \otimes \mathbb{1} - \sigma_z \otimes \sigma_z \otimes \mathbb{1} \rangle^2 + \\
&+ \langle \mathbb{1} \otimes \sigma_z \otimes \mathbb{1} - \mathbb{1} \otimes \mathbb{1} \otimes \sigma_z \rangle^2 \big) \geq 0
\end{aligned} \tag{3.24}$$

holds.

It is remarkable that the linear part of these inequalities is again a witness, already known from Section 2.7. It is the witness for the GHZ state given in Eq. (2.61) for the three-qubit case. This witness is improved by subtracting quadratic terms. Thus, it detects some states in addition.

*Proof of Lemma 3.5.* The proofs of these inequalities are the following. We write all variances as $\delta^2(M) = \langle M^2 \rangle - \langle M \rangle^2$. To prove then (3.18) we only have to show that $\langle \sigma_x \rangle^2 + \langle \sigma_z \rangle^2 \leq 1$ holds for all one-qubit states. This is clear from the geometry of the Bloch sphere, but we will give another proof which can be extended to the other cases. Let $\lambda = (\lambda_1, \lambda_2)$ be a real vector of length one and define $\Lambda = \lambda_1 \sigma_x + \lambda_2 \sigma_z$. It is easy to see that the maximal eigenvalue of $\Lambda$ equals one, thus $\lambda_1 \langle \sigma_x \rangle + \lambda_2 \langle \sigma_z \rangle \leq 1$ for all $\lambda$. This implies that the length of the vector $(\langle \sigma_x \rangle, \langle \sigma_z \rangle)$ is smaller than one, from which $\langle \sigma_x \rangle^2 + \langle \sigma_z \rangle^2 \leq 1$ follows. Eq. (3.19) follows directly from Eq. (3.18) and the fact that the variance is positive. Eq. (3.20) is proven similar as Eq. (3.18): We have to show that $\langle \sigma_x \otimes \sigma_x \rangle^2 + \langle \sigma_z \otimes \mathbb{1} \rangle^2 + \langle \mathbb{1} \otimes \sigma_z \rangle^2 \leq 2$. Again, we take a vector $\lambda = (\lambda_1, \lambda_2, \lambda_3)$ of length one and look at $\Lambda = \lambda_1 \sigma_x \otimes \sigma_x + \lambda_2 \sigma_z \otimes \mathbb{1} + \lambda_3 \mathbb{1} \otimes \sigma_z$. The largest eigenvalue of $\Lambda$ equals $\sqrt{1 + 2\lambda_2\lambda_3} \leq \sqrt{2}$. As before this implies that the length of $(\langle \sigma_x \otimes \sigma_x \rangle, \langle \sigma_z \otimes \mathbb{1} \rangle, \langle \mathbb{1} \otimes \sigma_z \rangle)$ is smaller than $\sqrt{2}$ which proves the claim. Eq. (3.21) can be proven similarly. Note that the uncertainty relation (3.10) can also be proved in this way. □

*Proof of Proposition 3.6.* The inequalities follow directly from the scheme of the LUR applied to the corresponding partitions and the Lemma 3.5. □

To conclude, we have seen that LURs are strong criteria to detect bipartite entanglement. Nevertheless the extension to multipartite systems is not so clear: It is not known up to now whether they are able to detect true multipartite entanglement. They have some further disadvantages: it is not clear which operators $A_i$ and $B_i$ one should choose to detect a given entangled state. Also, LURs can by construction characterize separable states only; they do not apply for other convex sets. In order to overcome these disadvantages, we introduce in the next section a more general approach which is dealing with variances of nonlocal observables.

## 3.5 Nonlocal uncertainty relations

A way to overcome the disadvantages of the local uncertainty relations is to consider *nonlocal* observables. For an experimental implementation one can decompose any nonlocal observable into local operators, as described in the previous chapter. The main idea to construct these nonlocal observables, is, to take them as projectors onto entangled states. With this we can also design uncertainty relations allowing the detection of a given state. Let us start with explaining this idea in the case of two qubits:

**Proposition 3.7.** Let $|\psi_1\rangle = a|00\rangle + b|11\rangle$ be an entangled two-qubit state written in the Schmidt decomposition, with $a \geq b$. Then, there exist $M_i$ such that for $|\psi_1\rangle$, $\sum_i \delta^2(M_i)_{|\psi_1\rangle\langle\psi_1|} = 0$ holds, while for separable states

$$\sum_i \delta^2(M_i) \geq 2a^2b^2 \tag{3.25}$$

is fulfilled. Thus, $|\psi_1\rangle$ violates the variance criterion. The $M_i$ are explicitly given (see below).

*Proof.* We define $|\psi_2\rangle = a|01\rangle + b|10\rangle$; $|\psi_3\rangle = b|01\rangle - a|10\rangle$; $|\psi_4\rangle = b|00\rangle - a|11\rangle$, and further $M_i := |\psi_i\rangle\langle\psi_i|, i = 1, ..., 4$. Then $\sum_i \delta^2(M_i)_{|\psi_1\rangle\langle\psi_1|} = 0$. We only need to prove the bound (3.25) for a pure product vector $|\phi\rangle$. We have $\sum_{i=1}^4 \delta^2(M_i)_{|\phi\rangle\langle\phi|} = 1 - \sum_i (|\langle\phi|\psi_i\rangle|^2)^2$. Now, we use the fact that $|\langle\phi|\psi_i\rangle|^2 \leq a^2$. This bound on the maximal squared overlap follows from the properties of the Schmidt decomposition, it has been proved in Section 2.6. With this bound we have $\sum_i (|\langle\phi|\psi_i\rangle|^2)^2 \leq (a^2)^2 + (1 - a^2)^2$ and finally $\delta^2(M_i)_{|\phi\rangle\langle\phi|} \geq 2a^2b^2$. □

Note that Eq. (3.25) allows to detect a mixed state of the form $\varrho(p) = p|\psi_1\rangle\langle\psi_1| + (1 - p)\mathbb{1}/4$ for $p > \sqrt{1 - 8a^2b^2/3}$. These are not all states of this family, since the states are entangled for $p > 1/(1 + 4ab)$. We will discuss this in more detail in Section 3.6.

This two-qubit example reveals already a main idea for the construction of the $M_i$ : We take one $M_i$ as the projector onto the range of $\varrho$, and the others as projectors onto a basis of the kernel of $\varrho$. This basis should contain only entangled vectors. As we will see now, this concept suffices to detect all pure entangled states, a family of bound entangled states and multipartite entanglement.

**Proposition 3.8.** Let $|\psi_1\rangle$ be an entangled pure state in a bipartite $N \times M$-system. Then $|\psi_1\rangle$ violates the variance criterion for properly chosen $M_i$.

*Proof.* Let $U$ be the space orthogonal to $|\psi_1\rangle$. It is clear that $U$ contains at least one entangled vector $|\psi_e\rangle$. We can choose a basis $|\psi_i\rangle, i = 2, ..., NM$ of $U$ which consists only of entangled vectors. (To do this, we choose an arbitrary, not necessarily orthogonal, basis. If it contains product vectors, we perturb them by adding $\varepsilon|\psi_e\rangle$). Then we take $M_i = |\psi_i\rangle\langle\psi_i|$. The only possible common eigenstates of the $M_i$ are the $|\psi_i\rangle$. So for product states the sum over all uncertainties is bounded from below, while it is zero for $|\psi_1\rangle$. □

The class of bound entangled states which we want to consider are those arising from an *unextendible product basis* (UPB) [38]. They can be constructed as follows: Let $\{|\phi_i\rangle = |e_i\rangle|f_i\rangle, i = 1, ..., n\}$ a set of product vectors in a bipartite $N \times M$-system, where all $|\phi_i\rangle$ be pairwise orthogonal, and do not span the whole space. If there is no other product vector orthogonal to all $|\phi_i\rangle$ they are said to form an unextendible product basis[4]. Then the state

$$\varrho_{UPB} := \frac{1}{NM - n}(\mathbb{1} - \sum_i |\phi_i\rangle\langle\phi_i|) \tag{3.26}$$

is a bound entangled state: It is clear that $\varrho_{UPB}$ is PPT since we have

$$\varrho_{UPB}^{T_A} = \frac{1}{NM - n}(\mathbb{1} - \sum_i |e_i^*\rangle\langle e_i^*| \otimes |f_i\rangle\langle f_i|) \geq 0. \tag{3.27}$$

Furthermore, they are entangled. They violate the range criterion, introduced in Section 1.1.3 in an extreme manner. There is no product vector in the range of $\varrho_{UPB}$, thus there is no set of product vectors spanning this range. However, all the UPB states can also be detected with our variance based method.

**Proposition 3.9.** Let $\varrho_{UPB}$ be a bound entangled state, constructed from a UPB. Then $\varrho_{UPB}$ violates the variance criterion for appropriate $M_i$.

*Proof.* Let $U$ be the subspace spanned by the UPB. Then $U$ must contain at least one entangled vector. To see this, note that subspaces containing only product vectors are of the form $\{|v\rangle = |a\rangle|b\rangle\}$ with a fixed $|a\rangle$ (or $|b\rangle$) for all $|v\rangle$. But then there would be product vectors orthogonal to $U$. Due to the existence of an entangled vector in $U$, there is, according to the proof of Proposition 3.8, an entangled basis $|\psi_i\rangle; i = 1, ..., n$ of $U$. We take $M_i = |\psi_i\rangle\langle\psi_i|; i = 1, ..., n$ and $M_{n+1} = \mathbb{1} - \sum_i |\phi_i\rangle\langle\phi_i|$. The common eigenstates of the $M_i$ are not product vectors, and we have $\sum_i \delta^2(M_i)_{\varrho_{UPB}} = 0$. □

Now we demonstrate that one can also detect true multipartite entanglement with uncertainties. We give a three-qubit example, allowing the detection of GHZ states. The inequalities are formulated so, that they can be checked with local measurements:

[4] This terminology is a little bit sloppy, since they do not form a basis of the whole space.

**Proposition 3.10.** Let $\varrho$ be a three-qubit state. We define a sum of variances as

$$\begin{aligned} E(\varrho) \ := \ & \frac{7}{8} - \frac{1}{8}\big(\langle \mathbb{1} \otimes \sigma_z \otimes \sigma_z\rangle^2 + \langle \sigma_z \otimes \mathbb{1} \otimes \sigma_z\rangle^2 + \langle \sigma_z \otimes \sigma_z \otimes \mathbb{1}\rangle^2 + \\ & + \langle \sigma_x \otimes \sigma_x \otimes \sigma_x\rangle^2 + \langle \sigma_x \otimes \sigma_y \otimes \sigma_y\rangle^2 + \langle \sigma_y \otimes \sigma_y \otimes \sigma_x\rangle^2 + \\ & + \langle \sigma_y \otimes \sigma_x \otimes \sigma_y\rangle^2\big)\,. \end{aligned} \tag{3.28}$$

If $E < 1/2$, the state $\varrho$ is fully tripartite entangled. If $E < 3/8$, the state $\varrho$ belongs even to the GHZ class. For the GHZ state $|GHZ\rangle = 1/\sqrt{2}(|000\rangle + |111\rangle)$ we have $E = 0$.

*Proof.* We define the eight states $|\psi_{1/5}\rangle = (|000\rangle \pm |111\rangle)/\sqrt{2}$; $|\psi_{2/6}\rangle = (|100\rangle \pm |011\rangle)/\sqrt{2}$; $|\psi_{3/7}\rangle = (|010\rangle \pm |101\rangle)/\sqrt{2}$; $|\psi_{4/8}\rangle = (|001\rangle \pm |110\rangle)/\sqrt{2}$ and as usual $M_i = |\psi_i\rangle\langle\psi_i|, i = 1, ..., 8$ and $E = \sum_i \delta^2(M_i)$. Then the proof is quite similar to the proof of Proposition 3.7. One only needs the maximal squared overlaps, which we know from the previous chapter: We have $\max_{|\phi\rangle\in BS} |\langle GHZ|\phi\rangle|^2 = 1/2$, and $\max_{|\phi\rangle\in W} |\langle GHZ|\phi\rangle|^2 = 3/4$ [46]. By decomposing the $M_i$ similar to the preceding chapter one gets $E$ in the form of Eq. (3.28). □

Let us compare the inequality (3.28) with the witness $\mathcal{W}_3^{(\mathrm{ghz})} = 1/2 \cdot \mathbb{1} - |GHZ\rangle\langle GHZ|$ and $\mathcal{W}_{GHZ} = 3/4 \cdot \mathbb{1} - |GHZ\rangle\langle GHZ|$. Assuming states of the type $\varrho(p) = p|GHZ\rangle\langle GHZ| + (1-p)\mathbb{1}/8$ the inequality (3.28) detects them for $p \geq \sqrt{3/7} \approx 0.65$ as tripartite entangled, and for $p \geq \sqrt{4/7} \approx 0.76$ as GHZ states. The witnesses detect them for $p \geq 3/7 \approx 0.43$ as tripartite entangled and for $p \geq 5/7 \approx 0.71$ as belonging to the GHZ class. Thus, inequality (3.28) seems to be slightly weaker. We will discuss the relationship between the witnesses and the variances in more detail in the next section.

The method presented here can be extended to more parties. For four qubits one can write down 16 orthogonal GHZ states of the type $|\psi_i\rangle = (|x_1^{(1)} x_1^{(2)} x_1^{(3)} x_1^{(4)}\rangle \pm |x_2^{(1)} x_2^{(2)} x_2^{(3)} x_2^{(4)}\rangle)/\sqrt{2}$ with $x_l^{(k)} \in \{0, 1\}$ and $x_1^{(k)} \neq x_2^{(k)}$. If we then define $M_i = |\psi_i\rangle\langle\psi_i|$ it follows as in the proof of Proposition 3.10 that for all biseparable states $\sum_i \delta^2(M_i) \geq 1/2$ holds. The same idea can be used to construct uncertainty relations for an arbitrary number of qubits, but the number of $M_i$ increases exponentially.

## 3.6 Comparison with witnesses

In this section, we want to compare the variance based criteria with entanglement witnesses. We also want to give some geometrical picture of the states which are detected by uncertainty relations. Let us first derive a very general uncertainty relation:

**Proposition 3.11.** Let $|\psi_i\rangle$ be an orthonormal basis of a bipartite Hilbert space. Let $c < 1$ be an upper bound for all the squared Schmidt coefficients of all $|\psi_i\rangle$. Then

$$\sum_i \delta^2(|\psi_i\rangle\langle\psi_i|) \geq 1 - \lfloor 1/c \rfloor c^2 - (1 - \lfloor 1/c \rfloor c)^2 \tag{3.29}$$

holds for all separable states[5].

*Proof.* The proof is a straightforward extension of the propositions from the last section. The bound on the squared Schmidt coefficients gives a bound on the overlaps. First, we have $\sum_i \delta^2(|\psi_i\rangle\langle\psi_i|) = 1 - \sum_i \langle\psi_i|\varrho|\psi_i\rangle^2$. The sum $\sum_i \langle\psi_i|\varrho|\psi_i\rangle^2$ is maximal, when the single terms are as big as possible, *i.e.* $\lfloor 1/c \rfloor$ of the $\langle\psi_i|\varrho|\psi_i\rangle$ satisfy the bound $\langle\psi_i|\varrho|\psi_i\rangle = c$, while one other $\langle\psi_i|\varrho|\psi_i\rangle$ is as big as possible. This proves the claim. We will give a proof of more general estimates based on convexity arguments later, see Theorem 3.21. □

The estimate of this proposition may be refined, if individual bounds on the Schmidt coefficients of each $|\psi_i\rangle$ are known. Also, the bound is in general not sharp. The proof of this proposition implies that the separability criterion given in Eq. (3.29) is not stronger than the witnesses $\mathcal{W}^{(i)} = c \cdot \mathbb{1} - |\psi_i\rangle\langle\psi_i|$. This is easy to see: If a state $\varrho$ cannot be detected by these witnesses we have $\langle\psi_i|\varrho|\psi_i\rangle \leq c$ for all $i$. Thus, $\varrho$ does also not violate the condition (3.29).

To give a simple example of the geometrical interpretation of the uncertainty based separability criteria, let us look at two qubits. We take $|\psi_1\rangle$ in Proposition 3.7 as the Bell state $|\psi_1\rangle = (|01\rangle - |10\rangle)/\sqrt{2}$, and the other $|\psi_i\rangle$ as the other Bell states. Let us denote them by $|BS_1\rangle = (|00\rangle + |11\rangle)/\sqrt{2}, |BS_2\rangle = (|00\rangle - |11\rangle)/\sqrt{2}, |BS_3\rangle = (|01\rangle + |10\rangle)/\sqrt{2}, |BS_4\rangle = (|01\rangle - |10\rangle)/\sqrt{2}$. Proposition 3.7 ensures now that for separable states the quadratic inequality

$$1 - \sum_i Tr(|BS_i\rangle\langle BS_i|)^2 \geq \frac{1}{2} \tag{3.30}$$

holds. We can write this inequality in the language of local measurement settings: If we define $i = Tr(\varrho\sigma_i \otimes \sigma_i)$ for $i = x, y, z$, we find $\langle BS_1|\varrho|BS_1\rangle = (1 + x - y + z)/4; \langle BS_2|\varrho|BS_2\rangle = (1-x+y+z)/4; \langle BS_3|\varrho|BS_3\rangle = (1+x+y-z)/4; \langle BS_4|\varrho|BS_4\rangle = (1-x-y-z)/4$. A straightforward calculation proves that in this notation, Eq. (3.30) reads

$$x^2 + y^2 + z^2 \leq 1. \tag{3.31}$$

This is the equation of a three-dimensional sphere. To compare this with the witnesses, let us look at the witnesses $\mathcal{W}^{(i)} = 1/2 \cdot \mathbb{1} - |BS_i\rangle\langle BS_i|$ for $i = 1, ..., 4$. A separable state has to fulfill $Tr(\varrho\mathcal{W}^{(i)}) \geq 0, i = 1, ..., 4$. Following the lines of Section 2.4.1 one can calculate that this requires

$$\begin{aligned} x + y + z \leq 1, &\qquad -x - y + z \leq 1, \\ x - y - z \leq 1, &\qquad -x + y - z \leq 1. \end{aligned} \tag{3.32}$$

These inequalities describe a tetrahedron in the three-dimensional space. Furthermore a physical state has to be positive, which implies that $Tr(\varrho|BS_i\rangle\langle BS_i|) \geq 0$ holds, for $i = 1, ..., 4$. This requires

$$\begin{aligned} x - y + z \leq 1, &\qquad -x + y + z \leq 1, \\ x + y - z \leq 1, &\qquad -x - y - z \leq 1. \end{aligned} \tag{3.33}$$

[5] The bracket $\lfloor x \rfloor$ denotes the integer part of $x$.

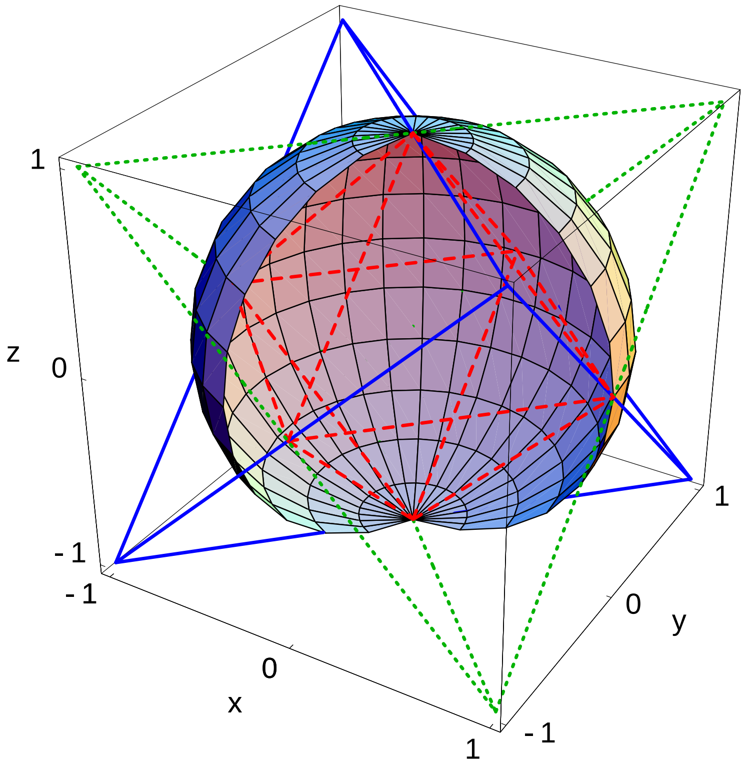

Figure 3.2. *Geometrical picture for the uncertainty based criteria for two qubits. The criteria from the Eqs. (3.31, 3.32, 3.33) are depicted in dependence on the coordinates $x, y$, and $z$. One tetrahedron (blue, solid lines) is the set of all states (3.33), the other tetrahedron (green, dotted lines) is the set of not necessarily positive matrices, which obey (3.32). The octahedron (red, dashed lines) is the intersection of these tetrahedra, it corresponds to the set of separable states. The ball represents the states which obey (3.31), all the states outside this ball are detected by the uncertainty relation.*

This is again a description of a tetrahedron. Eqs. (3.32, 3.33) together describe an octahedron as the intersection of two tetrahedra[6]. The proof of the Proposition 3.11 implies that if a state obeys (3.32) as well as (3.33), it also fulfills (3.31). Thus, the uncertainty criterion does not detect any states which are not detected by the witnesses $\mathcal{W}^{(i)}$. This is also depicted in Figure 3.2. Later, in Section 3.9 we will develop criteria, which interpolate between the inequalities (3.31) and (3.32). See also Figure 3.3.

When making a more sophisticated comparison between the general uncertainty relations and the witnesses, one has to be careful which questions one likes to pose. Since witnesses are capable of detecting all states, the uncertainty based methods can never be better in the sense that they detect more states.

[6] This geometrical picture was also studied in Refs. [110, 111].

A more interesting question is, if variances allow to detect more states than witnesses, if only a restricted set of measurements is available. Again, on has to be careful: A variance can be measured in different ways. To determine $\delta^2(M)$ one can setup an apparatus for the measurement of $M$ and calculate the variance of this observable from the fluctuations of the outcomes. One might also measure only the expectation values from the two observables $M$ and $M^2$ and calculate then $\delta^2(M)$. Note that for some cases these ways are equivalent. This happens, e.g. for Pauli matrices and projectors, since $\sigma_i^2 = \mathbb{1}$ and $|\psi\rangle\langle\psi|^2 = |\psi\rangle\langle\psi|$.

In our cases, we usually have nonlocal observables. Thus, the first way of measuring $\delta^2(M)$ is not easy to perform, since we cannot setup the apparatus for the nonlocal observable. One has to measure $M$ and $M^2$ separately with the help of local decompositions.

Now we want to answer the question whether uncertainty relations can improve the detection of entanglement, when only a restricted set of observables is accessible and when the variance is measured via determining $\langle M^2\rangle$ and $\langle M\rangle$. Let $A_i, i = 1, ..., n$ be the observables which one can measure. Without loosing generality, we can assume that they are linearly independent. Otherwise, they would be redundant and some of them could be omitted. Measurement of the observables gives rise to a map

$$\mathfrak{A} : \mathcal{B}(\mathcal{H}) \to \mathbb{R}^n, \quad \varrho \mapsto \mathfrak{A}(\varrho) = (\langle A_1\rangle_\varrho, ..., \langle A_1\rangle_\varrho), \tag{3.34}$$

which maps a state onto the set of expectation values. This map is linear and not injective. It maps the set $S$ of separable states onto the set $S' := \mathfrak{A}(S)$. An entangled state $\varrho$ with the property $\mathfrak{A}(\varrho) \in S'$ can not be detected with this reduced set of observables: There is a separable state $\varrho_s$ having the same $\mathfrak{A}(\varrho_s)$, thus $\varrho$ and $\varrho_s$ are indistinguishable. We can state:

**Proposition 3.12.** If a state $\varrho_e$ can be detected via the measurements of the $\langle A_i\rangle$, then there is a witness $\mathcal{W} = \sum_i w_i A_i$ which detects the state.

*Proof.* Since $\mathfrak{A}$ is linear, the set $S'$ is also convex. We have $\mathfrak{A}(\varrho_e) \neq S'$, otherwise the state could not be detected. Now we can the usual construction of witnesses: There must exist a hyper-plane separating $\mathfrak{A}(\varrho_e)$ from $S'$. This means that there is a vector $w = (w_1, ..., w_n)$ with $\langle w|\mathfrak{A}(\varrho_e)\rangle < 0$ while $\langle w|\mathfrak{A}(\varrho_s)\rangle > 0$ for all separable $\varrho_s$[7]. The observable $\mathcal{W} = \sum_i w_i A_i$ is now the desired entanglement witness, since $Tr(\varrho\mathcal{W}) = \langle w|\mathfrak{A}(\varrho)\rangle$. □

This proposition states that even when only a restricted set of measurements is available, the uncertainty relations cannot detect more states. Even other nonlinear criteria can not be better than the witness.

To end with an example, let us look at the LUR in Eq. (3.12) again. At first sight it seemed that this inequality improves the witness. This is of course true, but in view of Proposition 3.12 we can conclude that there must be witnesses using the same measurement settings which detect all states which are detected by the LUR. But these witnesses are not easy to construct. Thus we finally arrive at the following

[7] $\langle w|\mathfrak{A}(\varrho)\rangle$ denotes here the scalar product in $\mathbb{R}^n$.

conclusion: in principle the witnesses are better than the variance based methods, but the latter are sometimes much easier to construct.

## 3.7 Connection with Gaussian states

In this section we first formulate a criterion for separability in terms of so-called covariance matrices. Then we show that it is equivalent to the criterion given in Definition 3.2. However, as we will see, this new covariance matrix criterion has the advantage that it allows us to connect the separability problem for finite-dimensional systems with the separability problem for infinite-dimensional systems.

Let $\varrho = \sum_k p_k \varrho_k^A \otimes \varrho_k^B$ be a separable state and let $M_i$, for $i = 1, ..., n$ be some observables. We can define two functionals $Z(\varrho)[X]$ and $W(\varrho)[X]$ with $X = (x_1, ..., x_n) \in \mathbb{R}^n$ by:

$$\begin{aligned} Z(\varrho)[X] &:= \langle e^{\sum_{i=1}^n x_i M_i} \rangle, \\ W(\varrho)[X] &:= \ln(Z(\varrho)[X]). \end{aligned} \tag{3.35}$$

Similar constructions are known from Quantum Field Theory, there $Z$ is called the generating functional of the moments and $W$ the generating functional of the cumulants [112]. Due to the linearity of $Z$ we have $Z(\varrho) = \sum_k p_k Z(\varrho_k^A \otimes \varrho_k^B)$, and due to the concavity of the logarithm it follows that

$$W(\varrho) - \sum_k p_k W(\varrho_k^A \otimes \varrho_k^B) \geq 0 \tag{3.36}$$

holds. For $X = 0$ equality holds in this equation. Thus, for this value of $X$ the left hand side of Eq. (3.36) has a minimum. This implies that the matrix of the second derivatives with respect to $X$ of this left hand side is positive semidefinite at $X = 0$ :

$$\Big(\partial_i \partial_j \big[W(\varrho) - \sum_k p_k W(\varrho_k^A \otimes \varrho_k^B)\big]\big|_{X=0}\Big) \geq 0, \tag{3.37}$$

where we have denoted $\partial_i := \partial/\partial x_i$. In this matrix inequality the matrix of the second derivatives of $W$ occurs. This is the so-called covariance matrix (CM) $\gamma(\varrho, M_i)$. Its entries are:

$$\gamma(\varrho, M_i)_{kl} := \partial_k \partial_l W[X]\big|_{X=0}. \tag{3.38}$$

We will mention some properties of the CM later. First, we can state:

**Lemma 3.13.** Let $\varrho$ be separable, and let $M_i$ be observables. Then there exist product states $\varrho_k^A \otimes \varrho_k^B$ and $p_k$ such that

$$\gamma(\varrho, M_i) \geq \sum_k p_k \gamma(\varrho_k^A \otimes \varrho_k^B, M_i) \tag{3.39}$$

holds. We call a state *violating the covariance matrix criterion* for the observables $M_i$ iff there are no product states $\varrho_k^A \otimes \varrho_k^B$ and convex weights $p_k$, such that Eq. (3.39) can be fulfilled.

**Proposition 3.14.** The matrix $\gamma$ has the following properties:
(a) The entries are given by

$$\gamma_{kl} = \frac{\langle M_k M_l \rangle + \langle M_l M_k \rangle}{2} - \langle M_k \rangle \langle M_l \rangle. \tag{3.40}$$

(b) We have for an arbitrary $(x_1, ..., x_n) \in \mathbb{R}^n$

$$\sum_{i,j=1}^{n} x_i \gamma_{ij} x_j = \delta^2 (\sum_{i=1}^{n} x_i M_i) \geq 0. \tag{3.41}$$

In particular we have $\gamma \geq 0$.
(c) The commutation relations of the $M_i$ require:

$$\gamma(\rho) \geq \Omega(\rho) \qquad \Omega(\rho)_{kl} := \frac{1}{2} \langle [M_k, M_l] \rangle. \tag{3.42}$$

We have also $\gamma(\rho) \geq -\Omega(\rho)$.

*Proof.* (a) can be directly calculated from the definition with the help of the formula [113]

$$\partial_t e^{A(t)} = \int_0^1 ds e^{sA} (\partial_t A) e^{-sA} e^A \tag{3.43}$$

and the power expansion of the logarithm. (b) can be proven by a direct calculation of $\delta^2(\sum_{i=1}^n x_i M_i)$. To show (c), consider the matrix $\Theta_{ij} = \langle M_i M_j \rangle - \langle M_i \rangle \langle M_j \rangle$. As in (b) also for $\Theta$ one can prove $\sum_{i,j=1}^n x_i \Theta_{ij} x_j = \delta^2(\sum_{i=1}^n x_i M_i)$, thus $\Theta \geq 0$. Furthermore, we have $\Theta = \gamma + \Omega$ and $\Theta^T = \gamma - \Omega$, which proves the claim. Note that (c) is exactly the formulation of the uncertainty principle for several variables established in Ref. [106]. □

Armed with these insights, we can prove now the main result of this section: The variance criterion and the covariance matrix criterion are equivalent.

**Theorem 3.15.** A state $\varrho$ violates the variance criterion if and only if it violates also the covariance matrix criterion. The observables leading to a violation may differ.

*Proof.* A state violating the variance criterion violates also the covariance matrix criterion with the same $M_i$, since the sum of all variances is the trace of $\gamma$. To show the other direction, let us assume that $\gamma$ violates the covariance matrix criterion and look at the set of symmetric matrices of the form $T := \{\sum_k p_k \gamma(\varrho_k^A \otimes \varrho_k^B, M_i) + P\}$ where $P \geq 0$ is positive. $T$ is convex and closed. We have due to our assumption $\gamma \notin T$. As in the construction of entanglement witnesses there exists a symmetric matrix $W$ and a number $C \geq 0$ such that $Tr(W\gamma) < C$ while $Tr(W\mu) > C$ for all $\mu \in T$. Since $Tr(W\mu) > C$ for all $\mu \in T$ we have $Tr(WP) \geq 0$ for all $P \geq 0$. This implies $W \geq 0$. Now we use the spectral decomposition and write $W_{ij} = \sum_l \lambda_l a_i^l a_j^l$ with $\lambda_l \geq 0$. Let us define $N_l = \sqrt{\lambda_l} \sum_i a_i^l M_i$. Then $Tr(W\gamma) = \sum_l \delta^2(N_l) < C$ according to Eq. (3.41), while for all convex combinations of product states $\sum_k p_k \sum_l \delta^2(N_l)_{\varrho_k^A \otimes \varrho_k^B} > C$ holds. This is a violation of the variance criterion for the $N_l$. □

The criterion in terms of covariance matrices gives us the possibility to relate the separability problem for finite-dimensional systems with the separability problem for infinite-dimensional systems. Infinite-dimensional systems occur when the Hilbert spaces of Alice and Bob are not finite-dimensional, e.g. if they own a harmonic oscillator, or a free particle. These can happen for instance in quantum optical systems, where the harmonic oscillators can be represented by different modes of light. In these situations one can, of course, ask again, whether the state they share is entangled or not. In many experimental situations one deals with a special kind of infinite-dimensional states, the so-called Gaussian states. Let us recall some of their properties. A good introduction in this subject is given in Ref. [114].

Let us assume that $n$ harmonic oscillators (modes) are given, where the first $m$ belong to Alice and the last $n-m$ to Bob. To each of the modes belongs a pair of canonical conjugate observables $X_i, P_i$, corresponding to position and momentum. These can be written as a vector $R = (X_1, P_1, X_2, P_2, ..., X_n, P_n)$, and then the Weyl operators for $x = (x_1, ..., x_{2n}) \in \mathbb{R}^{2n}$ are defined as

$$\mathfrak{W}(x) := e^{i \sum_k x_k R_k}. \tag{3.44}$$

A state $\varrho$ can be described by a characteristic function, which is defined as

$$\chi(x) := Tr(\varrho \mathfrak{W}(x)). \tag{3.45}$$

Gaussian states are defined by the property that the characteristic function is a Gaussian distribution, *i.e.*, we can write

$$\chi(x) = \exp\big( -\frac{1}{4} \sum_{kl} x_k \tilde{\gamma}_{kl} x_k + i \sum_k d_k x_k \big). \tag{3.46}$$

Here, the vector $d$ is the so-called displacement of the state and $\tilde{\gamma} = (\tilde{\gamma}_{kl})$ is a real, positive $2n \times 2n$ matrix, the covariance-, or correlation matrix. This is nothing but the CM of the observables $X_i$ and $P_i$. It has to obey[8]

$$\tilde{\gamma} \geq i \bigoplus_{i=1}^{n} J \quad \text{where} \quad \mathrm{J} = \left( \begin{array}{cc} 0 & -1 \\ 1 & 0 \end{array} \right), \tag{3.47}$$

to be a proper CM. This is nothing but the incorporation of the Heisenberg uncertainty relation as in Eq. (3.42). One can show that the displacement can be set to zero by local transformations, thus all nonlocal properties of a Gaussian state are encoded in the CM $\tilde{\gamma}$.

The important fact for our approach is that the separability problem for Gaussian states is essentially solved. Namely the following facts are known.

In Ref. [115] it was shown that a Gaussian state described by $\tilde{\gamma}$ is separable if and only if there exist two CMs $\tilde{\gamma}_A$ and $\tilde{\gamma}_B$ for Alice and Bob such that

$$\tilde{\gamma} \geq \tilde{\gamma}_A \oplus \tilde{\gamma}_B \tag{3.48}$$

[8] The symbol $A \oplus B$ denotes a block diagonal matrix with the matrices $A$ and $B$ on the diagonal and zero matrices elsewhere: $A \oplus B = \left( \begin{array}{cc} A & 0 \\ 0 & B \end{array} \right)$.

holds. However, these $\tilde{\gamma}_A$ and $\tilde{\gamma}_B$ are in general not easy to find. The separability problem was the finally solved in [116] where a map

$$\tilde{\gamma}_1 \mapsto \tilde{\gamma}_2 \tag{3.49}$$

of one CM onto another CM was defined, which preserves the separability properties, *i.e.*, $\tilde{\gamma}_1$ is separable if and only if $\tilde{\gamma}_2$ is separable, too. By iterating this map one gets a CM for which separability is easy to check.

Now we can derive separability criteria similar to these for finite-dimensional systems. Let us first formulate a criterion similar to (3.48):

**Proposition 3.16.** Let $\varrho$ be separable and let $A_i, i = 1, ..., n$ and $B_i, i = 1, ..., m$ be observables on Alice's, respectively, Bob's space. Define $M_i = A_i \otimes \mathbb{1}, i = 1, ..., n$ and $M_i = \mathbb{1} \otimes B_i, i = n+1, ..., n+m$. Then there are states $\varrho_k^A$ and $\varrho_k^B$ and convex weights $p_k$ such that if we define

$$\begin{aligned} \kappa_A &:= \sum_k p_k \gamma(\varrho_k^A, A_i), \\ \kappa_B &:= \sum_k p_k \gamma(\varrho_k^B, B_i), \end{aligned} \tag{3.50}$$

the inequality

$$\gamma(\varrho, M_i) \geq \kappa_A \oplus \kappa_B \tag{3.51}$$

holds.

*Proof.* This inequality follows from Lemma 3.13 and the fact that for these special choice of the $M_i$ for product states: $\gamma(\varrho^A \otimes \varrho^B, M_i) = \gamma(\varrho^A, A_i) \oplus \gamma(\varrho^B, B_i)$ holds. □

The algorithm for the solution of the separability problem of Gaussian states [116] relies heavily on a equation of the type (3.48,3.51). Thus it is not very surprising that we can now also perform one step of this algorithm.

**Proposition 3.17.** Let $\varrho$ be separable and let $M_i$ be defined as in the preceding Proposition. We write $\gamma(\varrho, M_i)$ in a block structure:

$$\gamma(\varrho, M_i) = \begin{pmatrix} A & C \\ C^T & B \end{pmatrix}, \tag{3.52}$$

where $A = \gamma(\varrho, A_i)$ ( $B = \gamma(\varrho, B_i)$) is an $n \times n$- ($m \times m$)-matrix. Then there exist $\kappa_A$ and $\kappa_B$ as in the previous proposition with[9]:

$$A - CB^{-1}C^T \geq \kappa_A \tag{3.53}$$

$$B - C^T A^{-1} C \geq \kappa_B. \tag{3.54}$$

To prove this fact requires a lemma about positive matrices, which is known from the literature:

[9] Here, and for the rest of this section $X^{-1}$ denotes the pseudo-inverse of the matrix $X$, *i.e.*, the inversion on the range.

**Lemma 3.18.** Let $\gamma$ be matrix with a block structure as in (3.52). Equivalent are:
(a) $\gamma$ is positive: $\gamma \geq 0$;
(b) We have $\ker(B) \subseteq \ker(C)$ and $A - CB^{-1}C^T \geq 0$;
(c) We have $\ker(A) \subseteq \ker(C^T)$ and $B - C^T A^{-1} C \geq 0$.

*Proof.* This is proved in [116]. Similar statements are known in the context of the so-called Schur complement, see, e.g. [81, Theorem 7.7.6]. □

*Proof of Proposition 3.17.* These inequalities arise from the first step of the algorithm of [116]. We arrived already after the first step at a block diagonal matrix and thus at a fixed point. We prove it as in [116] with the help of Lemma 3.18. Applying the equivalence (a)-(b) to Eqs. (3.51, 3.52) yields with Proposition 3.14 $A - C(B-\kappa_B)^{-1}C^T \geq \kappa_A \geq 0$. With the equivalence (b)-(c) we get $B - \kappa_B - C^T A^{-1} C \geq 0$, which proves Eq. (3.54). The proof of (3.53) is similar. □

In view of the previous sections we can state that any pure entangled state and any bound entangled UPB state violates also the covariance matrix criterion. Identifying more general classes of states which allow a detection leads to the difficult problem of characterizing the possible $\sum_k p_k \gamma(\varrho_k^A \otimes \varrho_k^B, M_i)$ in Eq. (3.39), resp. the $\kappa_{A/B}$ in Eqs. (3.53, 3.54). This is for finite-dimensional systems not as simple as for Gaussian states, since the condition (3.42) is weaker and more difficult to handle than the condition (3.47). We only know some properties of the $\kappa_{A/B}$ : Taking $A_i = B_i = \sigma_i, i = x, y, z$ for two qubits, we know that $\sum_i \delta^2(A_i) \geq 2$. thus $Tr(\kappa_A) \geq 2$ holds. Applying this to the Werner states $\varrho(p) = p|\psi\rangle\langle\psi| + (1-p)\mathbb{1}/4$ with $|\psi\rangle = (|01\rangle - |10\rangle)/\sqrt{2}$ one can calculate that they are detected via Eq. (3.53) for $p \geq 1/\sqrt{3}$.

Of course, we also cannot expect Propositions 3.16 and 3.17 to be sufficient for separability, since we could have chosen the operators $M_i$ in an inconvenient way. But a remarkable property of Eq. (3.53) is that the restriction imposed by $\kappa_A$ is independent of the choice of the $B_i$. So, it might be a good tool to investigate entanglement in an asymmetric situation, e.g. when Alice owns a qubit and Bob's space is a high-dimensional one.

## 3.8 Entropic uncertainty relations

In this section we want to recall some facts we will need for the investigation of entropic uncertainty relations and their relationship to entanglement. We start by noting some properties of different types of entropies.

For a general probability distribution $\mathcal{P} = (p_1, ..., p_n)$ there are several possibilities to define an entropy. We will focus on some entropies, which are used often in the literature. We will use the *Shannon entropy* [117]

$$S^S(\mathcal{P}) := -\sum_k p_k \ln(p_k), \tag{3.55}$$

and the so-called *Tsallis entropy* [118, 119]

$$S_q^T(\mathcal{P}) := \frac{1-\sum_k (p_k)^q}{q-1}, \quad q > 1. \tag{3.56}$$

Another entropy used in physics is the *Rényi entropy* [120], which is given by

$$S_q^R(\mathcal{P}) := \frac{\ln(\sum_k (p_k)^q)}{1-q}, \quad q > 1. \tag{3.57}$$

Let us state some of their properties.

**Proposition 3.19.** The entropies $S^S, S_q^T, S_q^R$ have the following properties:
(a) They are positive and they are zero if and only if the probability distribution is concentrated at one $j$, *i.e.*, $p_i = \delta_{ij}$.
(b) For $q \to 1$ the Tsallis and the Rényi entropy coincide with the Shannon entropy:

$$\lim_{q\to 1} S_q^R(\mathcal{P}) = \lim_{q\to 1} S_q^T(\mathcal{P}) = S^S(\mathcal{P}). \tag{3.58}$$

Thus we often write $S_1^T := S^S$.
(c) $S^S(\mathcal{P})$ and $S_q^T(\mathcal{P})$ are concave functions in $\mathcal{P}$, *i.e.*, they obey $S(\lambda\mathcal{P}_1 + (1-\lambda)\mathcal{P}_2) \geq \lambda S(\mathcal{P}_1) + (1-\lambda)S(\mathcal{P}_2)$. The Rényi entropy $S_q^R(\mathcal{P})$ is not concave. $S_q^R(\mathcal{P})$ and $S_q^T(\mathcal{P})$ both decrease monotonically in $q$. Further, $S_q^R(\mathcal{P})$ is a monotonous function of $S_q^T(\mathcal{P})$:

$$S_q^R(\mathcal{P}) = \frac{\ln(1+(1-q)S_q^T(\mathcal{P}))}{1-q}. \tag{3.59}$$

(d) In the limit $q \to \infty$ we have

$$\lim_{q\to\infty} S_q^R(\mathcal{P}) = -\ln \max_j(p_j). \tag{3.60}$$

*Proof.* Property (a) is easy to see. Property (b) is shown in [119, 120]. For a discussion of (c) see [121]. Eq. 3.59 is given in [119] and (d) is proved in [120]. □

Now we can introduce more general entropic functions and note some facts about their relationship to majorization. Let $\mathcal{P} = (p_1, ..., p_n)$ and $\mathcal{Q} = (q_1, ..., q_n)$ be two probability distributions. We can write them decreasingly ordered, *i.e.*, we have $p_1 \geq p_2 \geq ... \geq p_n$. We say that $\mathcal{P}$ *majorizes* $\mathcal{Q}$ or $\mathcal{Q}$ *is more mixed than* $\mathcal{P}$ and write it as

$$\mathcal{P} \succ \mathcal{Q} \quad \text{or} \quad \mathcal{Q} \prec \mathcal{P} \tag{3.61}$$

iff for all $k$

$$\sum_{i=1}^{k} p_k \geq \sum_{i=1}^{k} q_k \tag{3.62}$$

holds[10]. If the probability distributions have a different number of entries, we can append zeroes in this definition. We can characterize majorization completely, if we look at functions of a special type, namely functions $S(\mathcal{P})$ of the form

$$S(\mathcal{P}) = \sum_i s(p_i) \tag{3.63}$$

[10] Note that the sign "$\succ$" is sometimes defined the other way round in the literature.

where $s : [0;1] \to \mathbb{R}$ is a concave function. Such functions are by definition concave in $\mathcal{P}$ and obey several natural requirements for information measures [121–123]. We will call them entropic functions and reserve the notion $S(\mathcal{P})$ for such functions. Note that the Shannon and the Tsallis entropy are of the type (3.63), while the Rényi entropy is not. The following connection between entropic functions and majorization will become important in the following.

**Lemma 3.20.** We have $\mathcal{P} \succ \mathcal{Q}$ if and only if for all entropic functions $S(\mathcal{P}) \leq S(\mathcal{Q})$ holds.

*Proof.* See [121] and references therein. □

It is a natural question to ask for a *small* set of concave functions $\{s_j\}$ such that if $\sum_i s_j(p_i) \leq \sum_i s_j(q_i)$ holds for all $s_j$, this already implies $\mathcal{P} \succ \mathcal{Q}$. Here, we only point out that the set of all Tsallis entropies is not big enough for this task, but there is two parameter family of $\{s_j\}$ which is sufficient for this task [124]. We will discuss this in more detail later.

Now we turn to entropic uncertainty relations. Let us assume that we have a non-degenerate observable $M$ with a spectral decomposition $M = \sum_i \mu_i |m_i\rangle\langle m_i|$. A measurement of this observable in a quantum state $\varrho$ gives rise to a probability distribution of the different outcomes:

$$\mathcal{P}(M)_\varrho = (p_1, ..., p_n); \quad p_i = Tr(|m_i\rangle\langle m_i|\varrho). \tag{3.64}$$

Given this probability distribution, we can look at its entropy $S(\mathcal{P}(M))_\varrho$. We will often write for short $S(M) := S(\mathcal{P}(M))_\varrho$, when there is no risk of confusion.

If we have another observable $N = \sum_i \nu_i |n_i\rangle\langle n_i|$ we can define $\mathcal{P}(N)_\varrho$ in the same manner. Now, if $M$ and $N$ do not share a common eigenstate, it is clear that there must exist a strictly positive constant $C$ such that

$$S^S(M) + S^S(N) \geq C \tag{3.65}$$

holds. Estimating $C$ is not easy, after early works [107, 125, 126] on this problem, it was shown by Maassen and Uffink [108, 127] that one could take

$$C = -2\ln(\max_{i,j} |\langle m_i | n_j\rangle|). \tag{3.66}$$

There are generalizations of this bound to degenerate observables [128], more than two observables [129], or other entropies than the Shannon entropy [130]. Also one can sharpen this bound in many cases [131, 132].

A few remarks about the entropic uncertainty relations are in order at this point. First, a remarkable fact is that the bound in Eq. (3.65) does not depend on the state $\rho$. This is in contrast to the usual Heisenberg uncertainty relation for finite dimensional systems. Second, as already mentioned, the Maassen-Uffink bound (3.66) is not optimal in general. Third, it is very difficult to obtain an optimal bound even for simple cases. For instance, for the case of two qubits, the determination of the optimal bound for arbitrary observables relies on numerical calculations at some point [132].

Let us finally mention that there are other ways of associating an entropy with the measurement of an observable. Given an observable $M$ one may decompose it as

$$M = \sum_i \eta_i |e_i\rangle\langle e_i| \qquad (3.67)$$

where a weighted sum of the $|e_i\rangle\langle e_i|$ forms a partition of the unity:

$$\sum_i \lambda_i |e_i\rangle\langle e_i| = \mathbb{1}, \quad \lambda_i \geq 0. \qquad (3.68)$$

Here the $|e_i\rangle\langle e_i|$ are not necessarily orthogonal, *i.e.*, the decomposition (3.67) is not necessarily the spectral decomposition. The expression (3.68) corresponds to a POVM, and by performing this POVM one could measure the probabilities $q_i = Tr(\varrho \lambda_i |e_i\rangle\langle e_i|)$ and determine the expectation value of $M$. This gives rise to a probability distribution $\mathcal{Q} = (q_1, q_2, ...)$ and thus to an entropy for the measurement via

$$S(M, \vec{\eta}, \vec{\lambda})_\varrho = S(\mathcal{Q}). \qquad (3.69)$$

This construction of an entropy depends on the choice of the decompositions in Eqs. (3.67, 3.68) which makes it more difficult to handle. Thus we will mostly consider the entropy defined by the spectral decomposition as in Eq. (3.64) in the following.

## 3.9 Entropies and entanglement I

What are the relations between entropic functions and entanglement? First, note that the majorization criterion introduced in Section 1.1.3 in nothing but a majorization relation as defined in Eqs. (3.61, 3.62). Thus, Lemma 3.20 can be applied, yielding inequalities for the entropies of the state and the reduced state. This approach is not new, is has been considered before in [133–136][11]. However, an experimental implementation of the resulting criteria requires always state reconstruction and then the computation of the eigenvalues of the total and the reduced state.

Another possibility lies in the entropy of a measurement $S(\mathcal{P}(M))_\varrho$. One may ask, whether this quantity can be used to say something about the entanglement properties of a state. $S(\mathcal{P}(M))_\varrho$ is an experimentally direct accessible quantity, thus no state reconstruction is needed.

The first work which raised the question whether the entropy of a measurement and entanglement are somehow connected was to our knowledge done in Ref. [137]. Recently, in Ref. [138], some separability criteria for bipartite systems in terms of entropic uncertainty relations were derived. However, their derivation relied at a crucial point on a numerical optimization.

The scheme we want to use for the detection of entanglement is similar to the idea for the variances: We take one or several observables $M_i$ and look at the sum

[11] Note, however, that the papers [133–135] were written before the majorization criterion was established.

of the entropies $\sum_i S(M_i)_\varrho$. For product states we derive lower bounds for this sum, which by concavity also hold for separable states. Violation of this bound for a state $\varrho$ thus implies that $\varrho$ is entangled. The difficulty of this scheme lies in the determination of the lower bound[12].

We will present two methods for obtaining such a bound here. In this section we will derive the first method based on the calculation of overlaps. Then we will investigate the resulting criteria for two and three qubits. In the next section we will derive and investigate the second method. There we will show how any entropic uncertainty relation can be translated into a necessary separability criterion.

The first of our methods applies if we look only at one $M$. If the set of the eigenvectors of $M$ does not contain any product vector, it is clear that there must be a $C > 0$ such that $S_q^T(\mathcal{P}(M)) \geq C$ holds for all separable states. From the Schmidt coefficients of the eigenvectors of $M$ we can determine $C$.

**Theorem 3.21.** Let $M = \sum \mu_i |m_i\rangle\langle m_i|$ be a non degenerate observable. Let $c < 1$ be an upper bound for all the squared Schmidt coefficients of all $|m_i\rangle$. Then

$$S_q^T(M) \geq \frac{1 - \lfloor 1/c \rfloor c^q - (1 - \lfloor 1/c \rfloor c)^q}{q-1} \tag{3.70}$$

holds for all separable states.

*Proof.* The maximal Schmidt coefficient of an entangled state is just the maximal overlap between this state an the product states, as proved in Section 2.6. Now, let $\mathcal{P}(M)_\varrho = (p_1, p_2, ...p_n)$. Then all the probabilities $p_i$ appearing there are bounded by $c$, if $\varrho$ is a projector onto a product vector. Furthermore, if there are two $p_i$, say $p_1$ and $p_2$, with $0 < p_1 \leq p_2 < c$ we can still decrease $S_q^T$, by increasing $p_2$ and decreasing $p_1$, since then

$$S_q^T(p_1 - \varepsilon, p_2 + \varepsilon, p_3, ..., p_n) < S_q^T(p_1, p_2, p_3, ..., p_n). \tag{3.71}$$

Thus, $S_q^T$ is minimized, when $\mathcal{P}(M)_\varrho$ is as peaked as possible, *i.e.*, $\lfloor 1/c \rfloor$ of the $p_i$ satisfy the bound $p_i = c$, while one other $p_i$ is as big as possible. This proves (3.70). Note that for $q = 2$ this is equivalent to Proposition 3.11. Note also that this Theorem can easily be extended to arbitrary entropic functions. □

Of course, for this approach due to Eq. (3.59) the Tsallis and the Rényi entropy are equivalent. The Rényi entropy will later be used to discuss the limit $q \to \infty$. Note also that a similar statement for the entropy defined via the corresponding POVM as in Eq. (3.69) can be derived, provided that a bound on the probabilities for the outcomes of the POVM is known.

To investigate Theorem 3.21, let us start with two qubits. Assume that we have a non-degenerate observable, which is Bell diagonal

$$M := \sum_i \mu_i |BS_i\rangle\langle BS_i| \tag{3.72}$$

[12] This scheme was also applied in [138], however, the determination of the lower bound was performed only numerically, by a minimization over all pure product states.

with $|BS_1\rangle = (|00\rangle + |11\rangle)/\sqrt{2}, |BS_2\rangle = (|00\rangle - |11\rangle)/\sqrt{2}, |BS_3\rangle = (|01\rangle + |10\rangle)/\sqrt{2}, |BS_4\rangle = (|01\rangle - |10\rangle)/\sqrt{2}$. Since the maximal squared overlap between the Bell states and and the separable states equals $1/2$, we can state:

**Corollary 3.22.** If $\varrho$ is separable, then for every $q > 1$

$$S_q^T(M)_\varrho \geq \frac{1 - 2^{1-q}}{q-1} \tag{3.73}$$

holds.

For the Rényi entropy the bound reads $S_q^R(M)_\varrho \geq \ln(2)$, thus this criterion becomes stronger when $q$ increases.

To investigate the power of this criterion, first note that Eq. (3.73) is for the case $q = 2$ equivalent to the variance based criterion $\sum_i \delta^2(|BS_i\rangle\langle BS_i|) \geq 1/2$ in Proposition 3.11. For other values of $q$ we can, as in Section 3.6 determine the expectation values of the $|BS_i\rangle\langle BS_i|$ by measuring three combinations of Pauli matrices. Again, we define $i = Tr(\varrho\sigma_i \otimes \sigma_i)$ for $i = x, y, z$. Then we arrive as before at a geometrical picture, see Figure 3.3.(a).

One can depict the border of the states which are not detected (for different $q$) in this three dimensional space. This has also been done in Figure 3.3. One can directly observe, that in the limit $q \to \infty$ the Corollary 3.22 enables one to detect all states, which are outside the octahedron. This is not by chance and can also be proven analytically: In the limit $q \to \infty$ Corollary 3.22 requires

$$\max_i \{p_i \in \mathcal{P}(M)_\varrho\} \leq \frac{1}{2} \tag{3.74}$$

from a state to escape the detection. This condition is equivalent, to a set of four witnesses: The observables

$$\mathcal{W}_i = \frac{\mathbb{1}}{2} - |BS_i\rangle\langle BS_i| \tag{3.75}$$

are all optimal witnesses, imposing the same condition on $\varrho$.

The results of Theorem 3.21 can also easily be applied to multipartite systems, e.g. three qubits.

**Corollary 3.23.** Let $M = \sum_i \mu_i |\psi_i\rangle\langle\psi_i|$ be an observable which is GHZ-diagonal, *i.e.* the $|\psi_i\rangle$ are of the form $|\psi_{1/5}\rangle = (|000\rangle \pm |111\rangle)/\sqrt{2}$; $|\psi_{2/6}\rangle = (|100\rangle \pm |011\rangle)/\sqrt{2}$; $|\psi_{3/7}\rangle = (|010\rangle \pm |101\rangle)/\sqrt{2}$; $|\psi_{4/8}\rangle = (|001\rangle \pm |110\rangle)/\sqrt{2}$. Then for all biseparable states

$$S_q^T(M)_\varrho \geq \frac{1 - 2^{1-q}}{q-1} \tag{3.76}$$

holds. For states belonging to the W class the entropy is bounded by $S_q^T(M)_\varrho \geq (1 - (3/4)^q + (1/4)^q)/(q-1)$.

*Proof.* Due to the concavity of the entropy we have to show the bound only for pure biseparable states. Then the proof follows directly from the fact that the

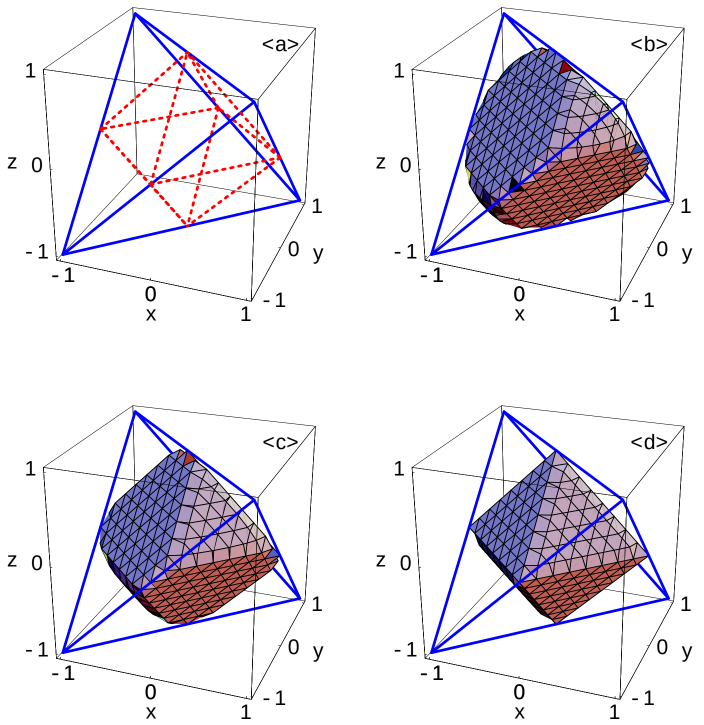

Figure 3.3. *Investigation of the criterion from Eq. (3.73) for different values of $q$: (a): The tetrahedron (blue, solid lines) of all states and the octahedron (red, dashed lines) which contains the separable states. (b): The subset of states which are not detected by Eq. (3.73) for $q = 2$. (c): As (b) but for $q = 4$. (d): As (b) but for $q = 15$.*

maximal overlap between the states $|\psi_i\rangle$ and the biseparable (resp. W class) states is $1/2$ and $3/4$, respectively (see Section 2.4.2). □

Again, as in the two qubit case, for $q = 2$ the criterion is equivalent to a criterion in terms of variances (see Proposition 3.10). Also one can show that this criterion becomes stronger, when $q$ increases, and in the limit $q \to \infty$ it is equivalent to a set of eight witnesses of the type $\mathcal{W}_i = 1/2 \cdot \mathbb{1} - |\psi_i\rangle\langle\psi_i|$ (respectively, $\mathcal{W}_i = 3/4 \cdot \mathbb{1} - |\psi_i\rangle\langle\psi_i|$).

## 3.10 Entropies and entanglement II

The second method for deriving lower bounds of the entropy for separable states, deals with product observables, which might be degenerate. If an observable $M$ is degenerate, the definition of $\mathcal{P}(M)$ is not unique, since the spectral decomposition is not unique. By combining eigenvectors with the same eigenvalue one arrives, however, at a unique decomposition of the form

$$M = \sum_i \eta_i X_i \tag{3.77}$$

with $\eta_i \neq \eta_j$ for $i \neq j$ and the $X_i$ are orthogonal projectors of maximal rank. Thus we can define for degenerate observables $\mathcal{P}(M)_\varrho$ by $p_i = Tr(\varrho X_i)$. To proceed, we need the following Lemma.

**Lemma 3.24.** Let $\varrho = \varrho_A \otimes \varrho_B$ be a product state on a bipartite Hilbert space $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B$. Let $A$ (resp. $B$) be observables with nonzero eigenvalues on $\mathcal{H}_A$ (resp. $\mathcal{H}_B$). Then

$$\mathcal{P}(A \otimes B)_\varrho \prec \mathcal{P}(A)_{\varrho A} \tag{3.78}$$

holds. Also $\mathcal{P}(A \otimes B)_\varrho \prec \mathcal{P}(B)_{\varrho B}$ is valid.

*Proof.* To prove the bound we use the fact that for two probability distributions $\mathcal{P} = \vec{p}$ and $\mathcal{Q} = \vec{q}$ we have $\mathcal{P} \succ \mathcal{Q}$ if and only if there is a doubly stochastic matrix $D$ (*i.e.* a matrix where all column and row sums equal one) such that $\vec{q} = D\vec{p}$ holds [139]. We will construct this matrix D.

Define $\mathcal{P} = \mathcal{P}(A)_{\varrho A} = \{p_i\}$ and $\mathcal{Q} = \mathcal{P}(B)_{\varrho B} = \{q_j\}$. Without loosing generality we can assume that $A$ and $B$ are non-degenerate and have both $n$ different outcomes. We only have to distinguish the cases where $A \otimes B$ is degenerate or non-degenerate.

If $A \otimes B$ is non-degenerate we have $\mathcal{R} = \vec{r} := \mathcal{P}(A \otimes B)_\varrho = \{p_i q_j\}$. Let us look at the $n^2 \times n^2$-matrix

$$\Lambda_0 = (\lambda_{ij}); \quad \lambda_{ij} = \mathbb{1}_n q_{((i+j-1) \bmod n)}. \tag{3.79}$$

$\Lambda_0$ is an $n \times n$ block matrix, the blocks are itself $n \times n$ matrices, the $\lambda_{ij}$. It is now clear, that

$$\vec{r} = \Lambda_0 \vec{p} \tag{3.80}$$

and $\Lambda_0$ is also doubly stochastic. To give an example, for two qubits Eq. (3.80) may look like

$$\begin{pmatrix} p_1q_1 \\ p_2q_1 \\ p_1q_2 \\ p_2q_2 \end{pmatrix} = \begin{pmatrix} q_1 & 0 & q_2 & 0 \\ 0 & q_1 & 0 & q_2 \\ q_2 & 0 & q_1 & 0 \\ 0 & q_2 & 0 & q_1 \end{pmatrix} \cdot \begin{pmatrix} p_1 \\ p_2 \\ 0 \\ 0 \end{pmatrix}. \tag{3.81}$$

Thus the claim is proved for the case that $A \otimes B$ is non-degenerate.

If $A \otimes B$ is degenerate, some of the $q_i p_j$ are grouped together since they belong to the same eigenvalue. This grouping can be achieved by successive contracting two probabilities:

$$\{p_i q_j \ , \ p_l q_m\} \rightarrow p_i q_j + p_l q_m. \tag{3.82}$$

Since $A$ and $B$ have nonzero eigenvalues we have here $i \neq l$ and $j \neq m$. We can now construct a new matrix $\Lambda$ from $\Lambda_0$ which is generates this contraction: Set

$$(\lambda_{11})_{il} = q_m; \;\; (\lambda_{m1})_{ll} = 0; \;\; (\lambda_{1m})_{ii} = 0; \;\; (\lambda_{mm})_{li} = q_m. \tag{3.83}$$

This corresponds to shifting the entry $q_m$ in the first block column up $\Lambda$ from block $\lambda_{m1}$ to $\lambda_{11}$ to obtain $p_i q_j + p_l q_m$. Then in the $m$-th block column of $\Lambda$ this index is shifted downwards to keep the resulting matrix doubly stochastic. Again, for two qubits this may look like

$$\begin{pmatrix} p_1 q_1 \\ p_2 q_1 + p_1 q_2 \\ 0 \\ p_2 q_2 \end{pmatrix} = \begin{pmatrix} q_1 & 0 & q_2 & 0 \\ q_2 & q_1 & 0 & 0 \\ 0 & 0 & q_1 & q_2 \\ 0 & q_2 & 0 & q_1 \end{pmatrix} \cdot \begin{pmatrix} p_1 \\ p_2 \\ 0 \\ 0 \end{pmatrix}. \tag{3.84}$$

By iterating this procedure one can generate any contraction, which is compatible with the fact that $A$ and $B$ have nonzero eigenvalues. The resulting $\Lambda$ is clearly doubly stochastic. □

With the help of this lemma we can derive separability criteria from entropic uncertainty relations:

**Theorem 3.25.** Let $A_1, A_2, B_1, B_2$ be observables with nonzero eigenvalues on Alice's resp. Bob's space obeying an entropic uncertainty relation of the type

$$S(A_1) + S(A_2) \geq C \tag{3.85}$$

or the same bound for $B_1, B_2$. If $\varrho$ is separable, then

$$S(A_1 \otimes B_1)_\varrho + S(A_2 \otimes B_2)_\varrho \geq C \tag{3.86}$$

holds.

*Proof.* We can write $\varrho = \sum_k \alpha_k \varrho_k^A \otimes \varrho_k^B$ as a convex combination of product states and with the help of Lemmata 3.20 and 3.24 we have: $S(A_1 \otimes B_1)_\varrho + S(A_2 \otimes B_2)_\varrho \geq \sum_k \alpha_k [S(A_1 \otimes B_1)_{\varrho_k^A \otimes \varrho_k^B} + S(A_2 \otimes B_2)_{\varrho_k^A \otimes \varrho_k^B}] \geq \sum_k \alpha_k [S(A_1)_{\varrho^A} + S(A_2)_{\varrho^A}] \geq C$. This proves the claim. Of course, the same result holds, if we look at three or more $A_i$. □

For entangled states this bound can be violated, since $A_1 \otimes B_1$ and $A_2 \otimes B_2$ might be degenerate and have a common (entangled) eigenstate. Note that the precondition on the observables to have nonzero eigenvalues is more a technical condition. It is needed to set some restriction on the degree of degeneracy of the combined observables. Given an entropic uncertainty relation, this requirement can always be achieved simply by altering the eigenvalues, since the entropic uncertainty relation does not depend on them.

This corollary shows how any entropic uncertainty relation can be transformed into a necessary separability criterion. On the other hand, if one is interested in numerical calculations, one can calculate bounds on $S(A_1 \otimes B_1)_\varrho + S(A_2 \otimes B_2)$ for separable states easily, since one only has to minimize the entropy for one party of the system.

To investigate the consequences of Theorem 3.25 for two qubits, we focus on the case that the observables for Alice and Bob are spin measurements in the $x$,$y$, or $z$-direction. First note that due to the Maassen-Uffink relation

$$S_1^T(\sigma_x)_\varrho + S_1^T(\sigma_y)_\varrho \geq \ln(2) \tag{3.87}$$

holds. This implies that for all separable states

$$S_1^T(\sigma_x \otimes \sigma_x)_\varrho + S_1^T(\sigma_y \otimes \sigma_y)_\varrho \geq \ln(2) \tag{3.88}$$

has to hold, too. This is just the bound which was numerically confirmed in Ref. [138]. In the same reference, also the separability criterion $S_1^T(\sigma_x \otimes \sigma_x)_\varrho + S_1^T(\sigma_y \otimes \sigma_y)_\varrho + S_1^T(\sigma_z \otimes \sigma_z)_\varrho \geq 2\ln(2)$ for has been asserted. In view of Theorem 3.25 this follows from the entropic uncertainty relation $S_1^T(\sigma_x)_\varrho + S_1^T(\sigma_y)_\varrho + S_1^T(\sigma_z)_\varrho \geq 2\ln(2)$, proved in [129].

It is now interesting to take the Tsallis entropy and vary the parameter $q$. We do this mainly numerically. We first compute entropic uncertainty relations by minimizing over all pure single-qubit states

$$\begin{aligned} S_{xy}(q) &= \min_\varrho (S_q^T(\sigma_x)_\varrho + S_1^T(\sigma_y)_\varrho) \\ S_{xyz}(q) &= \min_\varrho (S_q^T(\sigma_x)_\varrho + S_q^T(\sigma_y)_\varrho + S_q^T(\sigma_z)_\varrho). \end{aligned} \tag{3.89}$$

The results are shown in Figure 3.4. Some of the numerical results can, however, also be proved analytically:

**Remark 3.26.** For $q \in [2n-1, 2n], n \in \mathbb{N}$ we have:

$$S_{xy}(q) = \frac{1-2^{1-q}}{q-1}. \tag{3.90}$$

Numerically, this seems to be valid also for other values of $q$, except $q \in (2;3)$.

*Proof.* It is clear that the minimum of $X = S_q^T(\sigma_x) + S_q^T(\sigma_y)$ is obtained for a state in the $x$-$y$ plane. Calculating $X$ for a pure state in this plane one recognizes that minimizing $X$ is equivalent to maximizing $Y = \cos^{2q}(\alpha) + \cos^{2q}(\alpha + \pi/4) + \cos^{2q}(\alpha + \pi/2)^{2q} + \cos^{2q}(\alpha + 3\pi/4)$. Using now the formula

$$\begin{aligned} \cos^s(x) &= \frac{1}{2^{s-1}} \frac{\Gamma(s+1)}{(\Gamma(s/2+1))^2} \left\{ \frac{1}{2} + \frac{s}{s+2}\cos(2x) + \right. \\ &\quad \left. + \frac{s(s-2)}{(s+2)(s+4)}\cos(4x) + ... \right\} \end{aligned} \tag{3.91}$$

(see Ref. [140, p. 263]) one can rewrite $Y$ as a series of the form $Y = C\sum_k a_k \cos(8k\alpha)$. Then it is easy to show that if $q \in [2n-1, 2n], n \in \mathbb{N}$ for all $k$ the coefficients $a_k$ are positive, thus $Y$ is maximized if $\alpha = 0$. This proves that $X$ is minimized for an eigenstate of $\sigma_x$, which yields (3.90). □

Having derived these entropic uncertainty relations we can look at the corresponding separability criteria:

$$S_q^T(\sigma_x \otimes \sigma_x) + S_q^T(\sigma_y \otimes \sigma_y) \geq S_{xy}(q), \tag{3.92}$$

$$S_q^T(\sigma_x \otimes \sigma_x) + S_q^T(\sigma_y \otimes \sigma_y) + S_q^T(\sigma_z \otimes \sigma_z) \geq S_{xyz}(q). \tag{3.93}$$

To investigate the power of this criteria, let us look at Werner states $\rho(p) = p|\psi^-\rangle\langle\psi^-| + (1-p)\mathbb{1}/4$. We can make the following estimation: There are single qubit states with $\mathcal{P}(\sigma_x) = \mathcal{P}(\sigma_y) = \{(2-\sqrt{2})/4, (2+\sqrt{2})/4\}$. The lower bound

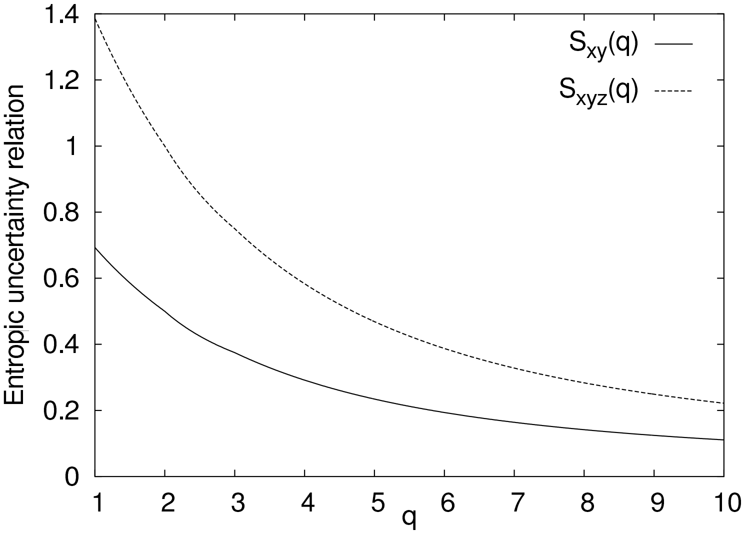

Figure 3.4. *Numerical lower bounds in Eq. (3.89) depending on q.*

$S_{xy}(q)$ must therefore obey $S_{xy}(q) \leq 2S_q^T(\{(2-\sqrt{2})/4, (2+\sqrt{2})/4\})$. For the Werner states we have $\mathcal{P}(\sigma_x \otimes \sigma_x) = \mathcal{P}(\sigma_y \otimes \sigma_y) = \{(1+p)/2, (1-p)/2\}$. From this one can easily calculate that Eq. (3.92) cannot detect them for $p \leq 1/\sqrt{2} \approx 0.707$. A similar argumentation shows that Eq. (3.93) has to fail for $p \leq 1/\sqrt{3} \approx 0.577$. The numerical results are shown in Figure 3.5. They show that indeed the Tsallis entropy for $q \in [2; 3]$ can reach this bound.

Here, it is important to note that Werner states are already entangled for $p > 1/3$. The criteria from Eqs. (3.92, 3.93) therefore fail to detect all Werner states, while the criterion from Eq. (3.74) is strong enough to detect all of them.

As already mentioned, the Tsallis entropy is not the only entropic function. A more general function is of the type:

$$S_{a,t}^{RC}(\mathcal{P}) := \sum_i f(p_i), \tag{3.94}$$

where $a \in [0; 1]$, and $t \in [0; \infty)$, and $f$ is defined as

$$f(x) := g_t(x-a) - (1-x)g_t(-a) - xg_t(1-a), \tag{3.95}$$

where

$$g_t(y) := -\frac{\ln(\cosh(ty))}{2t}. \tag{3.96}$$

One can show that $\mathcal{P} \succ \mathcal{Q}$ iff $S_{a,t}^{RC}(\mathcal{Q}) \leq S_{a,t}^{RC}(\mathcal{P})$ for all $a$ and $t$ [124]. This is a property which does not hold for the Tsallis entropy. But this does not mean that criteria based on $S_{a,t}^{RC}$ are stronger than criteria based on the $S_q^T$. With the use of the entropy $S_{a,t}^{RC}$ one can, of course, better use the property of Lemma 3.24. But since for the proof of Theorem 3.25 also the concavity of the entropy was used, one might loose this advantage there. In fact, by numerical calculations one can easily

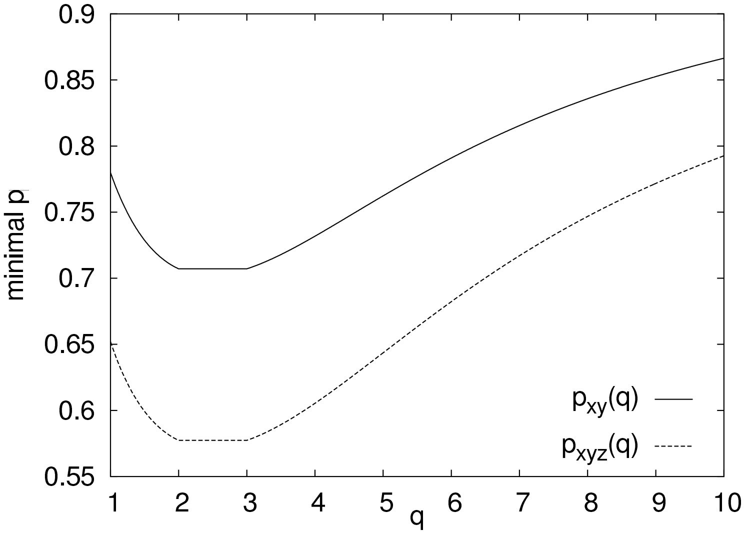

Figure 3.5. *Values of $p_{min}$ depending on $q$ such that for $p > p_{min}$ Werner states of the form $\rho(p) = p|\psi^-\rangle\langle\psi^-| + (1-p)\mathbb{1}/4$. are detected via Eqs. (3.92, 3.93). The curve $p_{xy}$ refers to the separability criterion in Eq. (3.92) and $p_{xyz}$ to Eq. (3.93). Note that Werner states are entangled for $p > 1/3$.*

show that for $a = 1/2$ and $t$ large the criterion using $S_{a,t}^{RC}$ and the measurements $\sigma_x \otimes \sigma_x$ and $\sigma_y \otimes \sigma_y$ (resp. $\sigma_x \otimes \sigma_x, \sigma_y \otimes \sigma_y$ and $\sigma_z \otimes \sigma_z$) reaches, as the Tsallis entropy, the best possible value $p = 1/\sqrt{2}$. (resp. $p = 1/\sqrt{3}$).

In order to show that Theorem 3.25 can be also applied for the detection of multipartite entanglement, we give an example which allows to detect the three-qubit GHZ state.

**Corollary 3.27.** Let $\varrho$ be a biseparable three-qubit state. Then for the Shannon entropy as well as for the Tsallis entropy for $q \in \{2, 3, 4, ...\}$ the following bounds hold:

$$S_1^T(\sigma_x \otimes \sigma_x \otimes \sigma_x)_\varrho + S_1^T(\sigma_z \otimes \sigma_z \otimes \mathbb{1})_\varrho + S_1^T(\mathbb{1} \otimes \sigma_z \otimes \sigma_z)_\varrho \geq \ln(2) \tag{3.97}$$

$$S_q^T(\sigma_x \otimes \sigma_x \otimes \sigma_x)_\varrho + S_q^T(\sigma_z \otimes \sigma_z \otimes \mathbb{1})_\varrho + S_q^T(\mathbb{1} \otimes \sigma_z \otimes \sigma_z)_\varrho \geq \frac{1-2^{1-q}}{q-1} \tag{3.98}$$

For the GHZ state $|GHZ\rangle = (|000\rangle + |111\rangle)/\sqrt{2}$ the left hand side of Eqs. (3.97, 3.98) is zero.

*Proof.* Again, we only have to prove the bound for pure biseparable states. If a state is A-BC biseparable, the bounds in Eq. (3.97) follows directly from Theorem 3.25 and the Maassen Uffink uncertainty relation, which guarantees that for the first qubit $S_1^T(\sigma_x) + S_1^T(\sigma_z) + S_1^T(\mathbb{1}) \geq \ln(2)$ holds. Eq. (3.98) follows similarly, using the fact that $S_q^T(\sigma_x) + S_q^T(\sigma_z) \geq (1-2^{1-q})/(q-1)$, see Remark 3.26. The proof for the other bipartite splittings is similar. □

Note that the observables used in Corollary 3.27 are so-called stabilizers of the GHZ state, which we already met in Section 2.7. The stabilizer formalism may again be useful to derive criteria in terms of entropic uncertainty relations.

Let us finally investigate, how robust against noise these criteria are. One can easily calculate that a state of the type $\varrho(p) = p|GHZ\rangle\langle GHZ| + (1-p)\mathbb{1}/8$ can be detected by Eq. (3.97) if $p \geq 0.877$. Eq. (3.98) seems to detect the most states for $q \in \{2, 3\}$, then they detect $\varrho(p)$ for $p \geq \sqrt{2/3} \approx 0.816$.

## 3.11 Conclusion

In conclusion, we have shown that uncertainty relations provide powerful tools for the detection of entanglement. This holds for the variance based formulations as well as for the entropic uncertainty relations. Our search for separability criteria in terms of uncertainty relations has, moreover, exhibited interesting facts about entanglement and the connection between uncertainty relations and entanglement on a fundamental level. One of these fact is that separability criteria for infinite-dimensional systems can be transferred to finite-dimensional systems. Another one is that any entropic uncertainty relation one one part of the system gives rise to a separability criterion on the composite system.

This leaves us with several questions which might be addressed further. One interesting question is whether there are entangled states, which do not violate the variance criterion. Finding an example for such states would surely lead to a better understanding of this criterion. Also the transfer of separability criteria from Gaussian states to finite-dimensional systems should be studied further. Especially a characterization of the states which are detected by the translation of the algorithm presented in Proposition 3.17 is of great interest.

Also for the entropic uncertainty relations many unsolved problems remain. One interesting question is, which entropies are best suited for special detection problems. We have seen that in some of our examples the Tsallis entropies with $q \in [2; 3]$ seemed to be the best. Clarifying the physical meaning of the parameter $q$ might help to understand this property.

Another important task is to find good (*i.e.* sharp) entropic uncertainty relations, especially for more than two observables. One the one hand, this is an interesting field of study for itself, on the other hand, this might help to explore the full power of the methods presented here.

Finally, one might speculate that there are connections between violation of entropic criteria and violation of local realism. It is remarkable that for the formulation of Bell inequalities as well as for the entropic uncertainty relations always non-commuting observables are required. Also, a violation of Theorem 3.25 can be viewed in the following way: Due to the quantum correlations the sum of the entropies of the total measurement is smaller than the sum of the entropies when only one party measures. This might be a violation of the locality condition needed in the derivation of Bell inequalities. There is, however, still a lot of work to do to justify these speculations.

# Bibliography

[1] A. Einstein, N. Podolski, and N. Rosen, Phys. Rev. **47**, 777 (1935). German translation in [141].

[2] E. Schrödinger, Naturwissenschaften **23**, 807 (1935); **23**, 823 (1935); **23**, 844 (1935). Reprinted in [141].

[3] R. Feynman, Int. J. Theor. Phys. **21**, 467 (1982).

[4] D. Deutsch, Proc. Roy. Soc. A **400**, 96 (1985).

[5] P. Shor, SIAM J. Comp. **26** 1484 (1997), [quant-ph/9508027].

[6] L. Grover, Phys. Rev. Lett. **79**, 325 (1997), [quant-ph/9706033]..

[7] C.H. Bennett and G. Brassard, in *Proc. IEEE Int. Conference on Computers, Systems, and Signal Processing*, (IEEE, New York, 1984).

[8] A.K. Ekert, Phys. Rev. Lett. **67**, 661 (1991).

[9] C.H. Bennett, G. Brassard, C. Crépeau, R. Jozsa, A. Peres and W.K. Wootters, Phys. Rev. Lett. **70**, 1895 (1993).

[10] A. Aspect, P. Grangier, and G. Roger, Phys. Rev. Lett. **47**, 460 (1981).

[11] J.-W. Pan, D. Bouwmeester, M. Daniell, H. Weinfurter, and A. Zeilinger, Nature **403**, 151 (2000).

[12] L.M.K. Vandersypen, M. Steffen, G. Breyta, C.S. Yannoni, M.H. Sherwood and I.L. Chuang, Nature **414**, 883 (2001).

[13] 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).

[14] A. Peres, *Quantum Theory: Concepts and Methods*, (Kluwer Academic Publishers, Dordrecht, 1995).

[15] D. Bouwmeester, A. Ekert, and A. Zeilinger (Eds.), *The Physics of Quantum Information,* (Springer, Berlin, 2000).

[16] M. Nielsen and I. Chuang, *Quantum Computation and Quantum Information,* (Cambrigde University Press, Cambridge, 2001).

[17] G. Alber, R. Beth, M. Horodecki, P. Horodecki, R. Horodecki, M. Rötteler, H. Weinfurter, R. Werner, and A. Zeilinger, *Quantum Information,* (Springer, Berlin, 2001).

[18] R.F. Werner, Phys. Rev. A **40**, 4277 (1989).

[19] A. Peres, Phys. Rev. Lett. **77**, 1413 (1996), [quant-ph/9604005].

[20] M. Horodecki, P. Horodecki, and R. Horodecki, Phys. Lett. A **223**, 1 (1996), [quant-ph/9605038].

[21] P. Horodecki, Phys. Lett. A **232**, 333 (1997), [quant-ph/9703004].

[22] M.A. Nielsen and J. Kempe, Phys. Rev. Lett. **86**, 5184 (2001), [quant-ph/0011117].

[23] M. Horodecki and P. Horodecki, Phys. Rev. A **59**, 4206 (1999), [quant-ph/9708015].

[24] O. Rudolph, Phys. Rev. A **67**, 032312 (2003), [quant-ph/0202121].

[25] K. Chen and L. Wu, Quant. Inf. Comp. **3**, 193 (2003), [quant-ph/0205017].

[26] A.C. Doherty, P.A. Parrilo, and F.M. Spedalieri, Phys. Rev. Lett. **88**, 187904 (2002), [quant-ph/0112007].

[27] A.C. Doherty, P.A. Parrilo, and F.M. Spedalieri, Phys. Rev. A **69**, 022308 (2004), [quant-ph/0308032].

[28] P. Horodecki and A. Ekert, Phys. Rev. Lett. **89**, 127902 (2002), [quant-ph/0111064].

[29] B.M. Terhal, Phys. Lett. A **271**, 319 (2000), [quant-ph/9911057].

[30] F. Hirzebruch and W. Scharlau, *Einführung in die Funktionalanalysis*, (Bibliographisches Institut, Mannheim, 1971).

[31] M. Lewenstein, B. Kraus, J.I. Cirac, and P. Horodecki, Phys. Rev. A **62**, 052310 (2000), [quant-ph/0005014].

[32] C.H. Bennett, G. Brassard, S. Popescu, B. Schumacher, J.A. Smolin, and W.K. Wootters, Phys. Rev. Lett. **76**, 722 (1996), [quant-ph/9511027].

[33] N. Gisin, Phys. Lett. A **210**, 151 (1996).

[34] M. Horodecki, P. Horodecki, and R. Horodecki, Phys. Rev. Lett. **80**, 5239 (1998), [quant-ph/9801069].

[35] T. Hiroshima, Phys. Rev. Lett. **91**, 057902 (2003), [quant-ph/0303057].

[36] W. Dür, J.I. Cirac, M. Lewenstein, and D. Bruß, Phys. Rev. A **61**, 062313 (2000), [quant-ph/9910022].

[37] D.P. DiVincenzo, P.W. Shor, J.A. Smolin, B.M. Terhal, and A.V. Thapliyal, Phys. Rev. A **61**, 062312 (2000), [quant-ph/9910026].

[38] C.H. Bennett, D.P. DiVincenzo, T. Mor, P.W. Shor, J.A. Smolin, and B.M. Terhal, Phys. Rev. Lett. **82**, 5385 (1999), [quant-ph/9808030].

[39] D. Bruß and A. Peres, Phys. Rev. A **61**, 30301 (2000), [quant-ph/9911056].

[40] K. Horodecki, M. Horodecki, P. Horodecki, and J. Oppenheim, quant-ph/0309110.

[41] W. Dür, G. Vidal, and J.I. Cirac, Phys. Rev. A **62**, 062314 (2000), [quant-ph/0005115].

[42] F. Verstraete, J. Dehaene, and B. De Moor, Phys. Rev. A **68**, 012103 (2003), [quant-ph/0105090].

[43] A. Acín, A. Andrianov, L. Costa, E. Jane, J.I. Latorre, and R. Tarrach, Phys. Rev. Lett. **85**, 1560 (2000), [quant-ph/0003050].

[44] V. Scarani, and N. Gisin, J. Phys. A **34**, 6043 (2001), [quant-ph/0103068].

[45] W. Dür, J.I. Cirac, and R. Tarrach, Phys. Rev. Lett. **83** 3562 (1999), [quant-ph/9903018].

[46] A. Acín, D. Bruß, M. Lewenstein, and A. Sanpera, Phys. Rev. Lett. **87**, 040401 (2001), [quant-ph/0103025].

[47] T. Eggeling and R.F. Werner, Phys. Rev. A **63**, 042111 (2001), [quant-ph/0010096].

[48] T.S. Cubitt, F. Verstraete, W. Dür, and J.I. Cirac, Phys. Rev. Lett. **91**, 037902 (2003), [quant-ph/0302168].

[49] W. Dür and J.I. Cirac, Phys. Rev. A **61**, 042314 (2000), [quant-ph/9911044].

[50] W. Dür and J.I. Cirac, J. Phys. A **34**, 6837 (2001), [quant-ph/0011025].

[51] A. Acín, E. Jane, W. Dür, and G. Vidal, Phys. Rev. Lett. **85**, 4811 (2000), [quant-ph/0007042].

[52] F. Verstraete, J. Dehaene, B. De Moor, and H. Verschelde, Phys. Rev. A, **65**, 052112 (2002), [quant-ph/0109033].

[53] J. Eisert and H.-J. Briegel, Phys. Rev. A **64**, 022306 (2001), [quant-ph/0007081].

[54] A. Peres, Found. Phys. **29**, 589 (1999), [quant-ph/9807017].

[55] R.F. Werner and M.M. Wolf Quant. Inf. Comp. **1**, 1 (2001), [quant-ph/0102024].

[56] D. Kaszlikowski and M. Żukowski, Int. J. Theor. Phys. **42**, 1023 (2003), [quant-ph/0302165].

[57] J.S. Bell, Physics **1**, 195 (1964). Reprinted in [142].

[58] J.F. Clauser, M.A. Horne, A. Shimony, and R.A. Holt, Phys. Rev. Lett. **23**, 880 (1969). Erratum: *ibid.* **24**, 549 (1970).

[59] S.J. Freedman and J.F. Clauser, Phys. Rev. Lett. **28**, 938 (1972).

[60] M.A. Rowe, D. Kielpinski, V. Meyer, C.A. Sackett, W.M. Itano, C. Monroe, and D.J. Wineland, Nature **409**, 791 (2001).

[61] N.D. Mermin, Phys. Rev. Lett. **65**, 1838 (1990).

[62] N. Gisin and H. Bechmann-Pasquinucci, Phys. Lett. A **246**, 1 (1998), [quant-ph/9804045].

[63] D. Collins, N. Gisin, S. Popescu, D. Roberts, and V. Scarani, Phys. Rev. Lett. **88**, 170405 (2002), [quant-ph/0201058].

[64] Z. Zhao, T. Yang, Y.-A. Chen, A.-N. Zhang, M. Żukowski, and J.-W. Pan, Phys. Rev. Lett. **91**, 180401 (2003), [quant-ph/0302137].

[65] J. Volz, C. Kurtsiefer, and H. Weinfurter, Appl. Phys. Lett. **79**, 869 (2001).

[66] Z. Zhao, Y.-A. Chen, A.-N. Zhang, T. Yang, H. Briegel, and J.-W. Pan, Nature **430**, 54 (2004), [quant-ph/0401096].

[67] M. Żukowski and D. Kaslikowski, Phys. Rev. A **56**, 1682 (1997), [quant-ph/9910045].

[68] P. Horodecki, Phys. Rev. Lett. **90**, 167901 (2003), [quant-ph/0111082].

[69] A. Sanpera, R. Tarrach, and G. Vidal, Phys. Rev. A **58**, 826 (1998), [quant-ph/9801024].

[70] P. Hyllus, PhD-Thesis, in preparation.

[71] D. Bruß, D.P. DiVincenzo, A. Ekert, C.A. Fuchs, C. Macchiavello, and J.A. Smolin. Phys. Rev. A **57**, 2368 (1998), [quant-ph/9705038].

[72] T.-C. Wei and P.M. Goldbart, Phys. Rev. A **68**, 042307 (2003), [quant-ph/0307219].

[73] H. Weinfurter and M. Żukowski, Phys. Rev. A, **64**, 010102 (2001), [quant-ph/0103049].

[74] M. Eibl, S. Gaertner, M. Bourennane, C. Kurtsiefer, M. Żukowski, and H. Weinfurter, Phys. Rev. Lett. **90**, 200403 (2003), [quant-ph/0302042].

[75] H. Briegel and R. Raussendorf, Phys. Rev. Lett. **86**, 910 (2001), [quant-ph/0004051].

[76] T.D. Howell, Linear Algebra and Appl. **22**, 9 (1978).

[77] J. Håstad, J. Algorithms **11**, 644 (1990).

[78] M.D. Atkinson and N.M. Stevens, Linear Algebra and Appl. **27**, 1 (1979).

[79] P. Pamfilos, Acta Math. Hung. **45**, 9 (1985).

[80] J.B. Kruskal, Linear Algebra and Appl. **18**, 95 (1977).

[81] R.A. Horn and C.R. Johnson, *Matrix analysis*, (Cambridge University Press, Cambridge, 1985).

[82] R.A. Horn and C.R. Johnson, *Topics in matrix analysis*, (Cambridge University Press, Cambridge, 1991).

[83] B.M. Terhal and P. Horodecki, Phys. Rev. A **61**,040301 (2000), [quant-ph/9911117].

[84] A. Sanpera, D. Bruß, and M. Lewenstein, Phys. Rev. A **63**, 050301 (2001), [quant-ph/0009109].

[85] F. Hulpke, D. Bruß, M. Lewenstein, and A. Sanpera, quant-ph/0401118.

[86] G. Tóth, Phys. Rev. A **69**, 052327 (2004), [quant-ph/0310039].

[87] R. Raussendorf and H. Briegel Phys. Rev. Lett **86**, 5188 (2001), [quant-ph/0010033].

[88] M.A. Nielsen, Phys. Rev. Lett. **83**, 436 (1998), [quant-ph/9811053].

[89] W. Dür, H. Aschauer, and H.-J. Briegel, Phys. Rev. Lett. **91**, 107903 (2003), [quant-ph/0303087].

[90] M. Hein, J. Eisert, and H.-J. Briegel, Phys. Rev. A **69**, 062311 (2004). [quant-ph/0307130].

[91] M. Barbieri, F. De Martini, G. Di Nepi, P. Mataloni, G.M. D'Ariano, and C. Macchiavello, Phys. Rev. Lett. **91**, 227901 (2003), [quant-ph/0307003].

[92] P.G. Kwiat, K. Mattle, H. Weinfurter, A. Zeilinger, A.V. Sergienko, and Y. Shih, Phys. Rev. Lett. **83**, 4725 (1999).

[93] N. Kiesel, M. Bourennane, C. Kurtsiefer, H. Weinfurter, D. Kaszlikowski, W. Laskowski, and M. Żukowski, J. Mod. Optics **50**, 1131 (2003).

[94] M. Eibl, N. Kiesel, M. Bourennane, C. Kurtsiefer, and H. Weinfurter Phys. Rev. Lett. **92**, 077901 (2004).

[95] J.L. Cereceda, quant-ph/0402198.

[96] H. Häffner, M. Riebe, C. Roos, W. Hänsel, M. Chwalla, J. Benhelm, F. Schmidt-Kaler and R. Blatt, Verhandl. DPG (VI) **39**, 7/148 (2004).

[97] M.D. Reid and P.D. Drummond, Phys. Rev. Lett. **60**, 2731 (1988).

[98] L.-M. Duan, G. Giedke, J.I. Cirac, and P. Zoller, Phys. Rev. Lett. **84**, 2722 (2000), [quant-ph/9908056].

[99] R. Simon, Phys. Rev. Lett. **84**, 2726 (2000), [quant-ph/9909044].

[100] P. van Loock and A. Furusawa, Phys. Rev. A **67**, 052315 (2003), [quant-ph/0212052].

[101] G. Tóth, C. Simon, and J.I. Cirac, Phys. Rev. A **68**, 062310 (2003), [quant-ph/0306086].

[102] H.F. Hofmann and S. Takeuchi, Phys. Rev. A **68**, 032103 (2003), [quant-ph/0212090].

[103] H.F. Hofmann, Phys. Rev. A **68**, 034307 (2003), [quant-ph/0305003].

[104] W. Heisenberg, Z. Phys. **43**, 172 (1927). Reprinted in [141].

[105] H.P. Robertson, Phys. Rev. **34**, 163 (1929).

[106] H.P. Robertson, Phys. Rev. **46**, 794 (1934).

[107] I. Białynicki-Birula and J. Mycielski, Commun. Math. Phys. **44**, 129 (1975).

[108] H. Maassen and J.B.M. Uffink, Phys. Rev. Lett. **60**, 1103, (1988).

[109] K. Życzkowski and H.-J. Sommers, J. Phys. A **34**, 7111 (2001), [quant-ph/0012101] .

[110] R. Horodecki and M. Horodecki, Phys. Rev. A **54**, 1838, (1996), [quant-ph/9607007].

[111] R.A. Bertlmann, H. Narnhofer, and W. Thirring, Phys. Rev. A **66**, 032319 (2002), [quant-ph/0111116].

[112] I. Montvay and G. Münster, *Quantum Fields on a Lattice*, (Cambridge University Press, Cambridge, 1994).

[113] R.M. Wilcox, J. Math. Phys. **8**, 962 (1967).

[114] G. Giedke, *Quantum information and Continuous Variable Systems*, (Dissertation, Leopold-Franzens-Unversität Innsbruck, 2001), avaiable at `http://www.phys.ethz.ch/~giedke/publications.html`.

[115] R.F. Werner and M.M. Wolf, Phys. Rev. Lett. **86**, 3658 (2001), [quant-ph/0009118].

[116] G. Giedke, B. Kraus, M. Lewenstein, and J.I. Cirac, Phys. Rev. Lett. **87**, 167904 (2001), [quant-ph/0104050].

[117] C. Shannon and W. Weaver, *The Mathematical Theory of Communication*, (University of Illinois, Urbana, 1949).

[118] J. Havrda and F. Charvat, Kybernetika **3**, 30 (1967).

[119] C. Tsallis, J. Stat. Phys. **52**, 479 (1988).

[120] A. Rényi, *Valószínűségszámítás*, (Tankönyvkiadó, Budapest, 1966). English translation: *Probability theory*, (North Holland, Amsterdam, 1970). German translation: *Wahrscheinlichkeitsrechnung*, (Deutscher Verlag der Wissenschaften, Berlin, 1979).

[121] A. Wehrl, Rev. Mod. Phys. **50**, 221 (1978).

[122] N. Canosa and R. Rossignoli, Phys. Rev. Lett. **88**, 170401 (2002).

[123] R. Rossignoli and N. Canosa, Phys. Rev. A **66**, 042306 (2002).

[124] R. Rossignoli and N. Canosa, Phys. Rev. A **67**, 042302 (2003).

[125] D. Deutsch, Phys. Rev. Lett. **50**, 631 (1983).

[126] K. Kraus, Phys. Rev. D **35**, 3070 (1987).

[127] H. Maassen, in *"Quantum Probability and Applications V"*, Lecture Notes in Mathematics 1442, edited by L. Accardi, W. von Waldenfels, (Springer, Berlin, 1988), p. 263.

[128] M. Krishna and K.R. Parthasarathy, Sankhya: The Indian Journal of Statistics, Series A, **64**, 842 (2002), [quant-ph/0110025].

[129] J. Sánchez, Phys. Lett. A **173**, 233 (1993).

[130] V. Majerník and E. Majerníková, Rep. Math. Phys. **47**, 381 (2001).

[131] J. Sánchez-Ruiz, Phys. Lett. A **244**, 189 (1998).

[132] G.-C. Ghirardi, L. Marinatto, and R. Romano, Phys. Lett. A **317**, 32 (2003), [quant-ph/0310120].

[133] N.J. Cerf and C. Adami, Phys. Rev. Lett. **79**, 5194 (1997), [quant-ph/9512022].

[134] R. Horodecki, P. Horodecki, and M. Horodecki, Phys. Lett. A **210**, 377, (1996).

[135] S. Abe and A.K. Rajagopal, Physica A **289**, 157 (2001), [quant-ph/0001085].

[136] K.G.H. Vollbrecht and M.M. Wolf, J. Math. Phys. **43**, 4299 (2002), [quant-ph/0202058].

[137] R. Horodecki and P. Horodecki, Phys. Lett. A **194**, 147, (1994).

[138] V. Giovannetti, Phys. Rev. A **70**, 012102 (2004), [quant-ph/0307171].

[139] R. Bhatia, *Matrix analysis*, (Springer, Berlin, 1997).

[140] E.T. Whittaker and G.N. Watson, *A Course of Modern Analysis* (Cambridge University Press, 1927).

[141] K. Baumann and R.U. Sexl, *Die Deutungen der Quantentheorie*, (Vieweg, Braunschweig, 1987).

[142] J.S. Bell, *Speakable and Unspeakable in Quantum Mechanics*, (Cambridge University Press, Cambridge 1988).

# List of Publications

[I] 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.

[II] 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.

[III] 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**, 1079-1102 (2003). Preprint: quant-ph/0210134.

[IV] O. Gühne and P. Hyllus: *Investigating three qubit entanglement with local measurements.* Int. J. Theor. Phys. **42**, 1001-1013 (2003). Preprint: quant-ph/0301162.

[V] O. Gühne, *Characterizing entanglement via uncertainty relations.* Phys. Rev. Lett. **92**, 117903 (2004). Preprint: quant-ph/0306194.

[VI] M. Bourennane, M. Eibl, C. Kurtsiefer, S. Gaertner, H. Weinfurter, O. Gühne, P. Hyllus, D. Bruß, M. Lewenstein, and A. Sanpera, *Experimental detection of multipartite entanglement using witness operators.* Phys. Rev. Lett. **92**, 087902 (2004). Preprint: quant-ph/0309043.

[VII] 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).

[VIII] O. Gühne and M. Lewenstein, *On entropic uncertainty relations and entanglement.* Phys. Rev. A (to be published). Preprint: quant-ph/0403219.

# Acknowledgements

I would like to thank the following people which made this thesis possible.

- First of all, I wish to thank Prof. Maciej Lewenstein for being the ideal supervisor for this thesis. His ideas, the discussions with him, and his interest in my ideas always motivated me and opened my mind to new directions.
- I am indebted to Priv.-Doz. Dagmar Bruß and Priv.-Doz. Anna Sanpera. They showed me many interesting topics to study and in many discussions with them I learned a lot of new things. I also want to thank Dagmar Bruß for refereeing this thesis.
- I wish to thank Philipp Hyllus for the nice and fruitful collaboration on entanglement witnesses, his help in many technical calculations, and for all the discussions on quantum theory.
- Then, I would like to express my gratitudes to the whole theoretical quantum optics group for the stimulating atmosphere. Among these people, I would like to mention Kai Eckert, Florian Hulpke, and, again, Philipp Hyllus for the careful proofreading of this thesis.
- Mohamed Bourennane, Harald Weinfurter and coworkers performed the experiments presented in this thesis. I would like to thank them for doing that and for showing me in this way that theoretical physics can be useful for experimentalists and vice versa.
- Géza Tóth convinced me of the elegance of the stabilizer formalism. I want to thank him for the good collaboration on that.
- I benefitted a lot from discussions with many people at several occasions. Among these people, I want to mention Marcos Curty, Manfred Eibl, Géza Giedke, Holger Hofmann, and Frank Verstraete.
- This work was supported by the Deutsche Forschungsgemeinschaft via the Graduiertenkolleg „Quantenfeldtheoretische Methoden in der Teilchenphysik, Gravitation, Statistischen Physik und Quantenoptik“. I like to thank the speaker of the Graduiertenkolleg, Prof. Lechtenfeld, and all the members. Especially, I would like to thank Marco Krohn for sharing the office with me and Sebastian Uhlmann for his help concerning LaTeX.

Zu guter Letzt möchte ich auch noch jene nennen, die mich in allen nichtphysikalischen Angelegenheiten unterstützt haben. Ich danke meinen Eltern, die immer hinter mir standen. Und natürlich Nicola, für alles.

# Lebenslauf

| | |
|---|---|
| 15. Mai 1975 | Geboren in Münster |
| August 1981- Juli 1982 | Besuch der Gottfried von Cappenberg-Grundschule in Münster |
| August 1982 - Juli 1985 | Besuch der Overberg-Grundschule in Warendorf |
| August 1985 - Juni 1994 | Besuch des Gymnasium Laurentianum in Warendorf |
| Juni 1994 | Abitur am Gymnasium Laurentianum in Warendorf |
| Juli 1994 - September 1995 | Zivildienst im Josephshospital in Warendorf. |
| Oktober 1995 - Februar 2001 | Studium der Mathematik und Physik an der Westfälischen Wilhelms-Universität Münster. |
| 26. September 1997 | Vordiplom in Mathematik |
| 26. September 1997 | Vordiplom in Physik |
| 22. Februar 2001 | Diplom in Physik |
| Juni 2001 - Mai 2004 | Promotionsstudium an der Universität Hannover. Doktorand in der Arbeitsgruppe von Prof. Dr. Maciej Lewenstein. Stipendiat am Graduiertenkolleg „Quantenfeldtheoretische Methoden in der Teilchenphysik, Gravitation, Statistischen Physik und Quantenoptik“ der Universität Hannover. |

*Abschied an den Leser*

*Wenn du von allem dem, was diese Blätter füllt,*
*Mein Leser, nichts des Dankes wert gefunden:*
*So sei mir wenigstens für das verbunden,*
*Was ich zurück behielt.*

Gotthold Ephraim Lessing