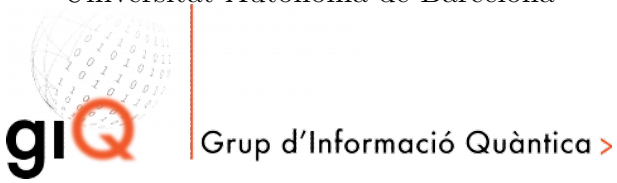
Correlations in the Generalized Bloch Picture & Applications in Entanglement Detection and Quantum Thermodynamics

Claude Köckl

Universitat Autònoma de Barcelona

Correlations in the Generalized Bloch Picture & Applications in Entanglement Detection and Quantum Thermodynamics

Thesis for the PhD in Physics at Fisica Teorica: Informacio i fenomens quantics, Universitat Autonoma de Barcelona



Claude Klöckl under supervision by: Andreas Winter & Marcus Huber

July 21, 2016

For Tamara,Gabi,Karim & Chris
I thank Marcus Huber & Andreas Winter for their support & advice

To

LIQUID

for making Barcelona what it was. Thank you Marcus, Stefan, Ima, Flo, Alexia, Pauli, Boris & Antoine.
I furthermore thank Amadeo & Janosch, Max, Kathrin for the joys of day to day life,
Chris, Maui, Sophie for their long friendships but also together with the whole U5 for sweat, blood and passion and guiding me through my first fight
Machine Manuel, Lock Lukas, Stricoletto, Thomas, Camilla, Leo, Sophie (again) for the longest nights and the most beautiful days
Maria for sharing three great cities and even more moments, Theo for motivation and pump it hard, Zahid & Romana for supporting me on the last bit and all the way, Iris for welcoming me at that time
Michi, Arno, Kathrin, Bella, Thomas, Bene & the whole crew for embarking on crazyness and carrying on the fire
Roberto, Jojo, Sergio and of course the team of GIQ for a great time even outside of LIQUID
Ferhat & the 6 comrades of this one long night I hope that you find real justice! Keep fighting!
Marlen, i thank you for the shared time and want to excuse me for the troubles this thesis may have caused you

Contents

Chapter 1

Introduction

The overarching theme of this thesis is the study of correlations. We approach the issue from two distinct perspectives: *entanglement theory* and *quantum thermodynamics*. We will relate correlations to entanglement, but also temperature and the work cost of creating correlations. Along the way, we will encounter and connect the description of correlations to the Bloch picture [4] and possible generalizations of the Pauli operators. Both can be seen as other reoccurring themes and we hope to motivate new investigations of these established but not exhausted topics through this thesis.

In the first and larger part of this thesis correlations will be used as tools to detect entanglement. Apart from the development of entanglement detection criteria, we hope to understand better how to best describe the correlations present within a general quantum mechanical system. Throughout this thesis, we propagate to represent correlations through the use of the well-known Bloch picture. It is equipped with a natural notion of locality, elegantly orders correlations by number of involved particles and allows for simple implementation of the partial trace. We advocate the employment of the Bloch decomposition not only for a single qubit, but for n-partite qudit systems. Traditionally the Bloch ball is applied to picture single qubits [4]. We want to convey our believe that the Bloch decomposition is not merely a visualization aid but a sensible technique that merits further study. Generalizing the Bloch ball to a higher dimensional or multipartite analogue is certainly a well studied subject (c.f.:[16, 31, 32, 3, 33, 34, 19]). Many of the authors writing about the subject are concerned with finding positivity conditions for higher dimensional systems, visualizing the set of physical states for qubits but also developing novel applications of the Bloch picture (even in rather unexpected applications like pattern recognition [41]). This thesis represents a contribution that tries to extend the application of the generalized Bloch decomposition to the field of multipartite entanglement detection. It can be seen as a continuation of recent progress in this direction (c.f.: [8],[9],[10]). When discussing past generalization attempts, it will become quickly apparent that the generalized Bloch decomposition is almost completely determined by the composition axiom of quantum mechanics. The only freedom arises when fixing the basis for the description of a single qudit system. This intimately connects multipartite entanglement detection with the generators of the SU(d) through the generalized Bloch decomposition. Entanglement detection problems motivate to consider some unconventional properties like requirements on rank or anti-commutativity when generalizing the Pauli basis. The canonical choices for extensions of the Pauli operators have

been derived with the objective of retaining either Hermiticity or unitarity. Hopefully by proposing new properties as sensible for generalization, we also renew interest in discussing how to best generalize the Pauli matrices.

The second part of this thesis represents a quantum thermodynamically motivated approach towards studying correlations. There we try to understand the limits imposed on the creation of correlations through thermodynamics. Before being able to answer this question, we will have to repeat some of the discussions that have been led in the recently active quantum thermodynamics community [18]. We have to clarify how notions such as temperature and work cost show up in a quantum information framework. Trying to answer the question of how much work is necessary to create correlations, will culminate in our search for optimally correlation creating unitaries. These unitaries turn out to be not easily constructable, but we can derive optimality conditions and point out connections to circulant matrices and the theory of unistochasticity.

1.1 Outline of this thesis

In Chapter 2 we introduce all notions and notations, that will be used throughout this thesis. We will be touching issues from quite different sub-fields of quantum physics. To reflect this Chapter 2 is structured into four sections, each aims to introduce one of the several theories we will be covering in this work. Section 2.1 gives a quick introduction into the theory of quantum mechanics. We briefly introduce a set of axioms for quantum mechanics subsection by subsection. Section 2.2 takes a closer look onto the implications of these axioms. We discuss possible ways to classify the physical state space, defining notions such as purity & mixedness as well as separability & entanglement. The second part 2.2.2 of Section 2.2 also gives a brief overview of entanglement theory. Section 2.3 is used to give some of the basic notions from the recently active field of quantum thermodynamics. We close this chapter with the explanation of some tools for the treatment of continuous variable systems in Section 2.4.
To reflect this Chapter 2 is structured into four sections, each aims to introduce one of the several theories we will be covering in this work. Section 2.1 gives a quick introduction into the theory of quantum mechanics. We briefly introduce a set of axioms for quantum mechanics subsection by subsection. Section 2.2 takes a closer look onto the implications of these axioms. We introduce possible ways to classify the physical state space, introducing notions such as purity & mixedness as well as separability & entanglement. The second part 2.2.2 of Section 2.2 also gives a brief overview of entanglement theory. Section 2.3 is used to give some of the basic notions from the recently active field of quantum thermodynamics. We close this chapter with the explanation of some tools for the treatment of continuous variable systems in Section 2.4.

Chapter 3 treats the way correlations are represented in the Bloch vector formalism. Section 3.1 reviews the classical case as well as possible ways to generalize a famous representation of quantum systems: the Bloch vector representation. For us the Bloch vector decomposition is one of the main technical tools, that we will employ for entanglement detection later on. We review the standard treatment of the qubit in Bloch representation in Section 3.1.1, then proceed with covering possible ways to generalize the Bloch representation to higher dimensional system in Section 3.1.2. Many authors have already contributed and we can only give a sample of the available work on this topic. Generalizing to qudits is not a straightforward affair and there is no unique way to do this, but it is

agreed upon to be a complicated endeavor. We state some of the properties, that in our opinion are necessary for a meaningful generalization of the Bloch vector decomposition. These properties will fix a higher dimensional representation up to a choice of basis. Finally, we cover the treatment of multipartite systems in Bloch representation. Once we fix the representation of a single qudit and demand a tensor product basis for multipartite systems, the basis choice for multipartite systems is obviously determined. We use Section 3.1.3 to introduce notation that will be of importance for multipartite entanglement detection. Furthermore we introduce the notion of the correlation tensor, that will be the starting point for a novel reference frame independent entanglement detection strategy.

As we have pointed out in the last paragraph, the freedom involved when generalizing the Bloch vector representation, essentially boils down to a choice of basis. In the qubit case the Pauli basis is clearly the unique choice of basis. In fact, the choice of basis should be adapted to the problem at hand and there exists a vast body of literature on so-called SU(d) generators. We use Chapter 3.2 to introduce the canonical choices of bases for treating qudit systems: the generalized Gell-Mann basis (cf.: Section 3.2.1) and the Heisenberg-Weyl basis (cf.: Section 3.2.2). Alongside, discussing the merits of these well established bases, we also try to give a new perspective on the topic by suggesting a symmetrized Heisenberg-Weyl basis in Section 3.2.3 and discussing another aspect of the Pauli basis, that usually is not treated as relevant, when generalizing to higher dimension: the anti-commutativity of the Pauli basis.

After fully developing the Bloch vector decomposition we will apply it to the problem of entanglement detection in Chapter 4. This chapter contains many of our main results. We begin in Section 4.1, with some general considerations, before proposing concrete entanglement detection schemes. There we reflect briefly what an entanglement detection criterion is, what notions of entanglement can be detected in the multipartite case and how the noise resistance is an essential measure of quality for every entanglement detection criterion. After setting the stage, we commence with a series of sections proposing three different strategies for multipartite entanglement detection:

- First, in Section 4.2, we study the possibility to detect entanglement via the multiplicativity of norms under tensor products. This is the most general strategy, that is basically always applicable.
- Second, in Section 4.3, we propose bounds of squared correlation tensor elements as a tool to certify entanglement. We start in Section 4.3.1 by discussing the relation between the correlation tensor, the 2-norm, LU-invariance and the purity as a bounding procedure. We proceed in Section 4.3.2 with some considerations on possible forms of correlation tensors. This will yield a series of bounds and culminate in a no-go result stating: there are no states with correlations concentrated into a single sector. The last part is Section 4.3.3. Here we put the bounds into practice and demonstrate the possibility of detecting entanglement with these bounds for example for the GHZ and the cluster state.
- Third, in Section 4.4, we propose an approach that utilizes the anti-commutativity of a basis to detect entanglement. After motivating briefly why the notion of anti-commutativity should even be useful at all for entanglement detection in Section 4.4.1, we prove this section's main theorem in Section 4.4.2, discuss its applicability and limitations in Section 4.4.2 and propose a possible way to extend the main theorem's region of validity by relaxing to the notion of almost anti-commutativity in Section 4.4.3.

The last Chapter 5 covers quantum thermodynamics. During Chapter 4, we have studied correlations represented in the correlation tensor extensively as a tool to detect entanglement. Now we change the viewpoint and study the role of correlations in a thermodynamic setting. The larger part of this chapter is contained in the first Section 5.1. There we concern ourselves with questions regarding work or energy cost of unitary transformations. Section 5.1.1 tries to answer the question, how we should treat notions of thermodynamics within quantum mechanics. These notions stem from outside of the established theory and it is seldom obvious, what a "quantum mechanical" analogue of for example temperature, heat or energy is. These questions have been extensively treated lately [18]. We will focus on the interplay between correlations (a quantum information notion) and work cost (a thermodynamical notion). It is generally not easy to find explicit constructions for unitary operations that achieve tasks like creating correlations with an optimal energy consumption. In Section 5.1.2, we try to find some conditions that would ensure optimality of such an unitary.

Before proceeding, to the main part of the text, we want to highlight the parts of this thesis that constitute novel work. In the process of writing this thesis three peer reviewed articles [28, 27, 2] were published. We also want to remark that an entanglement detection criterion proposed in [27] has already been cited and used for entanglement detection in a recent experiment (see [49]).

(i) The exposition of established notions in Chapter 2 naturally contains no new work. Only the very last Subsection 3.1.3 introduces the terminology of one of our papers [27] for use in Chapter 4.

(ii) Chapter 3.1 reviews well-known literature.

(iii) Chapter 3.2 mixes well established notions, with contributions from us. Obviously the sections on the generalized Gell-Mann basis and the Heisenberg-Weyl operators review the work of other authors, however we believe that we contribute an original perspective. Canonically the properties of unitarity and hermiticity are considered to be the desirable features of the Pauli basis, when considering generalizations. In addition to this, we introduce, the question how to generalize the non trivial anti-commutativity relations of the Pauli basis properly. To the best of our knowledge this is a novel approach to the topic. Furthermore, in Section 3.2.3 we propose the symmetrized Heisenberg-Weyl basis as an interesting basis for qudit systems. This material has been published in [2] We want to comment, that we are aware that the construction proposed in Section 3.2.3 is simple. Therefore it is surprising, that we know of no references for this basis. To the best of our knowledge it is new work, even though it seems possible that it has been considered before. We believe, that since the canonically studied properties of the symmetrized Heisenberg-Weyl operators and the generalized Gell-Mann basis coincide the community may have asserted the symmetrized Heisenberg-Weyl operators to be redundant. We think that only the consideration about full rank and anti-commutativity, that we introduce motivates further study in this direction. At the very least Section 3.2.3 represents an unconventional approach toward generalizing the Pauli basis.

(iv) Chapter 4 consists solely of original work. The first Section 4.2 is unpublished material contained only within this thesis. The later Sections 4.3 & 4.4 were published in [27, 2].

(v) Chapter 5 consists entirely of work published in [28].

Chapter 2

Definitions & Preliminaries

2.1 Quantum mechanics

In this section we aim to introduce quantum mechanics in a nutshell, as short and concise as possible. Clearly this can be only a very brief introduction into a well developed subject, without a lengthy discussion of all implications. More elaborated expositions can be found in many textbooks. For example, the standard introduction from a quantum information perspective is Nielsen & Chuang [39]. Alternatively, there is an introduction to quantum information theory by Wilde [52].

Aside from a few motivational remarks in the first section, we will restrain ourselves to simply state the fundamental axioms of quantum mechanics and then work one by one through each axiom in detail. We are giving exact definitions, introducing the relevant mathematical objects and some comments on why they appear in the way they do. In fact, one has to mention that even those fundamental "axioms" are not as fundamental as the name suggests. Many of these axioms could be also formulated differently. They are certainly sufficient to derive all the theory behind quantum mechanics, however there is even today considerable disagreement, regarding what particular formulation may be the most appealing one. Frequently some of the corollaries of our axioms are in fact equivalent to those very axioms and could be themselves be the base for another as sensible axiomatization of quantum mechanics. At the end of the day these fundamental issues are to a big extent a matter of personal taste and for many concrete problems in quantum mechanics irrelevant. Moreover, they are not the focus of this thesis.

What is quantum mechanics?

Prior to the more technically minded main part of the introduction, we commence by briefly giving an intuitive approach to motivate the concepts appearing later on. Just as the exact mathematical formulation of quantum mechanics, can differ significantly between different authors, the interpretation and intuition of quantum mechanic's fundamental laws vary widely. This paragraph's aim is much more to convey a general feeling for the later appearing mathematical definitions in order to allow readers first encountering them to grasp them more easily, than to give the "real" or "best" interpretation of quantum mechanics.

So, what is (quantum) mechanics?

Basically mechanics simply tries to describe the behavior the *things* that surround us. What exactly these things are, what they are composed of and why they act like they do, have been fundamental questions throughout all of history. It has been answered in many different fashions by different theories, one of them being quantum theory. Now more specifically whenever we want to introduce a description of something we want to study, we have to somehow label the object of study. This labeling is best done without assuming too much about its ontology. Whatever it may be, it is the one thing we are right now focusing on. We will refer to our current object of study with a quite general catch all phrase: *a physical system*. This notion is almost an axiom in itself, it simply refers some part of reality (or whatever it may be that determines our perceptions).

However a system is usually not just a single object, it is rather the collection of all relevant factors to describe an object. An object of study may behave very differently depending on the context presently surrounding it, like the same bit of water behaving vastly different at zero and at one hundred degrees. In a way this description of an object, including its presently relevant context, is what we will be referring to as the *state of a system* in quantum mechanics.

Note that we have used the term: presently! A state of a system somehow is also connected to a temporal description. Each state of a system is implicitly assumed to happen at a particular point in time. The behavior of everything we study may change with time, even though its development may be determined or at least influenced by the present or the past. These changes of a system between one point in time and another later one is something we will be studying as well. Formally we will call them *state transformations*.

Finally, quantum mechanics (in contrast to classical mechanics) separates strongly between the system we study and what we perceive of it. Many systems may be not directly perceivable by us, but only indirectly through the use of a device, an experiment or an easier visible effect. Quantum mechanics assumes, that what we can observe is at least influenced by something deeper unobservable and that our observations allow us to infer something systematic about those governing influences. These possibly covered aspects of the world are already described by the states of the system of study. In contrast the directly accessible part of human experience deserve of course special attention since they will the base of all reasoning we can do. Thus they are specially named and formalized into the notion of *the observable on a system*.

At its heart quantum mechanics consist of only these three fundamental building blocks: *states* ρ, *observables* O and *state transformations* U. Each of these building blocks is represented by a different class of operators, we will see that *states are positive Hermitian operators*, *observables are Hermitian operators* and finally *state transformations are unitary operators*.

2.1.1 Axiom 1

The first axioms specifies how a physical system is to be described, or maybe more exactly how the current *state of a system* is described at any given point in time.

We aim to describe a physical system by an operator. In order to properly speak about an

operator, we will also need to consider its domain. The domains will all be so-called Hilbert spaces.

Definition 2.1.1. (Hilbert space)
A **Hilbert space** $\mathcal{H}$ is a pair $(\mathcal{H}, \langle .|. \rangle)$, where $\mathcal{H}$ is a complex vector space and $\langle .|. \rangle$ is a positive semidefinite scalar product, such that with respect to the norm induced by it $\mathcal{H}$ is complete.

Note that the definition of the Hilbert space does not mention any condition on the dimension of the vector space $\mathcal{H}$. This is very much by intention, in order to allow for finite as well as infinite dimensional vector spaces. Even if all appearing domains may have the same mathematical properties, we will see that especially in the standard finite dimensional case, the size of the domains associated with any given operator is playing an important role and we will have to consider it carefully. Sometimes it is casually stated that all finite dimensional Hilbert spaces are the "same". This may be true if only concerning oneself with some more abstract mathematical properties of the space like topologies or norm equivalences. In this thesis we are considering many more physically motivated questions like entanglement detection or the number of measurements required to detect a state. These problems will exhibit a considerable gap not only between infinite and finite dimensional systems as usual, but also for example between the two-dimensional and the higher finite dimensional case. A large part of this thesis will try to discuss suitable bases for the underlying Hilbert spaces. Even though one can always convert one basis into all other bases, on a practical level they are far from equivalent. It turns out that some bases are more suited to represent some systems and other are more suited for others. A situation not at all unfamiliar from the standard coordinate system where the standard basis, cylindrical and spherical coordinate systems may be suited to different geometries within a system, even though related to each other by basis transformations. The two-dimensional Hilbert space is somewhat singled out, since there is one undisputed basis available to represent it: the Pauli basis. Higher dimensional systems allow different descriptions, where it is not so obvious to identify the right representation. Both cases will appear during this thesis, however in the field of quantum information generally finite dimensional systems are to be studied.

For the most part of this thesis and also most all of quantum information, we choose

$$\mathcal{H} := \mathbb{C}^d, \tag{2.1}$$

unless explicitly mentioned.

It is noteworthy that this is in a mathematical sense a separable Hilbert space, meaning that there exists a dense and countable basis for $\mathcal{H}$. Not to be confused with the physical notion of separability of a state (Note furthermore, that one notion labels a property of a space and the other a property of an operator). A countable basis is usually a prerequisite in order to use the standard formulation for the well-known bra-ket notation. In addition the base field of the vector field in quantum mechanics is required to be complex.

Summing up in a quantum mechanical setting the Hilbert spaces of choice are always required to be complex and separable vector spaces.

There are interesting related fields, namely quantum optics, where the standard setting treats infinite dimensional vector spaces that can not be identified with the above finite dimensional vector

space

$$\mathcal{H} = L^2(U) := \left\{ f : U \mapsto \mathbb{C} | \int_U |f(x)|^2 dx \leq \infty, U open \right\}, \tag{2.2}$$

the so-called **space of square integrable functions**. As we mentioned above all Hilbert spaces need to be defined in conjunction with a scalar product. As already apparent from the above definition of L^2, we use

$$\langle f|g\rangle_{L^2(U)} := \int_U f(x)\bar{g(x)}dx \tag{2.3}$$

on the function spaces arising in quantum optics.

We will address the exact definition of the continuous variable case, alongside a comparison to finite Hilbert spaces, at the end of the introductory chapter in 2.4.

We will describe it by its so-called *density matrix*, since it is the most general way to describe the state of a system.

Traditionally a slightly different construction for the physical states is used. First, so-called pure states, that are described as rays in Hilbert space are introduced. The convex combination of all pure states finally yields all admissible states. We work somewhat the other way round, first introducing a general description of all states and then characterizing a subset of them as pure later on. Both approaches are of course completely equivalent.

Definition 2.1.2. (Density matrix)

We define the **density matrix ρ of a quantum mechanical system** as a linear operator on a complex separable Hilbert space $\mathcal{H}$, that is required to be:

(i) (Normalized):

$$\mathrm{Tr}(\rho) = 1 \tag{2.4}$$

(ii) (Hermitian):

$$\rho = \rho^\dagger = (\rho^*)^t \tag{2.5}$$

(iii) (Positive semi-definite):

$$\rho \geq 0 \Leftrightarrow x^\dagger \rho x \geq 0 \forall x \in \mathcal{H} \tag{2.6}$$

Often the density matrix is in a simplified manner also referred to as the **state of ρ**.

The properties (Normalized) & (Positive semi-definite) should seem very convincing, once it is mentioned that they are basically directly imported from the definition of a probability. We simply require all measurements to follow a probability distribution, thus requiring no event can have a negative probability (Positive semi-definite) and all possible event's probabilities sum up to one (Normalized). Hermiticity of an operator implies, that it has real eigenvalues. Thus (Hermitian) ensures a real spectrum, a requirement that will be well motivated by the third axiom.

In addition to the scalar product of the Hilbert space, that is defined for vectors, we also need a scalar product for matrices.

Definition 2.1.3. (Hilbert-Schmidt product)

Let $A, B \in \mathbb{C}^{d\times d}$ be operators.

$$\langle A|B\rangle = \mathrm{Tr}(A^\dagger B) \tag{2.7}$$

as the so-called **Hilbert-Schmidt product**.

Finally, we can collect the above definitions to state the first axiom:

Definition 2.1.4. (Axiom 1)
A system is described by quantum mechanics by a pair $(\rho, \mathcal{H})$ consisting of a density matrix of a state, as well as a Hilbert space $\mathcal{H}$.

2.1.2 Axiom 2

The second axiom is concerned with finding a proper description of the rather abstract concept of an *observable*.

Definition 2.1.5. An observable O of a system $(\rho, \mathcal{H})$ is a Hermitian linear operator defined on the systems state space $\mathcal{H}$.

By this definition an observable is a very general object, where only Hermiticity is required. This is done to ensure a real spectrum. The reason for this will be apparent, when introducing the third axiom.

After introducing this very general description of generally perceivable quantities, we will give two more special cases of observables, that aim to formalize the concept of a *measurements*. Referring to a measurement, should indicate the first step of including a description of the instrument that is conducting the "observation" in our model.

The notion of a projective measurement boils down almost to a corollary of the definition of an observable O. Any complex Hermitian, thus diagonalizable, operator may be written in the eigenvalue decomposition. The possibility to diagonalize an observable may be used as a first motivation why to require the Hermiticity of O. Now simply labeling the resulting projectors yields our first formal description of measurement.

Definition 2.1.6. We call the projectors $|P_i\rangle\langle P_i|$ arising in the eigenvalue decomposition of an observable O, given by

$$O = \sum p_i |P_i\rangle\langle P_i|, \tag{2.8}$$

the **projective measurements** on O.

The measurements described by projective measurements are named **von-Neumann measurements**.

In a projective measurement, all possible outcomes will be associated to a different eigenvector-eigenvalue pair. The eigenvector's projector is furthermore the description of the system's state after measurement. Thus projective measurements are intimately connected to a description of the post-measurement states arising.

If we consider the properties of the above decomposition, we will see that only projectors arise in the decomposition. Furthermore, it is clear from the properties of the eigenvalue decomposition, that different eigenvectors are always orthogonal. This leads to very simple behavior of O, when repeatedly applying it. A projector applied to itself yields the very same projector, while it will vanish if applied to anything from the orthogonal complement. Thus the post measurement state can not change by repeated application of an observable.

Definition 2.1.7. (Axiom 2) Each quantity of the system that we can observe is represented as a so-called observable O acting on the state ρ.

O is a Hermitian linear Operator acting on $\mathcal{H}$ associated to ρ.

2.1.3 Axiom 3

The third axiom describes the expected result of measuring a observable O of a system ρ. We write expected result, because also this axiom describes measurements as random variables, that can only be describe statistics. Even more, it specifies how observable and system are related to give the proper statistics of the measurement result.

Definition 2.1.8. (Axiom 3)
Measuring system ρ in observable O, the expectation value of O is described by

$$\langle O \rangle := \mathrm{Tr}(\rho O) \tag{2.9}$$

Remember that since O is a Hermitian Operator it is expressible in its eigenvector decomposition. Measuring system ρ in observable O, the expectation value of O is described by

$$\mathrm{Tr}(\rho O) = \sum_i \lambda_i \underbrace{\mathrm{Tr}(|\lambda_i\rangle\langle\lambda_i|\rho)}_{:=p_i}. \tag{2.10}$$

The coefficients λ_i describe all possible measurement results, that could be attained.
The probability of measuring any particular λ_i is $p_i \overset{!}{=} \mathrm{Tr}(|\lambda_i\rangle\langle\lambda_i|\rho)$.
Furthermore, when λ_i has been measured, the post measurement state is $|\lambda_i\rangle\langle\lambda_i|$.
This statement is the famous: **Born rule**.

The Born rule explains the expected measurement statistics along side the corresponding post measurement state. After measurement, the state can only be found in some eigenstate of the observable. Measuring again, will result in a repeated measurement on an eigenstate, whose result is of course determined to be its (determined!) eigenvalue instead of a probability distribution.

2.1.4 Axiom 4

Up until now we have only described the *present* state of a system. This axiom is concerned in describing the *change* of a system, its so-called dynamics. In order to describe this we will develop the concept of so-called time *evolutions*, that describe the change of the system through time.

Probably the most important observable in physics is the Hamiltonian H. It is is intricately related to the concept of energy. In fact it is an observable that a fundamental observable that has to be supplied by formalizing empirical evidence about a certain system of interest. We will simply consider it as given, that it exists and describes the energy of a system.

Definition 2.1.9. We give the expectation value $\langle H \rangle$ a distinct name: the **energy of** ρ.

Apart from describing the energy of a system, the Hamiltonian describes another fundamental role. It describes the equations of motion of a system through the following differential equation.

Definition 2.1.10. (Liouville-von-Neumann equation) The time evolution of an initial state $\rho(0)$ under the Hamiltonian H is governed by the **Liouville-von-Neumann equation**

$$i\hbar \frac{d\rho}{dt} = [\rho, H], \tag{2.11}$$

where t is a parameter labeled **time**, H needs to be a Hamiltonian and $\rho(0)$ acts as the initial value of the initial value problem.

Note that the Liouville-von-Neumann equation is a matrix equation and the equivalent of the Schrödinger equation for general (i.e. mixed) states. The Schrödinger equation holds for pure states (i.e. vectors) and was historically the first equation to be used in quantum mechanics. We have chosen to directly define states over the more general density matrix, thus we also have to introduce time evolution by the more general matrix equation above.

Observation 2.1.11. *The solution to the Liouville-von-Neumann equation is*

$$\rho(t) = e^{\frac{i}{\hbar}Ht}\rho(0)e^{-\frac{i}{\hbar}Ht} \tag{2.12}$$

In the introduction we have stated that time evolutions in quantum mechanics are described by so-called unitary matrices.

Definition 2.1.12. A **unitary matrix** U is a matrix that fulfills:

$$U^\dagger U = \mathbb{1} \tag{2.13}$$

or equivalently

$$U^\dagger = U^{-1} \tag{2.14}$$

It is not immediately apparent how the von-Neumann equation 2.11 relates to unitaries. To see this, we state a lemma on the representability of unitary matrices.

Lemma 2.1.13. *Let U be a unitary, then we have for t' a scalar and $\tilde{H}$ be a Hermitian matrix,*

$$U = e^{it'H}. \tag{2.15}$$

Proof. Let us be given a Hermitian matrix H. Then we have

$$e^{itH}(e^{itH})^\dagger = e^{itH}e^{-itH} = e^{itH-itH} = 1, \tag{2.16}$$

where the first equality holds due to Hermiticity and the second one due to Baker-Campbell-Hausdorff formula and $[H, H] = 0$. We have proven that for any H, e^{itH} is unitary. The easy direction is all we need for the moment.
The other direction requires Lie theory and can be found in many textbooks, it is however not the focus of this thesis and omitted for simplicity. □

Now comparing the statement of the above lemma, to the solution of the von-Neumann-equation 2.12. We see that the matrices surrounding the initial state ρ are in fact unitaries just as above. Realizing this, we rewrite 2.12 into a simpler form and sum this section up to the next definition.

Definition 2.1.14. Let ρ be a state. Its **time evolution under** H is described by the mapping

$$\rho \to U\rho U^\dagger, \tag{2.17}$$

where $U = e^{itH}$ and t is a constant.

2.1.13 states that in fact every unitary is generated by a parameter t' together with a Hermitian matrix $\tilde{H}$. Implicitly, one can argue that all unitaries depend on these two quantities. If these quantities are fixed, then it is totally fine to drop them and write just U as above, whenever we only want to describe that ρ has evolved after some time to a desired target state. Of course then a lot of information about the system is lost since different t',H could lead to the same U. If we want to retain this information we need to include the time dependency explicitly in a $U(t)$ instead of U.

Definition 2.1.15. (Axiom 4)
Let ρ be a state at a certain point in time. After waiting for time t, the system is described by its **time evolution after** t **under** H. This is the mapping

$$\rho \to U(t)\rho U(t)^\dagger, \tag{2.18}$$

where $U(t) = e^{itH}$ and t is a parameter.

2.1.5 Axiom 5

We have now discussed how to describe a single system, how to describe measurements taken on that system and also how a single system changes with time. Intuitively, if you now have a complete description of two different isolated systems, one would expect that one should also be able to describe how they interact with each other. The final axiom formalizes this idea by stating that the tensor product is exactly the right operation to get a description of the joint system.

Definition 2.1.16. (Axiom 5)
Let $(\rho_1, \mathcal{H}_1), (\rho_2, \mathcal{H}_2)$ be two quantum mechanical systems, both fully described by a state-Hilbert space pair. The we define then the state space of the **composite system** is described by the tensor product

$$\mathcal{H}_{12} := \mathcal{H}_1 \otimes \mathcal{H}_2. \tag{2.19}$$

Also if the states of the two system are known to be ρ_1 and ρ_2, they can be combined into a composite state by the tensor product analogously. The tensor product applied to finite dimensional linear operators simplifies to the Kronecker product and $\rho_1 \otimes \rho_2$ is guaranteed to be a valid state on the composite system $\mathcal{H}_1 \otimes \mathcal{H}_2$. Still, the state space of $\mathcal{H}_1 \otimes \mathcal{H}_2$ is a significantly larger space and the Kronecker product of its marginal does not at all cover the whole composite state space. The study of all consequences arising from this axiom is not only one of the main points in this thesis, but in general a vast field in itself: *entanglement theory*. The following section 2.2 is therefore entirely devoted to a more in depth study of the geometric properties of composite state space(s).

2.1.6 Summary of the axioms for quantum mechanics

Summing up in one definition, what we have state in the previous pages, we formulate the fundamental axioms of quantum mechanics.

Definition 2.1.17. A Set of Axioms for Quantum Mechanics

(i) Each physical system is representable at a fixed point in time by a so-called state ρ in phase space $\mathcal{H}$.

ρ is required to be a positive semi-definite Hermitian linear operator defined on $\mathcal{H}$ a separable Hilbert space over the field $\mathbb{C}$

(ii) Each quantity of the system that we can observe is represented as a so-called observable O acting on the state ρ.

O is a Hermitian linear Operator acting on $\mathcal{H}$ associated to ρ.

(iii) The measurement of observables O on a state ρ is described through the scalar product defined on the phase space $(\mathcal{H}, \langle ., . \rangle)$. The action of the observable on the system is described by taking the Hilbert-Schmidt product of the observable on the state $\langle \rho O \rangle$. Each observable is decomposable into several events via spectral decomposition as $\langle O \rangle_\rho = \sum_i \lambda_i \langle |\lambda_i\rangle\langle\lambda_i| \rangle_\rho$. $p_i = \mathrm{Tr}(|\lambda_i\rangle\langle\lambda_i|\rho)$ are the probabilities to measure the event λ_i, when measuring O in state ρ.

(iv) The time evolution of all states is described by a so-called state transformation $U(t)$. After time evolution by time $t \in \mathbb{R}$, we find an initial state ρ exactly in the state $U(t)\rho U^\dagger(t)$.

(v) Two systems can be described together as a composite system. Composing systems is described by taking tensor products. Thus the composed system's space is the tensor product of its factors phase spaces.

2.2 Geometry of the state space

Up until now we have only considered the state space in very abstract terms. We have given precise mathematical properties required for logical consistency, however we have not yet developed a more qualitative intuition about it. In this section we want to collect what is known about the geometry of the state space. We address questions such as what do we know about its shape? We will also try to subdivide the state space into regions. In order to differentiate them we will identify physically relevant properties of states. Then we will identify, which regions are populated by states that exhibit these properties.

2.2.1 Purity & mixedness

In this part we introduce the classification of states into pure and mixed states.

Definition 2.2.1. Purity I (Projector formulation) A state ρ is called **pure** if it is a projector, i.e.

$$\rho^2 = \rho. \tag{2.20}$$

All non pure states are called **mixed**.

We have an in important alternative characterization of purity:

Lemma 2.2.2. *Purity II (Vector formulation) A state ρ is pure iff it can be written as the outer product of a vector $|\psi\rangle$,i.e.:*

$$\rho = |\psi\rangle\langle\psi| \tag{2.21}$$

Proof. By directly expressing $\rho = \sum_i \lambda_i |\lambda_i\rangle\langle\lambda_i|$ in eigenvalue decomposition in the above projector identity

$$\sum_i \lambda_i |\lambda_i\rangle\langle\lambda_i| = \rho = \rho^2 = \sum_i \lambda_i^2 |\lambda_i\rangle\langle\lambda_i|, \tag{2.22}$$

since $0 \leq \lambda_i \leq 1$ this can only hold for $\lambda_i = 1$. However this already saturates the bound imposed by normalization, thus forcing all coefficients in the above sum to be zero except one. □

By reformulating the above we get another important rephrasing of purity and mixedness in terms of the coefficients λ_i above.

Corollary 2.2.3. *The rank of a matrix ρ is the minimal number of non-zero coefficients λ_i in the decomposition into outer products of vectors ψ_i:*

$$\rho = \sum_i \lambda_i |\psi_i\rangle\langle\psi_i| \tag{2.23}$$

Note that the rank allows for a third formulation of purity due to:
A state of rank 1 is pure.
A state with rank larger than 1 is mixed.

Corollary 2.2.4. *Geometrical facts*

(i) The state space is convex.

(ii) ρ is pure $\Rightarrow$ ρ lies on the boundary of the convex state space

Proof. ad (i):
It is obvious that a convex combination of states is again normalized, positive and Hermitian, thus a state. It is also apparent from the above reformulation of purity in terms of rank that when writing a pure state as a convex combination it is an extremal state.
Furthermore, all non extremal states have a rank larger than one and are thus mixed. Since the whole state space is consisting only of pure and mixed states and can be obtained as a convex combination of the pure states the state space is convex.
ad (ii):
The state space is the convex hull of the pure states. Thus they are on the boundary. □

We need to point out that the converse of the property (ii) does not hold in general. It is true for qubits, but not for higher dimension. We will discuss the important consequences at length throughout the thesis starting with 3.1.

In fact the definition of being a projection can be turned into a trivial but important corollary.

Lemma 2.2.5. *For all states we have*

$$\frac{1}{d} \leq \mathrm{Tr}(\rho^2) = 1. \tag{2.24}$$

We call the quantity $\mathrm{Tr}(\rho^2)$ *the* ***purity*** *of* ρ.

Proof. Obvious from the definition of a projector and the normalization, we have $\mathrm{Tr}(\rho^2) = \mathrm{Tr}(\rho) = 1$ for all pure states. $\mathrm{Tr}(\rho^2)$ squares all the coefficients when expressing ρ in the eigenvalue basis. Thus it is the squared two norm of the coefficients. It is obvious that mixed states can only decrease the purity, since 2-norms are maximal for vectors with only one entry. Similarly, it can be seen that the minimum of a two norm is attained on a vector with uniformly distributed entries. □

In fact the quantity turns out to be a quantification of the initially qualitative concept of purity. We will have observed that the maximum of $\mathrm{Tr}(\rho^2) = 1$ is attained exactly on the pure states, while the minimum $\mathrm{Tr}(\rho^2) = \frac{1}{d}$ is attained on the identity $\frac{1}{d}\mathbb{1}$, which is why the identity is also sometimes referred to as the maximally mixed state.

The purity is often regarded as measure of distance from the maximally mixed state $\frac{1}{d}\mathbb{1}$, that minimizes it. The larger the purity of a density matrix, the farther away it is from the maximally mixed state. Apart from distance measures generally being interesting, this particular measure of distance has additional relevance. In fact it is often impossible to generate a real pure state in the lab, typically there is some noise in every preparation procedure. Realistically all states used in experiments can at best be "close" to pure states. This unavoidable noise is generally integrated into the theoretical description. There are many possible ways to model noise, but the following is almost a canonical choice for modeling noise without any further assumptions in the quantum info community.

Definition 2.2.6. Let ρ be a quantum mechanical state. We model so-called **white noise**, through an application of the depolarizing channel,by

$$\Delta_p(\rho) := p\frac{1}{d^n}\mathbb{1} + (1-p)\rho, \tag{2.25}$$

where p is the **noise parameter** and $\Delta_p(\rho)$ is the **noisy state** of ρ.

This definition of noise means that either with probability $(1-p)$ the correct target state ρ is produced, or with probability p we produce the noise term $\frac{1}{d^n}\mathbb{1}$. White noise refers to noise, that is completely random. All measurements on $\frac{1}{d^n}\mathbb{1}$ have uniform probability for all possible outcomes in all possible bases. Thus they are maximally random and $\frac{1}{d^n}\mathbb{1}$ properly models white noise. It is of course possible to consider other probability distributions in the noise term (so-called "coloured noise"). For simplicities sake, we will not go any further into alternative noise models during this thesis.
In many protocols we want to employ pure states $\rho = |\psi\rangle\langle\psi|$. Every mixing is then due to error. Thus whenever we aim to prepare pure states, the purity of the state becomes a measure of quality of the experimental preparation procedure.

2.2.2 Entanglement & separability

The first section has covered the two opposed concepts of a *mixed* and a *pure* state.

The next section is devoted to one of the most important phenomena of quantum mechanics, that is not present in the classical description. Again it will be a story about complementary notions: *entangled states* and *separable* states. These notions basically clarify if a system is describable by two correlated local descriptions or not.

One word of caution is that both pairs of complements do not interfere with each other, since they are sometimes confused when encountering the topic for the first time. Mixed states may be either entangled or separable, just as pure states may be either entangled or separable.

Entangled systems can be used to outperform classical systems in certain tasks and it has been conjectured that entanglement is the defining resource for quantum advantages in computations. Even though this is not yet strictly proven many researchers believe this hypothesis to be true. It has been demonstrated, that the violation of Bell inequalities and therefore entanglement is a requirement for the reduction of communication complexity in quantum mechanical systems (see for example the communication section in [5]). Also entangled systems are at least known to be sufficient to achieve a quantum speedup for computations [39]. Naturally this makes the certification, quantification and classification of entangled systems be of utmost importance to the quantum information community. How to achieve these objectives is the main topic of the following section.

Definition 2.2.7. (i) We call a state ρ **k-separable** if we can write the states density matrix as a sum of tensor products

$$\rho_{k-separable} := \sum_{i=1}^{m} p_i \bigotimes_{j=1}^{k} \rho_{i_j}, \tag{2.26}$$

with ρ_{i_j} the density matrix of dimension d_{i_j}, p_i the coefficients of a convex combination and$\prod_{j=1}^{k} d_{i_j} = d$.

(ii) We call every state that is at least 2-separable more simply **separable**.

(iii) We call a non-separable state ρ **entangled**.

(iv) If ρ is not k-separable for any k, we call it **genuinely multipartite entangled**.

A state's mixedness is quantified satisfyingly by the purity in 2.24. It would be desirable to have a similar quantity in order to quantify entanglement: a so-called *entanglement measure*.

It turns out to be much more complicated to quantify entanglement than purity. As a result, there is an abundance of entanglement measures developed in the literature, but none has been singled out to be the one right measure to capture the notion of entanglement perfectly. Due to the vast body of available literature it is impossible to list all available entanglement measure exhaustively in this thesis, for detailed study of the topic consider this review article [15].

Another consequence of the difficult search for a satisfying quantification procedure, is that many authors working in entanglement theory have changed their strategy. Abandoning the hope to capture notions such as *more* or *less* entangled, many try to only answer the binary question whether there is entanglement present within a system or not. This is usually done through the identification and subsequent violation of necessary properties of separable systems, leading to the sub-field of: *entanglement detection criteria*.

The development of new entanglement criteria has been one of the main motivations throughout this thesis and later chapters will present several new entanglement detection procedures. This is why we make use of the remaining section to introduce entanglement theory. Even though our contributions are within the topic of entanglement criteria, we also briefly introduce entanglement measures to contrast it to our sub-field.

It is controversial, what may be the correct axiomatization for an entanglement measure. At least it was possible to identify certain minimal requirements for a measure to be called an entanglement measure, that are generally agreed upon. This is usually achieved by finding an order relation on what is believed to be informally classical operations. Before being able to define this relation, we have to formally capture that notion of classicality.

LOCC operations: A partial order for entanglement

Definition 2.2.8. (LOCC Operations)
Let us consider a n-partite system.

A **local operations and classical communication** (LOCC) operation on the n-parties is defined as all operations representable by the following recursion:

- For the start of the recursion we define the so set of one-round LOCC operations by

$$LOCC_1 = \bigotimes_{i=1}^{n} M^i U^i \tag{2.27}$$

  where U are unitaries and $M^1, \ldots, M^n$ measurements.

- The measurements M will have certain outputs. These are written into the vector $\vec{1} \in \mathbb{R}^n$. $\vec{1}$ contains the measurement results of the 1-th round of LOCC.

- Let U_k be a function depending on the up to $k-1$ vectors of earlier measurement results. U_k outputs a specific unitary U, chosen in accordance to $\vec{k} \ldots \vec{1}$ all prior measurement results.

- Let M_k^i be a function depending on the up to $k-1$ vectors of earlier measurement results. M_k^i outputs a specific measurement M_k^i, chosen in accordance to $\vec{k} \ldots \vec{1}$ all prior measurement results. The measurements $M_k^1, \ldots, M_k^n$ will have certain outputs. These are written again into the vector $\vec{k} \in \mathbb{R}^n$. $\vec{1}$ contains the measurement results of the 1-th round of LOCC.

- Finally we define the recursion for the set of k-round LOCC operations by

$$LOCC_k = \bigotimes_{i=1}^{n} M_{k-1}^i U_{k-1} LOCC_{k-1} \tag{2.28}$$

- The operations representable by arbitrary many k-round iterations of the above recursion are the LOCC operations.

We restrict ourselves to stating the definition and giving an intuitive motivation why this definition is sensible. The class of LOCC operations is one of the core constituents of quantum information theory and has been extensively studied. For a detailed discussion of the classes definition and properties refer to [7].

Let us now briefly comment on the above definition. All operations performable in quantum mechanics are either measurements or state transformations. Any operation that can be written as above consists entirely of a sequence of tensor product measurements and state transformation. That means that all these measurements and transformation can only performed on a single system each. Thus the tensor product is capturing the notion of *local operations*.

Now we have to emphasize that the M_k^i, U^i should not be considered as fixed matrices, but rather as being chosen from a described set of possible operations in dependence of the earlier measurement results. This is the mathematical description of *classical communication*. One may measure, obtain a result and then use the information obtained to choose a proper state transformation for an measurement in the next step.

The definition of LOCC may capture the essence of the problem, but it is cumbersome to work with. It is not easy to check through definition directly whether a state can be transformed

into another one by LOCC. There exists however an important technical tool at least in the pure bipartite case, that allows us to characterize the state transformations possible in LOCC: the *Nielsen majorization theorem* [38]. Before stating the theorem, we have to introduce the notion of majorization.

Definition 2.2.9. Let x,y be to vectors in $\mathbb{R}^n$. Let $0 \leq x_i, y_i \leq 1$ and $\sum_{i=1}^n x_i = 1, \sum_{i=1}^n y_i = 1$ so that both vectors entries describe probability distributions.

We write

$$x \prec y \tag{2.29}$$

and say y **majorizes** x, if one of the following equivalent statements is true

1. Let $x^{\downarrow}, y^{\downarrow}$ denote the the vectors x, y ordered descendingly. For all entry wise sums of xx and y we have the relation
$$\sum_i^k x_i^{\downarrow} \leq \sum_i^k y_i^{\downarrow} \forall k. \tag{2.30}$$

2. There exists
$$y = Mx, \tag{2.31}$$
where M is a **doubly stochastic matrix**. M is doubly stochastic iff $\sum_i M_{ij} = \sum_j M_{ij} = 1$

3. There exists a finite sequence of **t-transforms** $T_1, \ldots, T_m$, with a t transform being defined as
$$Tx = \begin{pmatrix} x_1 \\ \vdots \\ x_i \\ \vdots \\ (1-t)x_j \\ \vdots \\ x_n \end{pmatrix}, 0 \leq t \leq 1 \tag{2.32}$$
such that
$$y = T_1, \ldots, T_m x \tag{2.33}$$

The intuition behind majorization, is that it defines a partial order on all probability distributions. This order describes the mixing of the probability distribution. x majorizes y, means x is less mixed than y. Majorization relations between vectors are a well studied subject with many applications to statistics and the theory of doubly stochastic transformations. For an exhaustive treatment of majorization theory from a classical statistics viewpoint see the textbook [36].

Theorem 2.2.10. *(Nielsen majorization theorem) Let ψ and ϕ be bipartite pure states. Then the following statements are equivalent:*

- *there exists a LOCC transformation*
$$\psi \overset{LOCC}{\longrightarrow} \phi \tag{2.34}$$

- *For the spectrum* $\sigma(\mathrm{Tr}_A(\psi))$,$\sigma(\mathrm{Tr}_A(\phi))$ *of* $\mathrm{Tr}_A(\psi), \mathrm{Tr}_A(\phi)$ *we have the majorization relation*

$$\sigma(\mathrm{Tr}_A(\psi)) \prec \sigma(\mathrm{Tr}_A(\phi)) \tag{2.35}$$

Note that Nielsen's theorem makes a comment about the spectrum of the reduced density matrices. We will later see that these spectra are exactly the important Schmidt coefficients 2.2.16. The case of the remaining partial trace Tr_B *is analogous.*

Proof. For the proof refer to Nielsen's original paper [38] □

Clearly Nielsen's Theorem defines an order relation on the set of pure bipartite states. The relation describes, whether a state can be transformed into another state by LOCC. Entanglement is a non classical property and it can be shown, that it can not increase by application of an LOCC operation. This means, that by moving "down" the transformability relation we can either leave the entanglement of the state constant or decrease it.

Thereby LOCC gives us a notion of more and less entangled. The next step is to ask us, is there is such a thing as a *least* or most *entangled* state?

For the *least* entangled state, the answer is trivial. There exists a whole class of separable state without any entanglement, so the answer is affirmative but far from uniquely defined.

For the *most* entangled state, the answer is much more intricate.

First, we observe that the Nielsen Theorem reduces that question of transformability on pure bipartite systems to the question of majorizability. Now in the special case of majorization, clearly there exists a unique maximal element. This is due to

$$\begin{pmatrix} \frac{1}{d} \\ \frac{1}{d} \\ \vdots \\ \frac{1}{d} \end{pmatrix} \in \mathbb{R}^n \succ \forall x \in \mathbb{R}^n. \tag{2.36}$$

This motivates us to define the notion the maximally entangled state. For a bipartite system we can state the maximally entangled state.

Definition 2.2.11. We define the d-dimensional **maximally entangled state** short ϕ^+ state as:

$$|\phi^+\rangle\langle\phi^+| := \sum_{i,j=0}^{d-1} \frac{1}{d}(|ii\rangle\langle jj|) \tag{2.37}$$

We have to caution, the notion of the maximally entangled state breaks down in multipartite systems. One first indication of this is the fact that the Nielsen Majorization Theorem is only defined for bipartite systems. So the above approach to argue for existence and uniqueness of a maximally entangled state via majorization is not available anymore. In fact, already for the tripartite qubit case a counter-example is available due to [1]. The authors provide two multipartite states, that can not be inter-converted into each other by LOCC, thus destroying any hope of uniqueness in general.

There have been recent attempts to extend the notion of the maximally entangled state to a set of maximally entangled states [11]. There the authors conclude that the maximally entangled set is of measure zero, but non-trivial for 3 qubits. Already the 4 qubit case shows a somewhat pathological behavior. Almost all states are maximally entangled in a trivial sense, since no LOCC conversion is possible at all on most of the states. However they argue that operationally the notion of LOCC convertibility is important, thus one should focus on the subset of the maximally entangled set that is LOCC convertible. They conclude that the set of LOCC convertible states is again of measure zero and describable.

We go on to introduce and discuss two famous examples of non inter-convertible multipartite states:

Definition 2.2.12. We define the n-partite d-dimensional **Greenberger-Horne-Zeilinger** short $GHZ(n,d)$ state as:

$$|GHZ(n,d)\rangle\langle GHZ(n,d)| := \frac{1}{d}\sum_{i,j=0}^{d-1} |\underbrace{i\dots i}_{ntimes}\rangle\langle\underbrace{j\dots j}_{ntimes}| \tag{2.38}$$

Definition 2.2.13. We define the n-partite **W state** as:

$$|W(n)\rangle\langle W(n)| := \frac{1}{n}\sum_{i,j=0}^{n-1} |0\dots\underbrace{1}_{i-thposition}\dots 0\rangle\langle 0\dots\underbrace{1}_{j-thposition}\dots 0| \tag{2.39}$$

We remark that we have included particle number and local dimension into the notation of the state. The bracket containing n or d will be sometimes dropped throughout the text of this thesis when this information is irrelevant or apparent from the context.

In [12] the case of 3 pure qubits has first been fully classified. Later on the more general case of 3 mixed qubits could be classified in [1]. These papers introduce the concept of classes of states that are connected by LOCC to either the W or GHZ state. Both classes of states differ considerably, thus signifying that finer classifications of the state space than entanglement or separability are in fact very relevant. Without going into much detail we want to exemplify their difference by one fundamental property: the *robustness* towards partial traces.

Observation 2.2.14. *Compare the result of tracing out one arbitrary single system X in different states.*

- *GHZ-state:*

$$\mathrm{Tr}_X(|GHZ(d,n)\rangle\langle GHZ(d,n)|) = \frac{1}{d}\mathrm{Tr}_X(\sum_{i,j=0}^{d-1}|i\rangle^{\otimes n}\langle j|^{\otimes n}) = \frac{1}{d}\sum_{i,j=0}^{d-1}|i\rangle^{\otimes n-1}\langle i|^{\otimes n-1} \tag{2.40}$$

  The result looks almost like an identity, however not all diagonal elements are one . It is a diagonal matrix with only d-nonzero entries all equal to $\frac{1}{d}$. All offdiagonal elements have been canceled by the partial trace.

  This means in the GHZ case that the erroneous loss of information about only one of potentially many systems is sufficient to leave us only with some diagonal state, that is not inside of

the form $|GHZ(d,n-1)\rangle\langle GHZ(d,n-1)|$. Even worse the resulting state is clearly separable! In that sense it can be said, that the GHZ class is not robust under particle loss.

- *W-state:*

$$\mathrm{Tr}_X(|W(n)\rangle\langle W(n)|) = \underbrace{\langle 0|_X W(n)\rangle\langle W(n)|0\rangle_X}_{\frac{n-1}{n}|W(n-1)\rangle\langle W(n-1)|} + \underbrace{\langle 1|_X W(n)\rangle\langle W(n)|1\rangle_X}_{\frac{1}{n}|0\ldots0\rangle\langle 0\ldots0|} \tag{2.41}$$

The W case is quite robust due to its combinatorics. The partial trace is an operator acting from the left and the right side simultaneously, thus only basis elements with an equal parity in the traced over system survive the mapping. The distribution of the ones and zeros of the W-state however differs a lot. Almost all basis entries are zeros. Regardless of particle number and the specific partial trace operator under consideration there, it is never possible to have more than one coefficient of 1's with equal parity. This is why only a tiny fraction of diagonal noise s generated, while taking the partial trace.

In contrast tracing out one system from $|W(n)\rangle$, maps a very large fraction to the corresponding $|W(n-1)\rangle$ with one less particle. Summing over $|0\rangle$ leaves $(n-1)$ nonzero coefficients. Also taking the partial traces, repeatedly is achieved very easy in an inductive manner only generating little noise in each step.

Even if the term robustness carries a positive connotation, whether this property is desirable or not is entirely up to the operational task. We have observed that W is robust under partial trace, while GHZ is not. Still, this does not allow us to judge one state to be generally more useful than the other. In many tasks like computation or communication the robustness of the W class would lead to a better error resistance, in cryptographic applications the fragility of GHZ turns instead into security because partial information gain of an eavesdropper is guaranteed to not yield any information. Maybe even more unexpectedly, it turns out that the robustness of the GHZ allows us to maximize analytically certain information theoretic quantities. This makes it of important theoretical use in information theory in general, but also more specifically and (more surprisingly!) in quantum thermodynamics.

We leave the discussion on the relevance of the LOCC classes with the concluding remark, that all of these classes seem to have benefits and disadvantages and it is unlikely to encounter one class that fits all needs. For this reason a further classification of multipartite systems seem desirable, if hard.

Detecting entanglement: Measures, witnesses & criteria

Finally, we have introduced the language to formulate the defining property of entanglement measures.

Definition 2.2.15. An **entanglement measure** $\mu(\rho)$ is a mapping: from the space of all density matrices $\mu : \rho \to \mathbb{R}$ that fulfills the following properties:

- μ is non-positive on the set of separable states
- μ is non-increasing under LOCC-operations.

As a first example of an entanglement measure, we introduce the so-called Schmidt Number. Prior to introducing the Schmidt Number, we have to introduce the well-known Schmidt decomposition.

Lemma 2.2.16. *A pure state $|\psi\rangle_{AB} \in \mathcal{H}_A \otimes \mathcal{H}_B$ can always be decomposed into its Schmidt decomposition by*

$$|\psi_{AB}\rangle = \sum_k \sqrt{\lambda_k}|i_k\rangle \otimes |j_k\rangle. \tag{2.42}$$

The marginals are expressible as

$$|\psi_A\rangle\langle\psi_A| = \mathrm{Tr}_B(|\psi_{AB}\rangle\langle\psi_{AB}|) = \sum_k \lambda_k |i_k\rangle\langle i_k| \tag{2.43}$$

or

$$|\psi_B\rangle\langle\psi_B| = \mathrm{Tr}_A(|\psi_{AB}\rangle\langle\psi_{AB}|) = \sum_k \lambda_k |j_k\rangle\langle j_k| \tag{2.44}$$

*The coefficients $\sqrt{\lambda_i}$ are named the corresponding **Schmidt coefficients**.*

Proof. First, we express each reduced system in an eigenvalue basis

$$\rho_B = \mathrm{Tr}_B(|\psi_{AB}\rangle\langle\psi_{AB}|) = \sum_k \lambda_k |i_k\rangle\langle i_k|, \rho_A = \mathrm{Tr}_A(|\psi_{AB}\rangle\langle\psi_{AB}|) = \sum_{k'} \lambda'_{k'} |j_{k'}\rangle\langle j_{k'}|. \tag{2.45}$$

This is of course always possible, it is however not initially clear that number,multiplicity and value of the coefficents $\lambda_k, \lambda'_{k'}$ coincide. Still, we are free to express the parent state in a tensor product basis of the marginals bases

$$|\psi_{AB}\rangle\langle\psi_{AB}| = \sum_{k,k'} \langle i_k(|\otimes\psi_{AB}\langle j_{k'}\rangle|i_k\rangle)^2 \langle i_k| \otimes |j_{k'}\rangle\langle j_{k'}| = \sum_{k,k'} c_{k,k'} |i_k\rangle\langle i_k| \otimes |j_{k'}\rangle\langle j_{k'}|. \tag{2.46}$$

Now by matching the above two expression we see

$$\sum_{k'} \lambda'_{k'} |j_{k'}\rangle\langle j_{k'}| = \rho_B = \mathrm{Tr}(\rho_{AB}) = \sum_{k,k'} c_{k,k'} \underbrace{\mathrm{Tr}_A(|i_k\rangle\langle i_k| \otimes |j_{k'}\rangle\langle j_{k'}|)}_{|j_{k'}\rangle\langle j_{k'}|} \tag{2.47}$$

and compare it to the other marginal

$$\sum_k \lambda_k |i_k\rangle\langle i_k| = \rho_A = \mathrm{Tr}(\rho_{AB}) = \sum_{k,k'} c_{k,k'} \underbrace{\mathrm{Tr}_B(|i_k\rangle\langle i_k| \otimes |j_{k'}\rangle\langle j_{k'}|)}_{|i_k\rangle\langle i_k|}. \tag{2.48}$$

Showing us that the sum of the coefficients have to coincide. □

We introduce one common example of an entanglement measure:

Definition 2.2.17. Let $|\psi\rangle\langle\psi|_{AB}$ be a pure state of an bipartite system.

$$r(|\psi\rangle\langle\psi|_{AB}) = \text{rank}(\text{Tr}_A(|\psi\rangle\langle\psi|_{AB})) = \text{rank}(\text{Tr}_B(|\psi\rangle\langle\psi|_{AB})) \tag{2.49}$$

The **Schmidt rank** of $|\psi\rangle\langle\psi|$ is the rank of its marginals.
(Note that due to the properties of the Schmidt decomposition, it is clear that the ranks of both matrices are equal!)

The Schmidt rank is probably one of the most famous entanglement measures, it is however limited to the rather restricted case of pure states. This is already unsatisfying, but there are other complications for example the lack of any corresponding notion for tri- or d-partitions. The restriction to the pure case can be remedied somewhat.

Definition 2.2.18. Let ρ be a (potentially mixed) state.
Then we define the **Schmidt number** $r(.)$ by

$$r(\rho) := \min_{\rho=\sum_{p_i|\psi_i\rangle\langle\psi_i|}} \max r(|\psi_i\rangle\langle\psi_i|), \tag{2.50}$$

where we minimize over all possible decompositions of ρ into pure states.

Note that ρ is decomposed into pure states $|\psi_i\rangle\langle\psi_i|$. On pure states we have defined the notion of Schmidt rank. $r(\rho)$ is the maximal Schmidt rank appearing in a decomposition, that minimize this number over all possible decompositions of ρ. For other examples of entanglement measures or an in-deptht discussion of the relevant theory refer to the exhaustive review article [15].

As usual whenever generalizing a pure state concept, we can consider the generalization by pure state decomposition. This approach is theoretically always possible, but the pure state decomposition is frequently incalculable. So even though we can define the Schmidt number, it is not always clear how to calculate it. Unfortunately, this is an archetypical problem arising when considering entanglement measures. These problems are not easy to circumvent, since the general entanglement detection problem is well-known to be NP-hard [21], we can be almost sure that there will not be one measure suitable for all states.

As stated above, the search for these entanglement measures is neither easy nor the focus of this work. One usual mathematical strategy, when encountering questions that are deemed too hard to tame is relaxation. So, probably we would like to drop some of the requirements of the entanglement measure if we want something easier to solve.

Since entanglement quantification is a hard task, that may also require large parts of the density matrix to be known we will content ourselves with the task of finding out whether they are non-zero. This leads to the concepts of entanglement witnesses, where entanglement is certified in a binary way. An entanglement witness may either show a state to be entangled or not, but without making any statement about how much entanglement is present within a system.

We start by a slightly more general consideration: what is a detection criterion?

In a way, when only trying to the retain the first defining property, what we are left with is already almost a criterion. Even more generally, we could consider different sets in the first definition, yielding a rather general formulation of a (general, not necessarily an entanglement) criterion.

The standard approach to detect entanglement, whenever an entanglement measure is deemed too complicated fo find, are entanglement witnesses. The large subject of entanglement witnesses is covered by several review articles (c.f. [20, 26]).

Definition 2.2.19. An **entanglement witness** W of a state ρ is required to be

1. a linear operator
2. $\langle W\rho_{ent}\rangle < 0$ for an entangled state ρ_{ent}
3. $\langle W\rho_{sep}\rangle \geq 0$ for all separable states ρ_{sep}

Even though the naming convention does not contain the word linear, it is generally considered part of the definition of an entanglement witnesses. Even though a large part of this thesis (specifically:4) is devoted to entanglement detection, it is not strictly covering entanglement witnesses in the above sense. Many of the detection methods we will study are not linear. Some authors would argue that these are by definition no witnesses!

For this reason, we introduce an even further relaxed notion:

Definition 2.2.20. An **entanglement detection criteria** $\mathcal{E}$ is a function of a state ρ is required to be

1. $\mathcal{E}(\rho) > 0$ for an entangled state ρ_{ent}
2. $\mathcal{E}(\rho_{sep}) \leq 0$ for all separable states ρ_{sep}

These detection criteria will be our main tool to certify entanglement. Note that the inequalities in 2.2.19 and 2.2.20 are reversed. We employ a different convention for entanglement criteria, because frequently the amount of a criterion's violation can be regarded as a type of robustness. We prefer to write this robustness as a positive quantity.

It can also be interesting to study a more specific notion of entanglement, such as non k-separability. Non k-separability detection criteria can be defined completely analogously. In any case, the question how to detect properties of various multipartite system is of great interest.We will address this question in Chapter 4.

2.3 Thermodynamic notions

This chapter is devoted to a quick introduction into quantum thermodynamics. Classical thermodynamics can be seen as the study of concepts such as *temperature*, *work*, *entropy* and *energy*.

Quantum thermodynamics has recently been a very active field. It tries to interpret notions derived from the classical thermodynamics in a quantum mechanical context. Frequently, also quantum informational notions are exported back into the thermodynamical description. This allows for an analysis of relations of concepts, that are usually not studied since they stem from different theoretical backgrounds such as entanglement and *temperature* or *work*. Entanglement is a notion totally absent from thermodynamics, in contrast *temperature* is an inherently statistical concept that is in many ways undefined in common quantum mechanical scenarios where only few particle interact. Especially at the level of single particle the notion of temperature seems to completely

break down. A pertinent question would be: *What is the temperature of a qubit?*

Other quantities are well-known in both theories, but the description and interpretation may vary vastly. It has been a celebrated result of statistical physics, that *entropy* appears as well in a thermodynamical context as well as in information theoretic questions. So it is no surprise to find entropy be one of the fundamental tools in quantum thermodynamics, that allows to connect the different theories. Finally, no branch of physics can be thought without the concept of *energy* and again energy appears in various formulations.

These questions on the interface of distinct and different theoretical models are very appealing, but have also led to a heated debate within the community how to answer them properly. For this reason we do want to point out, that the definitions we give for all these quantities seem useful to us. We claim in no way that they may be the only or best way to deal with these notions. For a more complete overview of the recent developments in the field of quantum thermodynamics please refer to the review [18].

Definition 2.3.1. We call the **entropy** of a state ρ the function

$$S(\rho) := -\operatorname{Tr}(\rho \log(\rho)) \tag{2.51}$$

We go on to address one of the defining notions of thermodynamics: *the temperature*.

Definition 2.3.2. A **thermal state** at Hamiltonian H is

$$\tau(\beta) := \frac{e^{-\beta H}}{\operatorname{Tr}(e^{-\beta H})} = \frac{e^{-\beta H}}{Z}. \tag{2.52}$$

We call β the **inverse temperature** and write in analogy with classical thermodynamics $\beta = \frac{1}{k_B T}$ with T the **temperature** and k_B the **Boltzmann constant**.
We call Z in a further parallel to classical thermodynamics the **partition function**.

It is quite remarkable, how temperature enters the picture here! First, we define the class of thermal states. The thermal states are parametrized by a single parameter. This parameter can be calculated by Jaynes principle and will then simply be defined to be the temperature. Only as an after thought, we introduce the temperature as some kind of order parameter of this class of states. On the other hand, it is remarkable that we can even find any equivalent concept of temperature, since quantum mechanics is a microscopic theory. Temperature usually only arises as an emergent property, thus it is sensible that it can only indirectly appear in a quantum mechanical treatment.

In fact, as familiar as it may seem from everyday life temperature is far from an unproblematic quantity even in the traditional setting. Temperature is a concept that is only defined for equilibrated systems, while in the non stationary case temperature is considered to be undefined. This makes it even more remarkable that a sensible quantum mechanical description can be possible.

Definition 2.3.3. (Energy) Let H be the Hamiltonian of system ρ. Consider the following two different notions of energy:

- We define the average energy of a ρ as

$$E(\rho) := \langle H \rangle_\rho = \operatorname{Tr}(H\rho) \tag{2.53}$$

- We define the free energy as

$$F_\beta(\rho) := \langle H \rangle_\rho - \frac{1}{\beta} S(\rho) = \langle H \rangle_\rho - k_B T_E S(\rho), \tag{2.54}$$

  where k_B labels Boltzmann's constant and T_E is the ambient temperature (the temperature of the environment of ρ).

Finally we define the work cost of a unitary U on a state by

$$W(\rho, U) = \langle H \rangle_{U\rho U^\dagger} - \langle H \rangle_\rho \tag{2.55}$$

Appreciate the similarity of the free energy to the classical case. There we have in complete analogy the relation

$$F_{classical} = E_{classical} - T S_{classical}. \tag{2.56}$$

Here we obtained the quantum mechanical F 2.54 from replacing the internal energy $E_{classical}$[1] by the average energy 2.53, the temperature T is in both approaches simply a scalar and we have replaced the classical notion of the Entropy by the von Neumann entropy. We will not work directly with the free energy within the scope of this manuscript, however we introduced it to motivate the use of our definition of work as 2.55. We will generally use this quantity instead of the free energy.

Please, appreciate that this is a sensible choice due to

$$\begin{aligned} W &= E(\rho) - E(U\rho U^\dagger) = \Delta E \stackrel{!}{=} \Delta F = \\ &= F(\rho) - F(U\rho U^\dagger) = \langle H \rangle_\rho - \frac{1}{\beta} S(\rho) - \langle H \rangle_{U\rho U^\dagger} + \frac{1}{\beta} S(\rho), \end{aligned} \tag{2.57}$$

showing that the above notion of work actually coincides with a free energy difference before and after a unitary transformation. It is therefore well grounded to use the quantity we call "work" 2.55 instead of difference of the free energy 2.54.

2.4 Fock space and the creation & annihilation operators

Typically Quantum Information concerns itself with studying discrete Hilbert spaces, such as the paradigmatic qubit or qudit. This is somewhat a departure from a more classical approach towards quantum mechanics, that had been frequently taught in the language of differential equations and functions satisfying those. In many parts of quantum mechanics the continuous variable viewpoint is still prevalent. Especially in the experimentally very important field of a quantum optics continuous objects are still used on a day to day basis. Even though all of those theories have to reach the same predictions, the mathematics involved can be quite different. Furthermore in one of our publications for this thesis [2] we discuss a finite dimensional basis together with its continuous variable limit. Due to this reason we introduce the language of the ladder operators, that provide a beautiful link between continuous functions, that constitute its elementary parts, and a countable infinite algebra, that yields the quantized behavior of the energy levels characteristic for quantum

[1] Frequently the internal energy is denoted by U in classical thermodynamics, we have refrained from using this convention due to the frequent use of U to denote a unitary in quantum information.

mechanics.

The basic idea behind the introduction is to find families of continuous functions, that obey two of the fundamental statements of quantum mechanics:

- First, that **energy is quantized**, hence appears in levels of finite differences.
- Second, that **energy is bounded from below**, hence there is one state of lowest energy: *the ground state*.

In order to speak of energy, we have to introduce an Hamiltonian that defines it. Along with every Hamiltonian comes a notion of energy and it from that follows a possible definition of the ladder Operators. In fact ladder operators can be defined for a wealth of differing Hamiltonians, each gives rise to a slightly different ground state that will be starting point to define all the ladder operators. We will restate the well-known general calculus of the ladder operators as well as state the explicit example of the best studied system: the harmonic oscillator.

We first specify the ground state (of the harmonic oscillator).

Definition 2.4.1. (Vacuum state)
We define the **vacuum state** as the lowest ground state of the Schrödinger equation with potential $V = \frac{m\omega^2}{2}x^2$. In the important case of the harmonic oscillator it can be written as

$$|0\rangle := ce^{-\frac{m\omega x^2}{2\hbar}} \tag{2.58}$$

with a suitable normalization constant $c := (\frac{m\omega}{\pi\hbar})^{\frac{1}{4}}$.

In a second step we specify the operations that will allow us to switch from one energy level to the other.

Definition 2.4.2. (The creation operator & the annihilation operator) We define:

(i) the **annihilation operator** a via the relation

$$a|0\rangle = 0 \tag{2.59}$$

and

$$a|n\rangle = \sqrt{n}|n-1\rangle. \tag{2.60}$$

(ii) We define the **creation operator** $a^\dagger$ via the relation

$$a^\dagger|n\rangle = \sqrt{n+1}|n+1\rangle \tag{2.61}$$

(iii) We define the **number operator** N via the relation

$$N := a^\dagger a \tag{2.62}$$

That these relations are achieved can be again seen by the example of the harmonic oscillator. There the choice

$$a = \sqrt{\frac{m\omega}{2\hbar}}(x + \frac{i}{m\omega}(-i\hbar\frac{d}{dx})), a^\dagger = \sqrt{\frac{m\omega}{2\hbar}}(x - \frac{i}{m\omega}(-i\hbar\frac{d}{dx})) \tag{2.63}$$

implements the Hamiltonian of the harmonic oscillator by

$$H = \hbar\omega(a^\dagger a + \frac{1}{2}) = -\frac{\hbar^2}{2m}\frac{d^2}{dx^2} + \frac{m\omega^2}{2}x^2. \tag{2.64}$$

Observation 2.4.3. *The creation & annihilation operators observe the so-called canonical commutation relations.*

$$[a, a^\dagger] = 1 \tag{2.65}$$

Once we have defined both a notion of a ground state and the prescription of switching to all its excited states, the definition of those modes follows logically and provides a general description.

Definition 2.4.4. (Number state) We define the n-th number State as

$$|n\rangle := \frac{(a^\dagger)^n}{\sqrt{n!}}|0\rangle \tag{2.66}$$

Observation 2.4.5. *Note that the number operator N gives the number of the number state*

$$N|n\rangle = n|n\rangle, \tag{2.67}$$

where N is an operator and n a scalar. Thus the above can also be seen as an eigenvalue problem.

Chapter 3

Correlations in the Bloch Picture

3.1 From the single qubit towards qudit systems: How to generalize the Bloch vector decomposition?

One of our main tools in the derivation of entanglement witnesses will be the Bloch vector decomposition of a physical state. We demonstrate that entanglement detection allows at least for some new applications of the well-known concept of the Bloch vector. Therefore we devote the following section to the introduction and review of the concept.

Traditionally the Bloch vector decomposition is a concept, that is applied to two dimensional single particle systems. In fact, even the traditional qubit case is usually not defined very formal, since due to the uniqueness of the Pauli basis, it may not seem necessary to elaborate too much on its definition. Mostly authors simply give the equation $\rho = \frac{1}{d}(\mathbb{1} + \sum_{i=1}^{d^2-1} c_i \sigma_i)$ as a definition, without explicitly mentioning what properties of the Pauli basis are the defining features of the definition. We will review the standard case and by highlighting the advantages of the Bloch vector in this simple setting in the first part of this Section3.1.1.Then we go on to establish a set of features that may be desirable to be kept when considering more general systems than the single qubit.

Next we treat the description of a single qudit in Section 3.1.2. Already for higher dimensional single particle systems, it is not possible anymore to talk of one uniquely defined Bloch vector decomposition. A lot of work has already been done on possible generalizations of the qubit Bloch vector formalism [31, 32, 3, 33, 34]. The qudit case is considerably more complicated due to complicated physicality constraints. As a consequence of this, certain properties of the formalism break down, if extended to a qudit system. The intuitive correspondence between the set of physical states and a generalized Bloch vector is no longer granted. This may lessen the pedagogical appeal of the formalism, but also more concrete problems follow. For example higher dimensional Bloch balls are not anymore closed under rotations.Nonetheless, we believe that the formalism has merits in the even more general case of n partite qudit systems, especially in the context of entanglement detection.

The treatment of the single qudit already shows all the fundamental problems arising, when

searching for a general Bloch vector description of qudit systems. In Section 3.1.2 we concern ourselves with the question, which properties of the Pauli basis can be generalized. We also try to justify why they should be generalized. In order to get a meaningful description for higher dimensional systems, the main challenge is to find a suitable set of SU(d) generators representing a given qudit. This is a well developed topic. With the generalized Gell-Mann bases a canonical choice exists (c.f. Section 3.2.1) alongside well studied alternative descriptions such as the Heisenberg-Weyl basis (c.f. Section 3.2.2). In principle many different bases can be constructed. Even though all of them are in some sense equivalent, it is by no means a straightforward task to single out descriptions that are deemed favorable. Since this is a diverse field, we devote the whole following section 3.2 to the comparison of different bases that may be used to define generalized Bloch vector decompositions. Alongside reviewing the well-known concepts, we also discuss an alternative to the generalized Gellman matrices: the symmetrized Heisenberg-Weyl operators (c.f. Section 3.2.3). These operators can be seen as a fresh viewpoint on how to generalize the Pauli matrices, that additionally considers properties like anti-commutativity and infinite dimensional scaling, besides the traditional focus on either Hermiticity or unitarity. It has been proposed in one of our publications [2] that led to this thesis and constitutes original work.

In Section 3.1.3 we end this section by considering composite systems. In contrast to the challenges faced while finding a suitable description for larger dimensional systems, studying large numbers of particles is a relatively straightforward affair. The Bloch picture can be easily extended via tensor products and the treatment of a composite system about as complicated as the treatment of its local constituents. In essence once a description for the single qudit is fixed, there is no freedom left when describing a system of qudits. Theoretically no real challenges are to be tackled, the only minor complications of multipartite systems are frequent combinatorial considerations and the required more cumbersome multi-index notation.

3.1.1 Single particle systems: the qubit case

The Bloch vector decomposition via Pauli matrices

Definition 3.1.1. Let ρ be a qubit density matrix. We call

$$\rho = \frac{1}{2}(\mathbb{1} + \sum_{i=1}^{3} c_i \sigma_i), \tag{3.1}$$

the **Bloch vector decomposition** [4] of ρ, where σ_i denotes the Pauli operators,

$$\sigma_x := \sigma_1 := \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \sigma_y := \sigma_2 := \begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix}, \sigma_z := \sigma_3 := \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{3.2}$$

We call $\vec{c} := (c_1, c_2, c_3)^t$ the **Bloch vector** [4] of ρ.

Expressing a qubit in the Bloch vector decomposition simply means to express it in terms of the basis of the identity and the Pauli operators. The arising coefficients of this basis representation are collected into a vector $\vec{c}$ and called the Bloch vector of ρ. Note that due to $c_i = \langle \sigma_i \rangle$ the coefficients are exactly the expectation values of a measurement in all non identity basis elements

of ρ. $\mathrm{Tr}(\rho^2) \leq 1$ enforces $\|\vec{c}\| \leq 1$.

It can be easily be seen that the Pauli operators are

- orthogonal,
- traceless,
- Hermitian,
- unitary,
- anti-commuting.

It is quite remarkable that the Pauli operators fulfill all of the above properties at once. This is in fact only possible in $d = 2$ dimensional spaces, larger spaces do not allow bases that fulfill all these properties at once. Arguably, the fact that $d = 2$ admits such a well behaved basis, is one of the key reasons why treating qubits is generally a significantly easier task than treating qudits.

Geometrical interpretation of the Bloch ball

Frequently the Bloch vector is used as a visualization aid to imagine the set of all physical states of a single qubit system. The idea is that the physical qubit state space is parametrized by an orthogonal basis and mapped one to one onto a ball.

Definition 3.1.2. We define the mapping

$$B : \rho \mapsto \vec{c}. \tag{3.3}$$

In the two-dimensional case the image of B on all states is the **Bloch ball**. In contrast to higher dimensions where the image of B is no longer a ball. There it will be a smaller convex subset of the hyper-ball.

It is obvious from $\mathrm{Tr}(\rho^2) \leq 1$ that $|\vec{c}| \leq 1$. Qubits exhibit three degrees of freedom, allowing us to work conveniently in $\mathbb{R}^3$, thus the above mapping really maps the state space to a ball for $\dim(\rho) = 2$.

These notable geometrical facts are known about the Bloch ball:

Lemma 3.1.3. *Let $\vec{c} \in \mathbb{R}^3, \|c\|_2^2 \leq 1$ be a Bloch vector in the Bloch ball.*

(i) The central point of the ball is the maximally mixed state.

(ii) All vectors within the Bloch ball correspond to physical states

(iii) The surface of the ball consists of pure states, while the interior of the ball consist of mixed states.

(iv) Convex combinations of points are equivalent to mixing the respective states.

(v) Antipodal points on the ball correspond to orthogonal states.

(vi) Applications of unitaries correspond to rotations of the ball. A rotation of a point on the ball can never leave the ball. Therefore a unitary can never map a state out of the physical state space.

(vii) Measurements can also be represented by vectors in the Bloch ball, just as states. They only differ in slightly different constraints.

Proof. (i) Follows from the definition

(ii) All Pauli matrices are Hermitian and normalized, any convex combination inherits these properties. The only non trivial thing to check for physicality is whether the resulting state is again positive. Considering the eigenvalue equation

$$0 = \det(\rho - \lambda \mathbb{1}) = \det \begin{pmatrix} (1-\lambda)+c_3 & c_1 + ic_2 \\ c_1 - ic_2 & (1-\lambda)-c_3 \end{pmatrix} = ((1-\lambda)^2 - c_3^2) - (c_1^2 + c_2^2)) = (1-\lambda)^2 - \|\vec{c}\|^2 ,$$

we see that for all $\|\vec{c}\|^2 \leq 1$ the spectrum always is positive and thus physical.

(iii) The purity requires: $1 = \mathrm{Tr}(\rho^2) = \frac{1}{2}(1 + \|c\|_2^2)$, clearly this holds if and only if ρ lies on the boundary of the ball.

(iv) Simply by writing out we see

$$p\psi + (1-p)\phi = p\frac{1}{2}(\mathbb{1} + \vec{c}_\psi \vec{\sigma}) + (1-p)\frac{1}{2}(\mathbb{1} + \vec{c}_\phi \vec{\sigma}) = \frac{1}{2}(\mathbb{1} + (p\vec{c_\psi} + (1-p)\vec{c}_\phi)\vec{\sigma}). \tag{3.4}$$

(v) Let p be point on the Bloch sphere, described by the Bloch vector $\vec{c}$. Its antipodal point is $p_A = -\vec{c}$. Now each of these points corresponds to a state. We want to show, that the states corresponding to p and p_A are orthogonal. Since the scalar product is invariant under rotations, we can w.l.o.g. rotate $p = \vec{c}$ to $(1,0,0)$ and its antipodal point $p_a = -\vec{c}$ to $(-1,0,0)$. These points correspond to the states $\frac{1}{2}(\mathbb{1} \pm \sigma_x)$. It can be easily checked that they are orthogonal.

(vi) U can be expressed in terms of Pauli operators. Due to the orthogonality of Pauli operators we have $U\vec{c}\vec{\sigma} = \vec{c}'\vec{\sigma}$, this leaves us with $U\rho U^\dagger = \frac{1}{2}(\mathbb{1} + \vec{c}U\sigma U^\dagger) = \frac{1}{2}(\mathbb{1} + \vec{c}'\sigma)$ demonstrating that the Bloch representation is closed under unitaries. Since all points that are representable by the Bloch representation lie on the sphere, they have to be connected by a rotation. In $d = 2$ the physical state space is the whole ball, so all rotations correspond to physical operations.

(vii) This is obvious, since the Pauli operators are a complete basis.

□

It is possible to define similar Bloch hyper-balls for qudits, but the situation is more complicated as will be addressed in 3.1.2. This will mainly be due to the fact that there is not a higher dimensional analogue of (*ii*). In the qubit case the Bloch mapping B is bijective, while in general B^{-1} may lead out of the state space for some $\vec{c}$ contained in the hyper-ball. As a consequence also vi in 3.1.3 breaks down for qudits.

3.1.2 Single particle systems: the qudit case

The special case of the single qutrit

Our aim in this work is to stay general when treating dimension and particle number of the systems we study. Therefore we tend to avoid working out special cases in detail. Of course sometimes particular examples may be beneficial for getting some intuition about the problem, therefore we want to briefly refer to existing work that specifically works out the positivity and geometrical aspects for fixed dimensions.

Several authors have studied the more specific qutrit case, that already exhibits all problems encountered for the fully general case. However the positivity constraints of the qutrit remain somewhat accessible and allowed them to find for example appealing visualizations as "Bloch ellipsoids" [34] or multiple Bloch balls [33], as well as comparably simple positivity constraints. Goyal et al. [19] give a case study with very detailed analysis of the sections in d=3, the work contains a classification of all possible sections of the qutrit state space, including plots of the sections. We are not aware of any explicit case studies for systems beyond the qutrit case up-to-date.

The case of the single general qudit

We give a quick review of the general single qudit case, before considering the general n qudit system. This section is inspired by the work of Bertlmann & Krammer [3], delivering a well readable exposition of the single qudit Bloch vector formalism in comparison to the qubit case. Alternatively, the reader may refer to any standard textbook about quantum mechanics.

A density matrix is like any matrix representable in many different bases. It is usually desirable to identify certain bases, that are well suited for the problem to be studied. Let us introduce the Bloch vector decomposition, that is one such Basis decomposition of a general matrix, that has been of great use in quantum mechanics.

Definition 3.1.4. Let ρ be a qudit density matrix. If ρ is expressed in terms of a basis

- containing the identity $\lambda_0 := \mathbb{1}$ as well as,
- $d^2 - 1$ orthogonal and traceless basis elements λ_i,

we call

$$\rho = \frac{1}{d}(\mathbb{1} + \sum_{i=1}^{d^2-1} c_i \lambda_i) \tag{3.5}$$

the **(generalized) Bloch vector decomposition** of ρ, with the so-called Bloch vector c and **Bloch vector** coefficients c_i.

In the case of $d = 2$ we obtain the classical Bloch vector decomposition. Here the orthogonal, traceless basis operators λ_i coincide with the well-known Pauli operators $\sigma_X, \sigma_Y, \sigma_Z$. In addition they exhibit the desirable property of being Hermitian and are uniquely defined up to a unitary.

Physicality conditions

Due to the fact that $\mathrm{Tr}(\rho^2) \leq 1$ and since the Bloch vector is 3-dimensional for qubits, it can be interpreted as the parametrization of a ball. In the qubit case every point contained in the Bloch ball represents a physical state. Furthermore the pure states are exactly the corresponding sphere, while the mixed states lie strictly in the interior of the Bloch ball.

In higher dimensions, the condition $\mathrm{Tr}(\rho^2) \leq 1$ still holds thus leaving us with a higher dimensional ball. However not all points belonging to this generalized Bloch ball are still physical states due to a possible violation of positivity. This means that in the general case the purity condition and the positivity do not coincide anymore. Furthermore, the question of how to properly choose a basis arises in higher dimensions. While in $d = 3$ the Gell-Mann operators are the standard choice, there are several known families of operators that might be suitable candidates. We elaborate on possible choices of basis in Section 3.2.

Let us consider the above conditions expressed as conditions of Bloch vector coefficients.

A Bloch vector $\vec{c}$ corresponds to a physical state $\rho = \frac{1}{d}(\mathbb{1} + \sum_{i=1}^{d^2-1} c_i \lambda_i)$ if $-1 \leq c_i \leq 1$. It is the $d = 2$ positivity condition $\sum_{i=0}^{3} \sigma_i^2 \leq 1$, that is hard to generalize to higher dimensions.

In $d = 3$ there are some successful parametrizations of the positivity (see: [33, 34]). In both papers, the author(s) present a similar chain of arguments. The first insight is in principle completely dimension independent, they use the fact that a Hermitian matrix (such as a states density matrix) is Hermitian if and only if its principal subminors determinants are positive. This result for the special case of the qutrit is one of the few explicit positivity constraints available beyond the qubit case.

It is however already evident why this approach will be complicated to generalize to arbitrary dimensions. Only in the qutrit case we have an easy direct expression for the three 2×2 subminors. Since the approach relies on constraints of determinants, we would have to use iteratively the determinant expansion formula in order to express a $d \times d$ subminor and derive positivity constraints. Furthermore the number of the subminors grows unfavorably with larger dimension. Even though one might consider continuing this approach in a tour de force for the ququart, it is quite clear that high dimensional systems expressed in this way will certainly lead to almost unhandleable bulky positivity constraints.

After choosing a particular basis for representation, they invoke the interlacing theorem to argue the positivity of all the principal sub-minors.

Consequences of positivity: Similarities and differences to the qubit

We explicitly state what properties of the qubit Bloch ball are generalizable and which are not, confer the lemma 3.1.3 where we state all properties relevant for this section.

Some of the geometrical properties of the qubit Bloch ball are conserved generally. Unfortunately arguably the two most important ones are lost: Property (ii) the easy *characterization of the physical state space* as well property (vi) the Bloch ball's *closedness under rotations.*

A unitary operation can always be interpreted as a rotation on the state space, regardless whether we study a three dimensional ball or a higher dimensional object. However only in three

dimensions the space state has the shape of a ball, generally it is a strictly smaller convex set. Even worse we know that all higher dimensional state spaces coincide on some points with the ball, but dont coincide on other points, thus ruling out the possibility of them being shaped like a smaller ball. Clearly, every set that is closed under rotations must be a ball, hence we can not expect any kind of closure for rotations acting on qudits.

3.1.3 Composite systems: multiple qubit or qudit systems

The case of the multipartite qudit system

Definition 3.1.5. In order to call a basis a **n-particle Bloch decomposition** we will require, it to be:

(i) being a tensor product basis, consisting of n local bases,
additionally we require that each local basis is

(ii) containing the identity,

(iii) orthogonal,

(iv) traceless.

The last three properties are of course the usual properties of a classical Bloch decomposition. However the first requirement is not a necessary choice. In theory a basis could be also composed of non-product operators. Our inclusion of (i) in Section 3.1.5 is justified, because this requirement ensures that by tracing out $n-1$ systems, we reobtain the standard Bloch vector decomposition. In fact (i) makes the generalized Bloch decomposition closed under partial traces. Tracing out an arbitrary number of subsystems leaves us again in a generalized Bloch decomposition.

Considering the many challenges faced in the last chapter, when generalizing the Bloch decomposition from a single qubit to a single qudit, it may be surprising to find that generalizing from a single particle to a system of particles is a straightforward affair. The description of the general n-partite qudit system is almost completely determined by the composition axiom (Treated in the first chapter in Definition 2.1.16) of quantum mechanics. Axiom 5 states that a composite system is treated as the tensor product of its constituents. Therefore there is now freedom left once the local description is known. The way we choose to describe the single qudit is only determined up to a choice of basis (c.f. Definition 3.5). Once the basis is chosen, we have implicitly determined already the general description of the general composite system.

If we are studying a composite n-partite qudit system, it is described by the pair of its density matrix and the composite Hilbert space $(\rho, \mathcal{H})$. According to 2.1.16 the Hilbert space necessarily is

$$\mathcal{H} := \mathcal{H}_1 \otimes \mathcal{H}_2 \otimes \cdots \otimes \mathcal{H}_n = \mathbb{C}_{d_1} \otimes \mathbb{C}_{d_1} \otimes \cdots \otimes \mathbb{C}_{d_n}, \tag{3.6}$$

where for each Hilbert space $\mathcal{H}_k = \mathbb{C}_{d_k}$ we choose a complete operator basis denoted as $\mathcal{B}_k = \{\lambda_{k_1}, \ldots, \lambda_{k_{d^2}}\}$. We will discuss several special choices of basis during the article, λ_{k_i} will always denote a traceless and orthogonal basis.

Similarly, the resulting tensor product basis (TPB) demanded in Definition 3.1.5 for $\mathcal{H}$ is given again by Axiom 5 (cf. Section 2.1.16) as

$$\mathcal{B} = \{\lambda_{1_1} \otimes \lambda_{2_1} \otimes \cdots \otimes \lambda_{n_1}, \ldots, \lambda_{1_{d^2}} \otimes \lambda_{2_{d^2}} \otimes \cdots \otimes \lambda_{n_{d^2}}\}, \tag{3.7}$$

where the local λ_i is provided by equation (3.5).

Now before explicitly writing down the generalized Bloch vector decomposition in the last Section 3.1.3 of the chapter, we interrupt briefly with a section devoted on notational issues.

Notation for qudit systems

In a way the main difficulty arising while treating multipartite system is not only a mathematical difficulty but the increasingly cumbersome notation. Therefore we close this section, by collecting some remarks on the notation involved when treating multipartite qudit systems.

We will employ the wide spread multi-index notation for treating multi-linear systems.

Definition 3.1.6. Let α be an n-tuple and $i_1, \ldots, i_n \in \mathbb{N}$ We call

$$\alpha = (i_1, \ldots, i_n) \tag{3.8}$$

an multi-index. Multi indices will always be denoted as single greek letters or directly as a n-tuple of latin letters.
Summation over a multi-index is to be understood as

$$\sum_{\alpha} = \sum_{i_1=1,\ldots,i_n=1}^{m_1,\ldots,m_n} \tag{3.9}$$

Due to their frequent use, it is also helpful to define a shorthand notation for the expectation values of a basis element. This "tensor product multi-index" notation is less standard, but proves to be useful.

Definition 3.1.7. Let λ be a basis, then we write

$$c_{i_1,i_2,\cdots,i_n} := \langle \lambda_{i_1,i_2,\ldots,i_n} \rangle = \mathrm{Tr}[\rho \bigotimes_{k=1}^{n} \lambda_{i_k}]\,. \tag{3.10}$$

Recall that we have required the basis of generalized Bloch vector decomposition to be a tensor product basis. Note that we will reserve the combination of c,λ with multi indices for basis elements and their coefficients. In the context of a basis, a multi-index will always refer to the indices of local bases in a tensor product. This will make it unambiguous, whenever we are referring to tensor product multi indices.

The 0-th basis element is always reserved for the identity $\lambda_0 = \mathbb{1}_d$. It is of special significance whether a local basis element is an identity or not, since the identity signifies no information in the

corresponding subsystem. The omission of multi-indices will be used equivalently writing $\mathbb{1}_d$ in the subsystem of the omitted index, thus we write

$$c_{i_j,i_l} := \langle \lambda_{i_j,i_l} \rangle = \langle \lambda_{0,\cdots,0,i_j,0,\cdots,i_l,0,\ldots,0} \rangle = \mathrm{Tr}[\rho \bigotimes_{k=1}^{n} \lambda_{(\delta_{i_j}+\delta_{i_l})i_k}] \tag{3.11}$$

$$= \mathrm{Tr}[\rho(\mathbb{1}_{d_1} \otimes \cdots \otimes \mathbb{1}_{d_{j-1}} \otimes \lambda_{i_j} \otimes \mathbb{1}_{d_{j+1}} \otimes \cdots \otimes \mathbb{1}_{d_{l-1}} \otimes \lambda_{i_l} \otimes \mathbb{1}_{d_{l+1}} \otimes \cdots \otimes \mathbb{1}_{d_n})], \tag{3.12}$$

with total analogous notation for more indices.

Observation 3.1.8. *We have only introduced the "multi-index tensor product" shorthand for correlation tensor coefficients (i.e: expectation values of basis elements). If we would extend the notation to general basis representations we may have a notation conflict between a single subsystem and the composite system with only identities except in one system.*

The correlation tensor & sectors of correlations

By using any of the bases provided in Chapter 3.2 we can also write out the density matrix in the familiar way.

Definition 3.1.9. Let ρ be a general n partite qudit state in $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2 \otimes \cdots \otimes \mathcal{H}_n := \mathbb{C}_{d_1} \otimes \mathbb{C}_{d_1} \otimes \cdots \otimes \mathbb{C}_{d_n}$, we can express ρ in the n-particle Bloch vector basis

$$\rho = \frac{1}{d_1 \ldots d_n} \sum_{i_1,i_2,\cdots,i_n=0}^{d-1} c_{i_1,i_2,\cdots,i_n} \bigotimes_{k=1}^{n} \lambda_{i_k}, \tag{3.13}$$

with $\lambda_0 = \mathbb{1}_d$. The coefficients $c_{i_1,i_2,\cdots,i_n}$ are now labeled by a multi-index $\alpha := (i_1, i_2, \cdots, i_n)$ and given by

$$c_{i_1,i_2,\cdots,i_n} = \mathrm{Tr}[\rho \bigotimes_{k=1}^{n} \lambda_{i_k}]\,. \tag{3.14}$$

We call the tensor that collects the indices of all local systems into one **the correlation tensor**

$$\mathcal{C}(\rho) := (c_{i_1,i_2,\cdots,i_n}). \tag{3.15}$$

We also define the **i-th sector** $\mathcal{C}_{|\alpha|=i}(\rho)$ of the correlation tensor of ρ as

$$\mathcal{C}_{|\alpha|=i}(\rho) := (c_\alpha). \tag{3.16}$$

The intuition here is that each sector $\mathcal{C}_{|\alpha|=i}(\rho)$ contains all the i-body correlations of the system described by ρ. In the definition of the generalized Bloch vector decomposition (c.f. Definition 3.1.5), we have required it to be a tensor product basis as well as including the identity in the local bases. Including the local identities in the bases gives us a way to express locality or globality of correlations directly in the indexing of the generalized Bloch bases coefficients c_α. The tensor product structure of our basis allows us to regroup the coefficients c_α into several subtensors $\mathcal{C}_{|\alpha|=i}(\rho)$, that contain the correlations in a structured way. Each subtensor $\mathcal{C}_{|\alpha|=i}(\rho)$ describes correlations between as many particles as the number of non-identity basis elements appearing in its respective

tensor product basis. By adopting the convention to drop all zero entries corresponding to local identities in a multi-index α labeling the coefficient of a tensor product basis, the multi-indices cardinality $|\alpha|$ denotes exactly the number of involved parties, whose correlations are described by α.

The state of a subsystem $\rho_\alpha = \mathrm{Tr}_{\overline{\alpha}}[\rho]$ containing systems $\alpha \subseteq \{1, 2, \cdots, n\}$ is completely determined by the tensor elements where all indices that are not part of α are set to zero. This yields a natural division of the whole correlation tensor $\mathcal{C}(\rho)$ into subsystem correlation tensors c_α, that encode correlations between all nontrivial observables within α, i.e. $||\mathcal{C}(\rho)|| = \sum_{\alpha \subseteq \{1,2,\cdots,n\}} ||c_\alpha||$. Thus it is possible to rewrite the generalized Bloch basis in terms of its sectors $\mathcal{C}_{|\alpha|=i}(\rho)$ to resemble the Bloch vector decomposition more closely by explicitly regrouping it into subtensors that include a fixed number of local identities.

Definition 3.1.10. Let ρ be a general n partite qudit state in $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2 \otimes \cdots \otimes \mathcal{H}_n := \mathbb{C}_{d_1} \otimes \mathbb{C}_{d_1} \otimes \cdots \otimes \mathbb{C}_{d_n}$, we can express ρ in general as

$$\|\rho\|_2^2 = \frac{1}{d^n}(\mathbb{1} + \underbrace{\sum_j \sum_{i_j}^{d-1} ||c_{i_j}||_2^2}_{||\mathcal{C}_1(\rho)||_2^2} + \underbrace{\sum_{j,l} \sum_{i_j,i_l}^{d-1} ||c_{i_j,i_l}||_2^2}_{||\mathcal{C}_2(\rho)||_2^2} + \cdots + \underbrace{\sum_{i_1,i_2,\cdots,i_n=1}^{d-1} ||c_{i_1,i_2,\cdots,i_n}||_2^2}_{||\mathcal{C}_{|\alpha|=n}(\rho)||_2^2)}, \tag{3.17}$$

Recall that by 3.12 the omission of indices in c means that they are set to zero. For example we have $c_{i_2,i_{n-1}} = c_{0,i_2,0,\cdots,0,i_{n-1},0}$

3.2 Possible bases

3.2.1 Gell-Mann basis

Definition 3.2.1. The **generalized Gell-Mann matrices** [1] as a basis for a Hilbert space of dimension d are divided into three sets of matrices; a symmetric set, an antisymmetric set and a diagonal set.

The set of **symmetric Gell-Mann matrices** is defined as

$$g_{jk}^S := \sqrt{\frac{d}{2}}(|k\rangle\langle j| + |j\rangle\langle k|) \tag{3.18}$$

The set of **antisymmetric Gell-Mann matrices** are defined as

$$g_{jk}^A := -i\sqrt{\frac{d}{2}}(|k\rangle\langle j| - |j\rangle\langle k|) \tag{3.19}$$

The set of **diagonal Gell-Mann matrices** is defined as

$$g_l^D := \sqrt{\frac{d}{l(l+1)}}\left(\sum_{j=1}^{l} |j\rangle\langle j| - l\,|l+1\rangle\langle l+1|\right) \tag{3.20}$$

[1] On a historical note it is not entirely clear to us who to attribute the generalized Gell-Mann matrices to. The $d = 3$ construction is due to Gell-Mann, however who is responsible for the first generalization is more unclear to us. We could trace back the mention of the generalized Gell-Mann matrices to [24]. We are happy about any comments on earlier work unknown to us.

where $1 \leq l \leq d-1$ and $1 \leq j < k \leq d$.

The generalized Gell-Mann matrices exhibit the properties of being

- complete,
- orthogonal,
- traceless,
- **Hermitian**,
- rank 2 non-trivially anti-commuting basis elements contained in sets of cardinality 2.

Correspondence to the Pauli matrices

The GGM can be seen as the Hermitian generalization of the Pauli matrices. We illustrate this way of generalizing the Pauli matrices by stating the GGM matrices and their correspondence to the Pauli matrices explicitly.
Note that we have omitted the normalization factors of the GGM below to improve the readability.
The symmetric class of generalized Gell-Mann matrices $g_{ij}^S 0 \leq i < j \leq d$ generalizes σ_x by

$$\sigma_x \to \left\{ \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \right\}$$

$$\to \left\{ \begin{pmatrix} 0 & 1 & 0 & \cdots 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \vdots & 0 & 0 & \ddots \end{pmatrix}, \begin{pmatrix} 0 & 0 & 1 & \cdots \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ \vdots & 0 & 0 & \ddots \end{pmatrix} \cdots \right\}$$

The anti-symmetric class of generalized Gell-Mann matrices $g_{ij}^A 0 \leq i < j \leq d$ generalizes σ_y by

$$\sigma_y \to \left\{ \begin{pmatrix} 0 & 0 & -i \\ 0 & 0 & 0 \\ i & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & -i & 0 \\ 0 & 0 & 0 \\ i & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix} \right\}$$

$$\to \left\{ \begin{pmatrix} 0 & -i & 0 & \cdots \\ i & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \vdots & 0 & 0 & \ddots \end{pmatrix}, \begin{pmatrix} 0 & 0 & -i & \cdots \\ 0 & 0 & 0 & 0 \\ i & 0 & 0 & 0 \\ \vdots & 0 & 0 & \ddots \end{pmatrix} \cdots \right\}$$

The diagonal class of general Gell-Mann matrices $g_i^D 1 \leq i <\leq d$ generalizes σ_z by

$$\sigma_z \to: \left\{ \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix} \right\}$$

$$\to \left\{ \begin{pmatrix} 1 & 0 & 0 & \cdots \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -2 & 0 \\ \vdots & 0 & 0 & \ddots \end{pmatrix}, \cdots, \begin{pmatrix} 1 & 0 & 0 & \cdots \\ 0 & 1 & 0 & 0 \\ \vdots & 0 & \ddots & \vdots \\ 0 & 0 & 0 & -d \end{pmatrix} \right\}$$

Spectrum

Note that the largest eigenvalue λ for the square of matrices element of these collections is given by

$$\max \lambda \left((g_{jk}^S)^2\right) = \max \lambda \left((g_{jk}^A)^2\right) = \frac{d}{2} \tag{3.21}$$

$$\max \lambda \left((g_l^D)^2\right) = \frac{dl}{l+1} \tag{3.22}$$

3.2.2 Weyl basis

Definition 3.2.2. The **Heisenberg-Weyl matrices** [50] as a basis for a Hilbert space of dimension d that consist of all powers of two matrices: the shift and the clock matrices.
The **Shift Matrix** X is defined as:

$$X := \sum_{i=1}^{d-1} |i+1\rangle\langle i| + |1\rangle\langle d| \tag{3.23}$$

The **Clock Matrix** Z is defined as:

$$Z := \sum_{i=1}^{d} \omega^{i-1} |i\rangle\langle i| \tag{3.24}$$

with ω denoting the d-th root of unity.

Observation 3.2.3. *Properties of the Heisenberg-Weyl Operators*
The Heisenberg-Weyl basis exhibit the properties of being

1. *complete,*
2. *orthogonal,*
3. *traceless,*
4. ***unitary**,*
5. *(anti-)commutation relation $Z^l X^m = X^m Z^l e^{i2\pi lm/d}$, at least in some cases anti commutative,*
6. *rank n non-trivially anti-commuting basis elements.*

Proof. Clearly X has only offdiagonals, thus they are traceless. $\sum_{i=1}^{d} \omega^{i-1} = 0$, because ω is the d-th root of unity, making Z traceless. By the elementary calculation

$$ZZ^\dagger = (\sum_{i=1}^{d} \omega^{i-1}|i\rangle\langle i|)(\sum_{i=1}^{d} -\omega^{i-1}|i\rangle\langle i|) = \mathbb{1}, XX^\dagger = (\sum_{i=1}^{d-1} |i+1\rangle\langle i| + |1\rangle\langle d|)(\sum_{i=1}^{d-1} |i\rangle\langle i+1| + |d\rangle\langle 1| = \mathbb{1}) \tag{3.25}$$

we see unitarity and the first part of orthogonality.

□

The paradigmatic difference between the Heisenberg-Weyl operators and the generalized Gell-man matrices is the unitarity of the Heisenberg-Weyl operators. In quantum information the generalized Gell-Mann matrices are frequently employed, this is not surprising since their property of being Hermitian matches the requirement for physical observables in the second fundamental axiom of quantum mechanics 2.1.2.

The utility of the unitarity that the Heisenberg-Weyl operators provide may be not as obvious. Generally speaking unitary operator allow for very easy invariance properties, that can be of great technical use

Surprisingly, this property turns out to be even useful when studying quantum thermodynamics. Many information theoretic quantities rely on combination of unitary invariant properties such as entropies. In our publication [28] and Chapter 5 , we connect these entropies with the thermodynamic notion of temperature. We aim to find the maximal possible amount of information within a system given a certain temperature. QI provides the GHZ state that is known to maximize certain relevant entropies. That state can be extended to a generalized Bell basis via the Heisenberg-Weyl operators and their unitarity keeps all relevant entropic quantities maximal.

Correspondence to the Pauli matrices

The shift class X generalizes σ_x by

$$\sigma_x \to X = \begin{pmatrix} 0 & 0 & \cdots & 1 \\ 1 & 0 & 0 & 0 \\ \vdots & \ddots & 0 & \vdots \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

The clock class Z generalizes σ_z by

$$\sigma_z \to Z = \begin{pmatrix} 1 & 0 & 0 & \cdots \\ 0 & \omega & 0 & 0 \\ \vdots & 0 & \ddots & \vdots \\ 0 & 0 & 0 & \omega^{d-1} \end{pmatrix} \tag{3.26}$$

The discrete displacement operators

The earlier introduced clock and shift operators can be collected into one operator. It describes The unitaries corresponding to discrete phase-space displacements for d-level systems are defined by

Definition 3.2.4. We define the **discrete displacement operator** as

$$\mathcal{D}(l,m) = Z^l X^m e^{-i\pi lm/d}. \tag{3.27}$$

Displacement operators hold a number of convenient properties. Principal among these is the completeness of displacement operators. That is, they form a complete non-Hermitian orthogonal basis satisfying the orthogonality condition

$$\mathrm{Tr}\{\mathcal{D}(l,m)\mathcal{D}^\dagger(l',m')\} = d\delta_{l,l'}\delta_{m,m'}. \tag{3.28}$$

Therefore, any bound operator, including density operators ρ, can be decomposed to

$$\rho = \frac{1}{d}\sum_{l,m=0}^{d-1} \mathrm{Tr}\{\rho\mathcal{D}(l,m)\}\mathcal{D}^\dagger(l,m) \equiv \frac{1}{d}(\mathbb{1} + \vec{\xi}\cdot\vec{\mathcal{D}}^\dagger), \tag{3.29}$$

from which the Bloch representation is already apparent.

In this formulation, however, the Bloch vector components, $\xi_{lm} = \mathrm{Tr}\{\rho\mathcal{D}(l,m)\}$ are generally complex as the displacement operators are not Hermitian. Therefore we need to determine $(d^2 - 1)$ complex parameters of the Bloch vector, $\vec{\xi}$, to fully characterize the density operator. The obvious question here is, could we find a minimal complete set of $d^2 - 1$ Hermitian operators whose expectation values with respect to the density operator are sufficient to fully characterize the state. In the following we develop a basis which has the above property.

3.2.3 Symmetrized Heisenberg-Weyl basis

In the last part we have introduced the Heisenberg-Weyl basis and commented that its lack of Hermiticity is unsatisfactory. In this part we follow the simple idea to force the Heisenberg-Weyl operators to be Hermitian by directly symmetrizing them.

Derivation of the symmetrized Heisenberg-Weyl basis

Here we are aiming to identify a minimal and complete set of Hermitian operators constructed from the HW operators, $\mathcal{D}(l,m)$. Our attempt begins by making an Ansatz solution of the form

$$\mathcal{Q}(l,m) := \chi\mathcal{D}(l,m) + \chi^*\mathcal{D}^\dagger(l,m), \tag{3.30}$$

which is Hermitian and traceless by construction. Just as the Heisenberg-Weyl or Gell-Mann matrices, we require the basis of observables in this Ansatz to be orthogonal.

Lemma 3.2.5. *The orthogonality condition*

$$Tr\{\mathcal{Q}(l,m)\mathcal{Q}(l',m')\} = d\delta_{l,l'}\delta_{m,m'}. \tag{3.31}$$

is satisfied only for the choice

$$\chi = \frac{(1 \pm i)}{2}. \tag{3.32}$$

Proof. We want to find χ defined in (3.30) such that

$$\mathrm{Tr}\{\mathcal{Q}(l,m)\mathcal{Q}(l',m')\} = d\delta_{l,l'}\delta_{m,m'}. \tag{3.33}$$

Therefore we expand

$$\begin{aligned}
&\mathrm{Tr}\{\mathcal{Q}(l,m)\mathcal{Q}(l',m')\} && (3.34)\\
&= |\chi|^2\mathrm{Tr}\{\mathcal{D}(l,m)\mathcal{D}^\dagger(l',m')\} + |\chi|^2\mathrm{Tr}\{\mathcal{D}^\dagger(l,m)\mathcal{D}(l',m')\} && (3.35)\\
&\quad \chi^2\mathrm{Tr}\{\mathcal{D}(l,m)\mathcal{D}(l',m')\} + \chi^{*2}\mathrm{Tr}\{\mathcal{D}^\dagger(l,m)\mathcal{D}^\dagger(l',m')\} && (3.36)\\
&= 2d|\chi|^2\delta_{l,l'}\delta_{m,m'} && (3.37)\\
&+ de^{i\pi(l-m)}(\chi^2 e^{-i\pi d} + \chi^{*2}e^{i\pi d})\delta_{l+l',d}\delta_{m+m',d} = d\delta_{l,l'}\delta_{m,m'}. && (3.38)
\end{aligned}$$

In order to satisfy the orthogonality condition we need to have that

$$|\chi|^2 = \frac{1}{2} \quad \text{and} \quad \chi^2 = -\chi^{*2}. \tag{3.39}$$

Simple algebra leads to the solution $\chi = (1 \pm i)/2$.

□

We established $d^2 - 1$ orthogonal and traceless observables which are linearly independent. The $d^2 - 1$ observables plus identity matrix, $\mathcal{Q}(0,0) = \mathbb{1}_d$ form a basis acting on a d dimensional Hilbert space.

Definition 3.2.6. The **symmetrized Heisenberg-Weyl matrices** are a basis for a Hilbert space of dimension d. They are defined defined as:

$$\mathcal{Q}(l,m) := \frac{(1+i)}{2}\mathcal{D}(l,m) + \frac{(1-i)}{2}\mathcal{D}^\dagger(l,m), \tag{3.40}$$

with $\mathcal{D}(l,m)$ denoting the discrete d-dimensional displacement operator.

This basis simply reduces to Pauli matrices for $d = 2$. We henceforth refer to its elements as "Heisenberg-Weyl observables" or "Symmetrized Heisenberg-Weyl basis" .

Properties of symmetrized Heisenberg-Weyl basis

The **symmetrized Heisenberg-Weyl matrices** have the following properties:

- Orthogonal
- Traceless
- **Hermitian**
- **Anti commutation relation**
- Rank n non-trivially anti-commuting basis elements contained in sets of cardinality 3.

They have distinct properties from those of the standard Heisenberg-Weyl matrices. Note that after symmetrizing the Heisenberg-Weyl basis, it is no longer unitary. Effectively we are trading unitarity for Hermiticity. Therefore the symmetrized Heisenberg-Weyl basis exhibits the defining property of the Gell-Mann matrices, instead of the original Heisenberg-Weyl matrices.

Even though symmetrizing a quantity to enforce Hermiticity seems not to be an unusual idea, it seems that these are rarely studied. Even though the construction method of the symmetrized Heisenberg-Weyl operators seems so obvious, we where not able to find literature on them up to date. We can only guess, that this is due to the fact, that the Gell-Mann matrices already are Hermitian. Since the Heisenberg-Weyl operators lose their defining trait of unitarity after symmetrization, it may seem that we have not gained anything over the more standard GGM. Possibly, this sentiment has led to the community not studying symmetrized Heisenberg-Weyl observables.

However HW observables in contrast to GGM generically have full-rank, this allows them to inherit the *non trivial anti-commutativity relation* from the standard HW observables. It also leads to the slightly larger size of anti-commuting sets contained within the symmetrized Heisenberg-Weyl. For HWO these sets consist of 3 elements instead of 2 in the GGM case. Furthermore the full rank makes them more efficient in sparsely characterizing states with a lot of coherence.

The possibility to "cover" an observable by anti-commuting sets gains importance due to a theorem we will be proving in Section 4.4.2. There we show that the squared coefficients of an anti-commuting basis can be conveniently bounded. For the HWO the bound depends essentially on the spectrum, but not on the dimension or the particle number! This could be a way to circumvent many problems arising in multipartite entanglement detection. There dimensionality dependent factor blow up bounds that need to be violated to detect entanglement. Removing these factors by dimension independent methods such as the anti-commutativity bound is highly desirable. Unfortunately, the theorem of Section 4.4.2 only applies to the anti-commuting part of a basis. The theorem can also be applied to a basis that can be covered by several anti-commuting sets by repeated use if the theorem, however each repeated use leads to worse bounds. For this reason we ware interested in bases that can be covered by as few anti-commuting sets as possible. It is easier to cover a basis the closer the rank of the single basis element is to the rank of the observable we are studying. This makes the rank n elements of the HWO appealing.

Infinite dimensional limit of symmetrized Heisenberg-Weyl basis

Notably, HW observables can be systematically extended to the continuous limit of infinite dimensional systems holding analogous properties, and with this also all corresponding results can be extended in this limit. Up until now we have worked only with finite dimensional systems, since our Heisenberg-Weyl observables are constructed from discrete displacement operators. The symmetrized Heisenberg-Weyl basis inherits their infinite dimensional limit: the continuous displacement operator. We employed two real discrete indices l, m in order to label $\mathcal{Q}(l, m)$. Generally we will keep to this convention since it allows for easy checking of commutativity and anti-commutativity, however in order to study the limit of $\mathcal{Q}$ a continuous indexing is necessary. This is best achieved by switching notation to complex numbers α.

Definition 3.2.7. We define the **continuous displacement operator** $\mathcal{D}(\alpha)$ as

$$\mathcal{D}(\alpha) := e^{\alpha a^\dagger - \alpha^* a} \tag{3.41}$$

with $a^\dagger, a$ the creation-annihilation operators and $\alpha \in \mathbb{C}$ as the **displacement amplitude**. We can define in total analogy to 3.40 the **continuous symmetrized Heisenberg-Weyl** operators by

$$\mathcal{Q}(\alpha) := \frac{1+i}{2}\mathcal{D}(\alpha) + \frac{1-i}{2}\mathcal{D}^\dagger(\alpha). \tag{3.42}$$

If

$$\alpha \in \mathbb{C} = \sqrt{\frac{\pi}{d}}(m + il) \tag{3.43}$$

then α is also the **discrete displacement amplitude** and generalizes the discrete case by $\mathcal{Q}(l, m) = \mathcal{Q}(\mathrm{Re}(\alpha), \mathrm{Im}(\alpha)))$.

Thus we can write

$$\rho = \frac{1}{d}\Big(\mathbb{1} + \sum_{\alpha\in\mathcal{S}} \langle Q(\alpha)\rangle \mathcal{Q}(\alpha)\Big), \tag{3.44}$$

where we used $\mathcal{S} := \{\alpha : \alpha_I = \sqrt{\frac{\pi}{d}}l, \alpha_R = \sqrt{\frac{\pi}{d}}m\}$. To facilitate a smooth transition to infinite dimensions one can consider $\hat{x} = Q\sqrt{2\pi/d}$ and $\hat{p} = P\sqrt{2\pi/d}$ as the position and momentum operators. In this case, $X^m \equiv e^{-ix\hat{p}}$ indicates position displacement by $x = m\sqrt{2\pi/d}$. Similarly $Z^l \equiv e^{ip\hat{x}}$ displaces the momentum by $p = l\sqrt{2\pi/d}$. Therefore Eq. (3.27) can be rewritten

$$\mathcal{D}(p,x) \equiv e^{ip\hat{x}}e^{-ix\hat{p}}e^{-ixp/2}. \tag{3.45}$$

In the limit $d \to \infty$ we recover the Heisenberg commutation relation for position and momentum of a continuous variable system, $[\hat{x},\hat{p}] = i$. We can then use the special form of Baker-Campbell-Hausdorff formula for exponential operators, i.e. $e^{A+B} = e^Ae^Be^{-[A,B]/2}$ where $[A,[A,B]] = 0 = [B,[A,B]$. Thus, in this limit the definition (3.45) can be written $\mathcal{D}(p,x) = e^{ip\hat{x}-ix\hat{p}}$, the form of which is valid only in the infinite dimensional limit. An equivalent reformulation of this is $\mathcal{D}(\alpha) = e^{\alpha a^\dagger - \alpha^* a}$ with orthogonality condition $\mathrm{Tr}\{\mathcal{D}^\dagger(\alpha)\mathcal{D}(\alpha')\} = \pi\delta^2(\alpha-\alpha')$ where $a^\dagger(a)$ denotes creation (annihilation) operators of a bosonic mode and α is the displacement amplitude. Therefore, the continuous analog of equation 3.34 is

$$\mathrm{Tr}\{\mathcal{Q}(\alpha)\mathcal{Q}(\alpha')\} = \pi\delta^2(\alpha-\alpha'). \tag{3.46}$$

The discrete-continuous transition is therefore identified by the replacement $\frac{1}{d}\sum_\alpha \to \frac{1}{\pi}\int d^2\alpha$. This shows that our observable basis developed for discrete systems can be systematically extended to continuous variable systems.

For example we can now extend the anti-commutativity approach to entanglement detection from discrete systems to continuous variable systems. As an example for an continuous variable system, let us introduce the well-known coherent state state.

Definition 3.2.8. A **Coherent state** $|\alpha\rangle$ is defined as

$$|\alpha\rangle := e^{(\alpha a^\dagger - \alpha^* a)}|0\rangle = D(\alpha)|0\rangle, \tag{3.47}$$

with α the **amplitude** of the coherent state $|\alpha\rangle$.

A **2-mode squeezed state** $|\xi\rangle$ is defined as

$$|S_2(\xi)\rangle := e^{(\xi^* ab - \xi a^\dagger b^\dagger)}|0\rangle := S_2(\xi)|0\rangle, \tag{3.48}$$

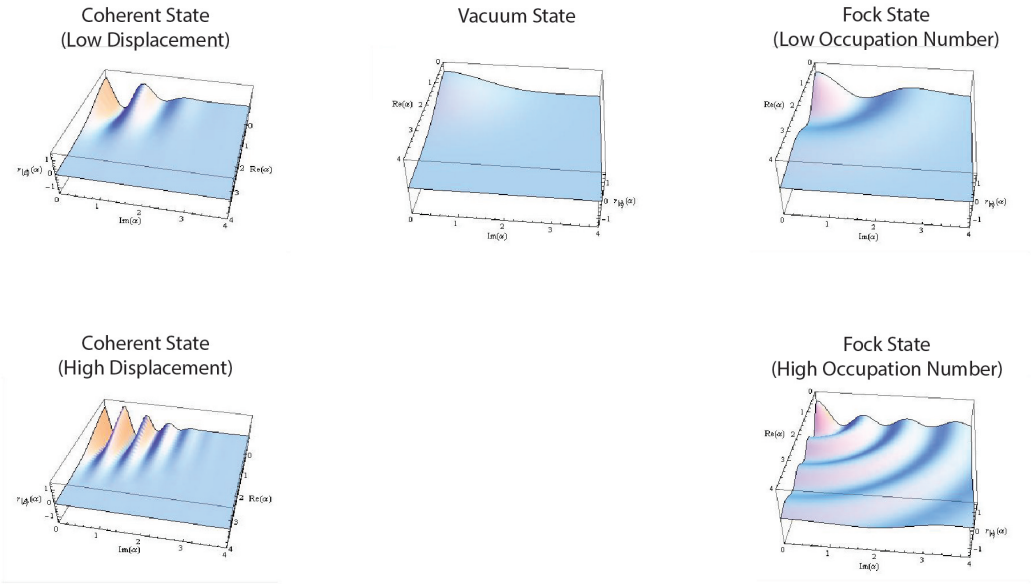
with ξ the so-called **squeezing parameter**.

Finally we define the **squeezed coherent states** as

$$|\alpha,\xi\rangle := D(\alpha)S(\xi)|0\rangle \tag{3.49}$$

where r denotes a constant describing a damping factor, $a, b, a^\dagger, b^\dagger$ are the creation annihilation operators of two fixed modes. $e^{ir(a^\dagger b^\dagger + ab)}$ corresponds to **two mode squeezing**, a well-known operation from quantum optics

Figure 3.1: We compare several typical states in Heisenberg-Weyl Observable representation.



We can make a suitable explicit choice of HW observables for entanglement detection in a Gaussian system. Then we check that this choice remains valid if the dimension d of our system goes towards infinity and finally calculate the expectation values. We will later proof a theorem, that will allow us to employ the Heisenberg-Weyl observable for entanglement detection (c.f. Theorem 4.4.1).

We start by calculating a displacements operators expectation value for a Fock state for later use. Assume without loss of generality $m \geq n$, else the following argument can be obtained by letting all operators act in reverse directions.

$$\langle m|D(\alpha)|n\rangle = \langle m|D(\alpha)\frac{(a^\dagger)^n}{\sqrt{n!}}|0\rangle = \langle m|\frac{(a^\dagger-\alpha^*)^n}{\sqrt{n!}}D(\alpha)|0\rangle = \langle m|\frac{(a^\dagger-\alpha^*)^n}{\sqrt{n!}}|\alpha\rangle, \tag{3.50}$$

Where we made use of $D(\alpha)a^\dagger = (a^\dagger - \alpha^*)D(\alpha)$, as a consequence of $D(\alpha)a^\dagger D(\alpha)^\dagger = a^\dagger - \alpha^*$. By letting act $a^\dagger$ to the left and applying the binomial theorem we derive

$$\langle m|\frac{(a^\dagger-\alpha^*)^n}{\sqrt{n!}}|\alpha\rangle = \sum_{k=0}^{n}\binom{n}{k}\langle m|\frac{(a^\dagger)^{n-k}(-1)^k(\alpha^*)^k}{\sqrt{n!}}|\alpha\rangle = \tag{3.51}$$

$$= \sum_{k=0}^{n}\binom{n}{k}(-1)^k(\alpha^*)^k\frac{\sqrt{m\cdot(m-1)\cdot\ldots\cdot(m-n+k)}}{\sqrt{n!}}\langle m-n+k|\alpha\rangle, \tag{3.52}$$

where only how to evaluate the overlap of an arbitrary Fock state with a given coherent state $\langle m-n+k|\alpha\rangle$ is yet unclear.

Due to $\langle n|\alpha|\alpha\rangle = \langle n|a|\alpha\rangle = \sqrt{n+1}\langle n+1|\alpha\rangle$ we have the recurrence relation

$$\langle n|\alpha\rangle = \frac{\alpha}{\sqrt{n}}\langle n-1|\alpha\rangle = \ldots = \frac{\alpha^n}{\sqrt{n!}}\langle 0|\alpha\rangle \tag{3.53}$$

between all Fock states. This reduces the computation of any Fock state's overlap to the calculation of the overlap between the vacuum state with the coherent state. The vacuum case in turn can be easily treated by Baker-Campbell-Hausdorff and letting act the annihilation operators to the right, while applying creation to the left

$$\langle 0|\alpha\rangle = \langle 0|D(\alpha)|0\rangle = e^{-\frac{1}{2}|\alpha|^2}\langle 0|e^{\alpha a^\dagger}e^{-\alpha^* a}|0\rangle = e^{-\frac{1}{2}|\alpha|^2}\langle 0|(1+\alpha a^\dagger+\ldots)(1-\alpha^* a+\ldots)|0\rangle = e^{-\frac{1}{2}|\alpha|^2}. \tag{3.54}$$

Collecting the above equations together we finally obtain for an expression of the displacement operators expectation value for any Fock state

$$\langle m|D(\alpha)|n\rangle = \tag{3.55}$$

$$= \sum_{k=0}^{n}\binom{n}{k}(-1)^k(\alpha^*)^k\frac{\sqrt{m\cdot\ldots\cdot m-n+k}}{\sqrt{n!}}\underbrace{\langle m-n+k|\alpha\rangle}_{\frac{\alpha^{m-n+k}}{\sqrt{(m-n+k)!}}e^{-\frac{1}{2}|\alpha|^2}} = \tag{3.56}$$

$$= e^{-\frac{1}{2}|\alpha|^2}\sum_{k=0}^{n}\binom{n}{k}(-1)^k(|\alpha|)^{2k}\alpha^{m-n}\frac{\sqrt{m\cdot\ldots\cdot m-n+k}}{\sqrt{n!}\sqrt{(m-n+k)!}}. \tag{3.57}$$

More specifically for the $m=n$ case, the diagonal elements of the state simplify to the following form

$$\langle n|D(\alpha)|n\rangle = e^{-\frac{1}{2}|\alpha|^2}\underbrace{\sum_{k=0}^{n}\binom{n}{k}(-1)^k(|\alpha|)^{2k}\frac{\sqrt{n\cdot(n-1)\cdot\ldots\cdot k}}{\sqrt{n!}\sqrt{k!}}}_{=:L_n(|\alpha|^2)}. \tag{3.58}$$

Note that the term we have denoted $L_n(|\alpha|^2)$ above, are exactly the well-known *Laguerre Polynomials* of order n evaluated at $|\alpha|^2$! We are now ready to compute expectation values of our newly defined symmetrized Heisenberg-Weyl observables. For this we can basically reuse the above formula to insert in

$$\langle m|Q(\alpha)|n\rangle = \frac{1+i}{2}\langle m|\mathcal{D}(\alpha)|n\rangle + \frac{1-i}{2}\langle m|\mathcal{D}(-\alpha)|n\rangle. \tag{3.59}$$

Realizing that both expectation values only differ in α we compare

$$\langle m|\mathcal{D}(\alpha)|n\rangle + \langle m|\mathcal{D}(-\alpha)|n\rangle = \tag{3.60}$$

$$= e^{-\frac{1}{2}|\alpha|^2}\sum_{k=0}^{n}\binom{n}{k}(-1)^k(|\alpha|)^{2k}\frac{\sqrt{m\cdot\ldots\cdot m-n+k}}{\sqrt{n!}\sqrt{(m-n+k)!}}(\alpha^{m-n}+(-\alpha)^{m-n}) \tag{3.61}$$

$$i(\langle m|\mathcal{D}(\alpha)|n\rangle - \langle m|\mathcal{D}(-\alpha)|n\rangle) = \tag{3.62}$$

$$= e^{-\frac{1}{2}|\alpha|^2}\sum_{k=0}^{n}\binom{n}{k}(-1)^k(|\alpha|)^{2k}\frac{\sqrt{m\cdot\ldots\cdot m-n+k}}{\sqrt{n!}\sqrt{(m-n+k)!}}i(\alpha^{m-n}-(-\alpha)^{m-n}). \tag{3.63}$$

Note that due to

$$(\alpha^{m-n} + (-\alpha)^{m-n}) = \begin{cases} 0 | m-n| \equiv 1(2) \\ 2(\alpha)^{m-n} | m-n| \equiv 1(2) \end{cases}, \tag{3.64}$$

and

$$i(\alpha^{m-n} - (-\alpha)^{m-n}) = \begin{cases} 2i(\alpha)^{m-n} | m-n| \equiv 1(2) \\ 0 | m-n| \equiv 1(2) \end{cases}, \tag{3.65}$$

we have

$$(\alpha^{m-n} - (-\alpha)^{m-n}) + i(\alpha^{m-n} - (-\alpha)^{m-n}) = 2(-i\alpha)^{m-n}. \tag{3.66}$$

We are left with

$$\langle m|Q(\alpha)|n\rangle = \frac{1+i}{2} \langle m|\mathcal{D}(\alpha)|n\rangle + \frac{1-i}{2} \langle m|\mathcal{D}(-\alpha)|n\rangle = \tag{3.67}$$

$$= e^{-\frac{1}{2}|\alpha|^2} \sum_{k=0}^{n} \binom{n}{k} (-1)^k (|\alpha|)^{2k} \frac{\sqrt{m \cdot \ldots \cdot m-n+k}}{\sqrt{n!}\sqrt{(m-n+k)!}} (-i\alpha)^{m-n}. \tag{3.68}$$

Explicit representation

For the convenience of the reader here an explicit matrix representation of the HW observables is given in low dimension. For the sake of a compact notation let $\chi = (1+i)/2$ and $\omega = e^{\frac{2i\pi}{3}}$.
The symmetrized Heisenberg-Weyl basis for d=3:

$$\mathcal{Q}(0,0) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \mathcal{Q}(0,1) = \begin{pmatrix} 0 & \chi^* & \chi \\ \chi & 0 & \chi^* \\ \chi^* & \chi & 0 \end{pmatrix} \quad \mathcal{Q}(0,2) = \begin{pmatrix} 0 & \chi & \chi^* \\ \chi^* & 0 & \chi \\ \chi & \chi^* & 0 \end{pmatrix}$$

$$\mathcal{Q}(1,0) = \begin{pmatrix} \chi+\chi^* & 0 & 0 \\ 0 & \chi\omega + \chi^*\omega^* & 0 \\ 0 & 0 & \chi\omega^* + \chi^*\omega \end{pmatrix} \quad \mathcal{Q}(1,1) = \begin{pmatrix} 0 & -\chi^*\omega & -\chi\omega \\ -\chi\omega^* & 0 & -\chi^* \\ -\chi^*\omega^* & -\chi & 0 \end{pmatrix}$$

$$\mathcal{Q}(1,2) = \begin{pmatrix} 0 & \chi\omega^* & \chi^*\omega^* \\ \chi^*\omega & 0 & \chi \\ \chi\omega & \chi^* & 0 \end{pmatrix} \quad \mathcal{Q}(2,0) = \begin{pmatrix} \chi+\chi^* & 0 & 0 \\ 0 & \chi\omega^* + \chi^*\omega & 0 \\ 0 & 0 & \chi\omega + \chi^*\omega^* \end{pmatrix}$$

$$\mathcal{Q}(2,1) = \begin{pmatrix} 0 & \chi^*\omega^* & \chi\omega^* \\ \chi\omega & 0 & \chi^* \\ \chi^*\omega & \chi & 0 \end{pmatrix} \mathcal{Q}(2,2) = \begin{pmatrix} 0 & \chi\omega & \chi^*\omega \\ \chi^*\omega^* & 0 & \chi \\ \chi\omega^* & \chi^* & 0 \end{pmatrix}$$

The symmetrized Heisenberg-Weyl basis for d=4:

$$\mathcal{Q}(0,0)=\begin{pmatrix}1&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&1\end{pmatrix}\quad \mathcal{Q}(0,1)=\begin{pmatrix}0&\chi^*&0&\chi\\\chi&0&\chi^*&0\\0&\chi&0&\chi^*\\\chi^*&0&\chi&0\end{pmatrix}\quad \mathcal{Q}(0,2)=\begin{pmatrix}0&0&1&0\\0&0&0&1\\1&0&0&0\\0&1&0&0\end{pmatrix}$$

$$\mathcal{Q}(0,3)=\begin{pmatrix}0&\chi&0&\chi^*\\\chi^*&0&\chi&0\\0&\chi^*&0&\chi\\\chi&0&\chi^*&0\end{pmatrix}\quad \mathcal{Q}(1,0)=\begin{pmatrix}1&0&0&0\\0&-1&0&0\\0&0&-1&0\\0&0&0&1\end{pmatrix}\quad \mathcal{Q}(1,1)=\frac{1}{\sqrt{2}}\begin{pmatrix}0&-i&0&1\\i&0&-1&0\\0&-1&0&i\\1&0&-i&0\end{pmatrix}$$

$$\mathcal{Q}(1,2)=\begin{pmatrix}0&0&-i&0\\0&0&0&i\\i&0&0&0\\0&-i&0&0\end{pmatrix}\quad \mathcal{Q}(1,3)=\frac{1}{\sqrt{2}}\begin{pmatrix}0&-i&0&-1\\i&0&1&0\\0&1&0&i\\-1&0&-i&0\end{pmatrix}\quad \mathcal{Q}(2,0)=\begin{pmatrix}1&0&0&0\\0&-1&0&0\\0&0&1&0\\0&0&0&-1\end{pmatrix}$$

$$\mathcal{Q}(2,1)=\begin{pmatrix}0&-\chi&0&\chi^*\\-\chi^*&0&\chi&0\\0&\chi^*&0&-\chi\\\chi&0&-\chi^*&0\end{pmatrix}\quad \mathcal{Q}(2,2)=\begin{pmatrix}0&0&-1&0\\0&0&0&1\\-1&0&0&0\\0&1&0&0\end{pmatrix}\quad \mathcal{Q}(2,3)=\begin{pmatrix}0&-\chi^*&0&\chi\\-\chi&0&\chi^*&0\\0&\chi&0&-\chi^*\\\chi^*&0&-\chi&0\end{pmatrix}$$

$$\mathcal{Q}(3,0)=\begin{pmatrix}1&0&0&0\\0&1&0&0\\0&0&-1&0\\0&0&0&-1\end{pmatrix}\quad \mathcal{Q}(3,1)=\frac{1}{\sqrt{2}}\begin{pmatrix}0&-1&0&-i\\-1&0&-i&0\\0&i&0&1\\i&0&1&0\end{pmatrix}\quad \mathcal{Q}(3,2)=\begin{pmatrix}0&0&i&0\\0&0&0&i\\-i&0&0&0\\0&-i&0&0\end{pmatrix}$$

$$\mathcal{Q}(3,3)=\frac{1}{\sqrt{2}}\begin{pmatrix}0&1&0&-i\\1&0&-i&0\\0&i&0&-1\\i&0&-1&0\end{pmatrix}$$

Spectrum of the symmetrized Heisenberg-Weyl basis

The expectation value of a squared HW observable is given by

$$\sqrt{\langle\mathcal{Q}^2(l,m)\rangle}=\sqrt{1+\mathrm{Im}\langle\mathcal{D}(2l,2m)\rangle}. \tag{3.69}$$

As the displacement operator is a unitary operator its eigenvalue is bounded by 1 and thus we need to take the imaginary part of the displacement operator. Therefore we have that

$$\sqrt{\langle\mathcal{Q}^2(l,m)\rangle}\leq|q_{\max}|\leq 2. \tag{3.70}$$

where $q_{\max}$ is the maximum eigenvalue. In the special case $d=2$, $\mathrm{Im}\,\mathcal{D}(2l,2,m)=0$, which is simply the case where the HW observables reduce to Pauli operators. For higher dimensions however $\mathrm{Im}\,\mathcal{D}(2l,2,m)$ is non-zero and the maximum eigenvalue is bounded by $\sqrt{2}$. The absolute values of the eigenvalues for $\sqrt{\mathcal{Q}^2(l,m)}$ for all HW observables are

$$|q_n|=\sqrt{1+\sin\frac{4\pi n}{d}} \tag{3.71}$$

for $n=0,\cdots,d-1$.

Chapter 4

Application: Multipartite Entanglement Detection for Qudits

In this chapter, we explore the question how to experimentally detect various notions of entanglement in multipartite qudit systems. Entanglement detection is a widely studied subject, however there is a tremendous gap between the theory of bipartite system and general multipartite systems [20] [15] [26]. Typically the bipartite qubit case is considered to be answered quite satisfyingly. There are easy to implement and strong entanglement criteria at hand. We will consider the well understood bipartite case as an special case and cover the topic always hoping to find more general answers. The chapter consist of material published in the articles [2, 27] during the writing of this thesis.

Apart from many technical difficulties, when studying multipartite qudit entanglement theory one is faced with the break down of the concept of entanglement & separability into sub-concepts. In fact there are at least combinatorially many different concepts of non-k-separability for higher particle numbers. Not even mentioning the dimensionality induced division for example into different LOCC classes or states with certified ranks. In other words, the general scenario opens up a plethora of potentially interesting properties to be identified, certified and at best quantified.

4.1 A framework of entanglement detection criteria: How to derive them and how to asses multipartite detection criteria?

For sure a fully general entanglement criterion is an almost unsolvable challenge due to [21]. There will not be one single entanglement criterion that captures and differentiates between all possible notions of entanglement. For this reason we will try to present an at least partially systematic framework of these notions, that allows to express and subsequently certify the various multipartite entanglement notions. We will see that there is a tremendous amount of freedom involved. Small details such as different norms, partitions and of course the exact notion of study influence the

performance of entanglement criteria. However before tackling these issues, we want to step back and ask again: *what is an detection criterion?*

4.1.1 What is a detection criterion?

The general form of a the linear detection criterion for a certain property is given by

$$\mathcal{E}_{property,W} = \max_{\rho_{\neg property}} \mathrm{Tr}(\rho_{\neg property} W) = \quad (4.1)$$

$$= \max_{\rho_{\neg property}} \sum_{i_1,\ldots,i_n} c_{i_1,\ldots,i_n} \mathrm{Tr}(\lambda_{i_1,\ldots,i_n} \rho_{\neg property}) = \max_{\rho_{\neg property}} \sum_{i_1,\ldots,i_n} c_{i_1,\ldots,i_n} \langle \lambda_{i_1,\ldots,i_n} \rangle_{\rho_{\neg property}}, \quad (4.2)$$

where $\rho_{\neg property}$ is a state not possessing the desired property. We expressed the observable W by a linear combination of the correlation tensor elements $\langle \lambda_{i_1,\ldots,i_n} \rangle$. If the results obtained by measuring the observable W exceed $\mathcal{E}_{property,W}$ then the systems state ρ must posses the property in question!

Typically one want to show, that a specific state possesses a desired property. In this case

$$\mathcal{E}_W(\rho) = \mathrm{Tr}(\rho W), \quad (4.3)$$

is also of importance. It is simply the expected value of ρ under the same observable W as used in the detection criterion $\mathcal{E}_{property,W}$.

Observation 4.1.1. *In order to be in principle able to certify a property of a certain state ρ by a detection criterion $\mathcal{E}(.)$. We need*

$$\mathcal{E}_W(\rho) - \mathcal{E}_{property,W} > 0. \quad (4.4)$$

The expectation value $\mathcal{E}_W(\rho)$ of the state under consideration needs to exceed the maximum possible on the set of states that do not possess the target property.

Generally the derivation of an entanglement criterion involves the maximization of an arbitrary chosen linear combination of observables on the set of all states not possessing the target property. By choosing $c_{i_1,\ldots,i_n} = \mathrm{sgn}(\langle \lambda_{i_1,\ldots,i_n} \rangle)$, we make a simplification that may fail to capture some properties, but it will be much easier to deal with as we will see later. We get

$$\sum_{i_1,\ldots,i_n} c_{i_1,\ldots,i_n} \langle \lambda_{i_1,\ldots,i_n} \rangle_{\rho_{\neg property}} = \sum_{i_1,\ldots,i_n} |\mathrm{Tr}(\lambda_{i_1,\ldots,i_n} \rho_{\neg property})|. \quad (4.5)$$

Now this is about everything, we can say about a completely general property. At this point we have to specify more closely, what we want to detect and exploit as much of the structure we get from this choice.

As we have already stated we are generally interested in entanglement criteria. It can be generally assumed in all derivations of entanglement criteria that some form of separability holds for the state $\rho_{\neg property}$. The exact type of separability we assume at this stage determines the type of entanglement we detect by the subsequently arising inequality's violation. In fact it turns out in the next Section 4.1.2, that the amount of achievable violation is of importance for the noise resistance of the criterion. How the assumptions on separability yield different detection criteria is illustrated after the following section in part 4.2.1.

4.1.2 Noise resistance of detection criteria

In the last section we have provided numerous inequalities of experimentally measurable quantities. Their violation in an experiment allows us to infer the presence of certain features of the underlying system. One could argue that in principle, this is all you need for the experimental verification of these features. In practice, however it is of great importance to not just find inequalities such as 4.4, but to find those that are additionally easy to implement. Up until now we have not mentioned: *ease of implementation.*

One concept that tries to capture the important aspect of usability is the so-called *noise resistance* of an detection criterion. Any state we write down in a mathematical fashion is necessarily an idealized concept. This is especially true for the case of pure states. Whenever we ask an experimenter to prepare a certain state, he will try to do this best to manufacture this state in his lab, but inevitably he will make smaller or larger errors in the preparation process. In the end, the physical state we will be working with will be a distorted version of the ideal state we had asked to prepare. To account for this we use a theoretical noise model, that tries to include the effect of inevitable implementation errors into the description. We have given the most standard way to model noise in the introduction: the mixture with white noise model (confer: 2.2.6).

Observation 4.1.2. *Let ρ be a state where we want certify a property.*
If we want to include white noise in our description, equation 4.4 changes to

$$R(\mathcal{E}_W, \rho) := \mathrm{argmin}_p(\mathcal{E}_W(\Delta_p(\rho)) - \mathcal{E}_{property,W} \geq 0) \tag{4.6}$$

Larger values of p convexly mix the target state with the maximally mixed state. The maximally mixed state is always a separable state. Thus mixing any entangled target state ρ_{target} with it will lead to a new state, that will be closer to the set of separable states.

In theory we could solve for p . This would give us the amount of white noise, where detection under criteria $\mathcal{E}$ first becomes impossible. In practice quantities like $\mathcal{E}_{entanglement,W} = \max_{\rho_{separable}} \mathrm{Tr}(\rho_{separable} W)$ are incomputable. The separable set has a complicated structure, so that generally efficient optimization is impossible. If we have a state independent upper bound for the incomputable quantity $\mathcal{E}_{entanglement,W} \leq \mathcal{E}'_{separable}$, we can obtain a smaller but state independent noise resistance $R(\mathcal{E}'_{separable}, \rho)$, that can be computed in practice. We have to accept a trade-off between computability and size of the noise resistance. A noise resistance always depends on a *lower* and an *upper* bound, the sharpness of the upper bound directly affects the amount of attainable noise resistance. This motivates why good upper bounds for $\mathcal{E}_{entanglement,W}$ are necessary.

4.2 Multipartite entanglement detection: The norm approach

4.2.1 A framework for deriving multipartite entanglement criteria

We have already mentioned in the introduction 2.2.7, that in multipartite entanglement theory various distinct notions of entanglement are studied. In this section, we illustrate how to derive

bounds, that can serve as entanglement criteria for any of those notions. What kind of entanglement is detected by a bound, depends entirely on the assumptions on separability used during the bounding procedure. Many choices are possible:

Consider for example

- **Biseparability criterion**
  We start with the least complicated notion of multipartite entanglement. Biseparability simply mirrors the notion of bipartite entanglement. As soon as there is any way to partition a system, we call it biseparable.

  Assume that ρ is a product state with respect to the partition A and $\bar{A}$, i.e. $\rho = \rho_A \otimes \rho_{\bar{A}}$. We can therefore write

$$\begin{aligned} \|\vec{\lambda}_{bisep}\|_1 &= \sum_i |\langle \lambda_A^i \otimes \lambda_{\bar{A}}^i \rangle| && (4.7)\\ &= \sum_i |\langle \lambda_A^i \rangle \langle \lambda_{\bar{A}}^i \rangle| && (4.8)\\ &= \langle |\vec{\lambda}_A| | |\vec{\lambda}_{\bar{A}}| \rangle && (4.9)\\ &\leq \|\vec{\lambda}_A\| \|\vec{\lambda}_{\bar{A}}\| := \mathcal{E}_{bisep} && (4.10) \end{aligned}$$

  where in the second line we used the assumption that ρ is product. As the right hand side depends on ρ in a convex way we can also conclude the above bound also holds for all convex combination of product states, i.e. for all states that are biseparable with respect to the partition A and $\bar{A}$.

- **non k-separability criterion**

$$\sum_{i_1,\ldots,i_n} |\operatorname{Tr}(\lambda_{i_1,\ldots,i_n}\rho_{k-sep})| = \sum_{i_1,\ldots,i_n} |\operatorname{Tr}(\bigotimes_{m=1}^{n}\lambda_{i_m}\bigotimes_{j=1}^{k}\rho_j)| = \sum_{i_1,\ldots,i_n} |\operatorname{Tr}(\bigotimes_{j=1}^{k}(\bigotimes_{i_m\in j}\lambda_{i_m})\rho_j)|. \tag{4.11}$$

  In the last equality we used that n is partitioned into k parts, so that the tensor product can be split accordingly. The separability of the state, always translates into the multiplicativity of the states norm.

$$\sum_{i_1,\ldots,i_n} |\operatorname{Tr}(\bigotimes_{j=1}^{k}(\bigotimes_{i_m\in j}\lambda_{i_m})\rho_j)| = \sum_{i_1,\ldots,i_n}\Pi_{j=1}^{k}\left|\operatorname{Tr}((\bigotimes_{i_m\in j}\lambda_{i_m})\rho_j)\right| = \sum_{i_1,\ldots,i_n}\Pi_{j=1}^{k}|\langle\lambda_{i_m\in j}\rangle_{\rho_j}|, \tag{4.12}$$

  Now all notions of k-separability leads to a factorization of the norm into k-parts.

$$\begin{aligned} \sum \Pi_{j=1}^{k}|\langle\lambda_{i_m\in j}\rangle_{\rho_j}| &= \left\langle \langle\lambda_{i_m\in k}\rangle_{\rho_k} | \Pi_{j=1}^{k-1}\langle\lambda_{i_m\in j}\rangle_{\rho_j}\right\rangle && (4.13)\\ &\leq \left\|\langle\lambda_{i_m\in k}\rangle_{\rho_k}\right\|_p \left\|\Pi_{j=1}^{k-1}\langle\lambda_{i_m\in j}\rangle_{\rho_j}\right\|_q := \mathcal{E}_{k-sep}, && (4.14) \end{aligned}$$

where p,q are conjugated norms that fulfill the condition $\frac{1}{p}+\frac{1}{q}=1$. Note that since we sum over the factorized norm, hence the quantity above can be formally written as the result of a scalar product. Scalar products allow the application of the Hölder inequality. We can choose to apply the Hölder inequality for any conjugated p-q norm. At the moment it is not clear, how to choose p,q in particular. For the moment we leave this question open, but we will return later during the chapter to discuss different possible choices. There is some more freedom aside from the norm considerations due to

$$\left\langle \langle \lambda_{i_m\in k}\rangle_{\rho_k} |\Pi_{j=1}^{k-1} \langle \lambda_{i_m\in j}\rangle_{\rho_j} \right\rangle = \left\langle \Pi_{j=k-1}^{k} \langle \lambda_{i_m\in k}\rangle_{\rho_k} |\Pi_{j=1}^{k-2} \langle \lambda_{i_m\in j}\rangle_{\rho_j} \right\rangle \tag{4.15}$$

$$= \ldots =$$

$$= \left\langle \Pi_{j=2}^{k} \langle \lambda_{i_m\in k}\rangle_{\rho_k} | \langle \lambda_{i_m\in j}\rangle_{\rho_j} \right\rangle \tag{4.16}$$

The Hölder inequality splits a scalar product into two parts. The k-fold product that results from k-separability can be distributed freely onto both sides of the scalar product. The different ways to partition the system have to be taken into account leading to a whole family of related equations. Since the systems are labeled and thus distinguishable, we can not simply count them via stars and bars. There are $k!$ permutations of the various systems, every permutation may be split in (k-1) different ways, leading to $(k-1)k!$ different partitions. Each partition leads in principle to a slightly different equation. In order to certify non-k-separability, we have to violate at least one of these inequalities.

We sum up the above equations and summarize into the following statement:

$$\left\| \vec{\lambda}_{k-separable} \right\|_1 = \sum_i \langle \lambda_A^i \otimes \lambda_B^i \otimes \ldots \otimes \lambda_K^i \rangle \tag{4.17}$$

$$= \sum_i \langle \lambda_A^i \rangle \langle \lambda_B^i \otimes \ldots \otimes \lambda_K^i \rangle \leq \|\lambda_A\|_p \|\lambda_B \ldots \lambda_K\|_q := \mathcal{E}_{k-sep}. \tag{4.18}$$

This tells us, that basically k-separability factorizes the 1-norm into product of conjugated p-q-norms. It is however only one way to partition the separable factors and we may not forget the others. For $1<l<k$ and an arbitrary permutation σ of the factors, we have the family of bounds

$$\left\| \vec{\lambda}_{k-separable} \right\|_1 \leq \| \prod_{i=1}^{l} \lambda_{\sigma(i)} \|_p \| \prod_{j=l}^{k} \lambda_{\sigma(j)} \|_q. \tag{4.19}$$

Any of these bounds may be violated to certify non-k-separability.

Many other possible partitions of the system $A \ldots K$ would be possible, we opted to separate a single system.

In our framework above, we have explicitly provided the upper bounds $\mathcal{E}_{bisep}$ 4.7 and $\mathcal{E}_{k-sep}$ 4.17 instead of $\max_{\rho SEPARABLE} \mathcal{E}_{property}((\rho)$. The sharper we bound in our framework, the lower these upper bounds get and the better our noise resistance becomes. This makes it important to keep their right hand sides as small as possible. Let us have a closer look at the objects on the right hand side of equation (4.7),

$$\|\vec{\lambda}_A\|_p = \sqrt[p]{\sum_{i\in\mathcal{X}} \langle \lambda_A^i \rangle^p} \tag{4.20}$$

where $\mathcal{X} \subseteq \mathcal{S}_{\mathcal{B}}$. With that in mind we can try to fix the various not yet determined parameters. Specifically we have to address how to choose the norm parameters p,q and the choice of basis λ_i in the above framework.

The choice of basis

In section 3.2 we discuss several sensible choices of basis. We review the canonical case of the generalized Gell-Mann matrices, the well-known Heisenberg-Weyl operators and we close by proposing a new construction of our own: the symmetrized Heisenberg-Weyl operators (published in:[2]).

The choice of norm: 1-norm & ∞-norm

The choice of p and q is relatively free in the biseparable case. We are more restricted in the k-separable case, if we want to pursue an analogous strategy, because the Hölder inequality is only applicable to $\|.\|_1$ norms. Since we want to iteratively apply the Hölder inequality on a k-partite system, we are bound to chose $p = \infty$ and $q = 1$ at least $k-2$ times , resulting in

$$\left\|\vec{\lambda}_{k-separable}\right\|_1 \leq \|\lambda_A\|_\infty \|\lambda_B \dots \lambda_K\|_1 \leq \underbrace{\|\lambda_A\|_\infty}_{maxEVof1stBasis} \quad \dots \quad \underbrace{\|\lambda_I\|_\infty}_{maxEVof(k-1)-thBasis} \|\lambda_J\|_p \|\lambda_K\|_q \tag{4.21}$$

Here it becomes apparent, that we are forced to frequently employ bounds involving the $\|.\|_\infty$ norm. This is rather unpleasant due to the relation

$$\|.\|_\infty \leq \|.\|_2 \leq \|.\|_1 . \tag{4.22}$$

At least for a given basis of an arbitrary subsystem S $\left(\lambda_S^i\right)_{i=0}^{d_S}$ the $\|.\|_\infty$ can be easily calculated, since it is simply the maximum eigenvalue of the basis. The maximum eigenvalue of any local basis will of course depend on the choice of basis. Here it becomes apparent why the choice of basis is important in multipartite entanglement detection. Finding bases with suitable flat spectra makes all bounds involving the $\|.\|_\infty$-norm sharper. In section 3.2 we calculate the maximal eigenvalues, in order to give concrete values for the equation 4.21.

Still, the last two norms remain always as unconstrained as in the bipartite case. Should one be in the lucky position to have a small maximal eigenvalue, the choosing the 1-norm and the ∞-norm seem good candidates. We will see however that also the 2-norm has appealing advantages, mainly the fact that 2-norms are inherently connected to the correlation tensor of the system. In 4.3 we further discuss the possibilities of 2-norms in entanglement detection.

Example: k-separability via GGM

We go on to illustrate the entanglement detection properties of the Gell-Mann basis. Now consider a matrix M that can be decomposed as a tensor product of matrices, element of any of the three

collections defining the GGM. Then the biggest eigenvalue of the square of such a matrix is given by

$$\max\lambda\left(M^2\right)=\left(\prod_{l=1}^{d-1}\left(\frac{dl}{l+1}\right)^{a_l}\right)\times\left(\frac{d}{2}\right)^{b+c} \tag{4.23}$$

where a, b and c are the numbers specifying how often a matrix of the diagonal (a_l), the anti-symmetric (b) and the symmetric (c) collection appear in the decomposition of M respectively. To achieve a full orthogonal basis we add the identity operator as $g_0^D := \mathbb{1}_d$ to the diagonal collection. This leaves equation (4.23) invariant as $\max\lambda\left(\mathbb{1}_d^2\right)=1$.

We provide explicit values on the k-separability 4.21 bound given above. $\|.\|_\infty$ can now be easily calculated, since we know that the Eigenvalues of the diagonal basis elements are larger than all others. Within the diagonals we also know the maximal eigenvalue. Above we calculated the ev of the squares of matrices, so below we can use the square root. We choose $p=2$ and $q=2$, due to the connection of the $\|.\|_2$ norm to the purity bound $\|\tau\|\leq\sqrt{d^n-1}$, we get

$$\left\|\vec{\lambda}_{k-separable}\right\|_1 \leq \underbrace{\|\lambda_A\|_\infty}_{\left(\frac{d(d-1)}{(d-1)+1}\right)^{\frac{n}{2}}} \ldots\|\lambda_I\|_\infty\|\lambda_J\|_2\|\lambda_K\|_2=(d_A-1)^{\frac{n}{2}}\ldots(d_I-1)^{\frac{n}{2}}(d_J^n-1)^{\frac{1}{2}}(d_K^n-1)^{\frac{1}{2}} \tag{4.24}$$

Example: k-separability of the GHZ

As an example we would like to know the maximum attainable by the GHZ state. The diagonal elements of the GHZ have a relatively ugly expression in arbitrary dimensions. To simplify we calculate only the offdiagonal part of GHZ.

Lemma 4.2.1. *Let ρ_{GHZ} denote a generalized n-partite qudit GHZ state with a decomposition $\rho_{GHZ}=\rho_{GHZ}^{diag}+\rho_{GHZ}^{off}$, dividing it into a diagonal part and $\rho_{GHZ}^{diag}=\frac{1}{d}\sum_{i=0}^{d-1}|i\ldots i\rangle\langle i\ldots i|$ an offdiagonal part $\rho_{GHZ}^{off}=\frac{1}{d}\sum_{i,j\neq 0}^{d-1}|i\ldots i\rangle\langle j\ldots j|$. Let $\lambda_{GHZ(n,d)}$ be the coefficient vector of the state, written in the generalized Gellman basis.*

$$|\left\{(\lambda_{GHZ(n,d)})_i\neq 0\right\}|=\sum_{k=0}^{\lfloor\frac{n}{2}\rfloor}\binom{n}{2k} \tag{4.25}$$

Proof. We proof the lemma in three steps.

- First we show that any nonzero offdiagonal element of the GHZ(n,d) is expressible in a tensor product basis containing no diagonal basis elements. We prove this by contradiction. Assume $\lambda_k^{D_l}$ to be a tensor product basis with a diagonal basis element D_l in the k-th term. For $l+1\leq j$ we have $\langle j|D_l=0$ in this case the following equation is zero due to the linearity of

the tensor product. Else we either get $\langle j|D_l = \langle j|$ or $\langle j|D_l = -l\langle j|$.

$$\mathrm{Tr}(\rho_{GHZ}^{off}\lambda_k^{D_l}) = \mathrm{Tr}((\frac{1}{d}\sum_{i,j\neq 0}^{d-1}|i\dots i\rangle\langle j\dots j|)\lambda_1\otimes\dots\otimes D_l\otimes\dots\otimes\lambda_n) =$$

$$= \mathrm{Tr}(\frac{1}{d}\sum_{i,j\neq 0}^{d-1}|i\dots i\rangle\langle j|\lambda_1\otimes\dots\otimes\underbrace{\langle j|D_l}_{\langle j|}\otimes\dots\otimes\langle j|\lambda_n = \mathrm{Tr}(\frac{1}{d}\sum_{i,j\neq 0}^{d-1}|i\dots i\dots i\rangle\langle i'\dots j\dots i'|) = 0 \tag{4.26}$$

Here the i' may or may not be equal to i, but we use that at least in the k-th tensor product we have $i \neq j$ forcing all elements to be offdiagonal. Thus there is no diagonal element that the trace could sum over and the above equation reduces to 0.

- Second we illustrate that the tensor product basis can only consist of basis elements with common support. For $n = 2$

$$\mathrm{Tr}(\rho_{GHZ}^{off}\lambda_k^{D_l}) = \mathrm{Tr}((\frac{1}{d}\sum_{i,j\neq 0}^{d-1}|ii\rangle\langle jj|)(|k\rangle\langle l| + |l\rangle\langle k|)\otimes(|k'\rangle\langle l'| + |l'\rangle\langle k'|) \tag{4.27}$$

$$= \mathrm{Tr}((\frac{1}{d}\sum_{i,j\neq 0}^{d-1}|ii\rangle\langle jj|)(|kk'\rangle\langle ll'| + |ll'\rangle\langle kk'|) \tag{4.28}$$

  can only be nonzero for $k = k'$ and $l = l'$. This means that the tensor product may feature only symmetric or antisymmetric basis elements which share their support. Higher n follow by induction.

- In the first two steps we argued that we only have to consider tensor product bases consisting of symmetric and antisymmetric local bases with the same support. Since the GHZ is invariant to transposition, we know that the correlation tensor elements of the antisymmetric elements have to be zero. Note however $\mathrm{Tr}(g_{kl}^S)\otimes Tr(g_{kl}^S) = \mathrm{Tr}(g_{kl}^A)\otimes Tr(g_{kl}^A)$, yielding again an symmetric basis element. This means that exactly the tensor products containing no or an even number of antisymmetric local bases are nonzero. Including 0, there are $\sum_{k=0}^{\lfloor\frac{k}{2}\rfloor}\binom{n}{2k}$ ways to pick an even number out of n, proving the claim.

□

Making use of the above lemma we can now calculate, the maximum norm we can attain by measuring only the off-diagonal contributions of a 3-partite qudit GHZ.

$$\left\|\lambda_{GHZ(3,d)}^{off}\right\|_2 \geq \sqrt{\frac{1}{d^2}\underbrace{(\frac{d}{2})^3}_{Eigenvalue}\underbrace{(\frac{d^2-d}{2}}_{No.2-dimensional subspaces}\underbrace{\sum_{k=0}^{\lfloor\frac{3}{2}\rfloor}\binom{3}{2k})}_{Equivalent Bases}} = \sqrt{(d^3-d^2)} \tag{4.29}$$

Explicitly writing equation 4.7 yields

$$\|\lambda_{2-sep}\|_1 \leq \|\lambda_{AB}\|_2\|\lambda_C\|_2 = \sqrt{(d^2-1)(d-1)} = \sqrt{d^3-d^2-d+1}, \tag{4.30}$$

the bound we need to certify non biseparability. Furthermore either by direct application of the multiplicativity we get

$$\|\lambda_{3-sep}\|_1 \leq (d-1)^{\frac{3}{2}} = \sqrt{(d^3 - 2d^2 + d)}, \tag{4.31}$$

or alternatively by our framework utilizing the Hölder inequality 4.21 we get the slightly better

$$\|\lambda_{3-sep}\|_1 \leq \|\lambda_A\|_\infty \|\lambda_B\|_2 \|\lambda_C\|_2 \sqrt{(\frac{d}{2})(d-1)^2} = \sqrt{\frac{d^3 - 2d^2 + d}{2}}, \tag{4.32}$$

bound for non triseparability. Evidently equation 4.29 violates the bounds for biseparability 4.30 as well as for triseparability 4.31 4.32 for all d, thus proving their utility for the general qudit case.

4.3 Multipartite entanglement detection: The correlation tensor approach

In this section we take up a line of work propagating the correlation tensor as a mean to derive entanglement detection criteria. We want to point out that this chapter is based and aims to continue the work for bipartite systems in [8, 9], that has been subsequently generalized to multipartite systems in [10]. (For other related approaches see for example [22, 40, 47, 48])

The treatment of multipartite system requires quite some notation. All notational conventions relevant to this section have been already covered in the chapter about Bloch vector decompositions for n-partite qudit systems (c.f. Section 3.1.3). We ask the reader to familiarize himself with the notation before proceeding.

4.3.1 Bounding 2-norms of correlation tensors: The purity bound

2-norms and the purity bound

We have the following connection of the correlation tensor 2-norm with squared density matrices. We know that the convex state space's boundary consists of pure states and any convex function attains its maximum on the pure states. If a state is pure, it is also a projector and thus squaring it leaves a pure state invariant. This is way the connection between the 2-norm and squared density matrices, can also bee seen as connecting norms of the correlation tensor to the pure states. To highlight this we name the following bound on all correlation tensor elements the **purity bound**.

Theorem 4.3.1. *(Purity bound)*
The correlation tensor $\mathcal{C}(\rho)$ of any state ρ fulfills

$$\|\mathcal{C}(\rho)\|_2 \leq \sqrt{d^n - 1}. \tag{4.33}$$

Proof. The set of all physical states is known to be convex. Therefore all convex functions of physical states attain their maximum on the boundary of the state space. In particular all p-norms with $p \geq 2$ are convex functions, so those norms can also only attain their maximum on the boundary.

This boundary consists exactly of the pure states. Using only normalization and purity we have

$$1 = \mathrm{Tr}(\rho) \geq \mathrm{Tr}(\rho^2) = \mathrm{Tr}((\frac{1}{d^n} \sum_{i_1,i_2,\cdots,i_n=0}^{d-1} c_{i_1,i_2,\cdots,i_n} \bigotimes_{k=1}^{n} \lambda_{i_k})^2). \tag{4.34}$$

We can always express ρ in a suitable basis that fulfills the normalization $\mathrm{Tr}(\lambda_i \lambda_j) = d\delta_{ij}$, we can simplify the above further by

$$(\frac{1}{d^n})^2 \sum_{i_1,i_2,\cdots,i_n=0}^{d-1} \mathrm{Tr}(c^2_{i_1,i_2,\cdots,i_n} \bigotimes_{k=1}^{n} \lambda^2_{i_k}) = (\frac{1}{d^n})^2 \sum_{i_1,i_2,\cdots,i_n=0}^{d-1} c^2_{i_1,i_2,\cdots,i_n} \mathrm{Tr}(\bigotimes_{k=1}^{n} \lambda^2_{i_k}) \tag{4.35}$$

$$= (\frac{1}{d^n})^2 \sum_{i_1,i_2,\cdots,i_n=0}^{d-1} c^2_{i_1,i_2,\cdots,i_n} \Pi_{k=1}^{n} \underbrace{\mathrm{Tr}(\lambda^2_{i_k})}_{d^n} = \frac{1}{d^n} \sum_{i_1,i_2,\cdots,i_n=0}^{d-1} c^2_{i_1,i_2,\cdots,i_n} = \|\mathcal{C}(\rho)\|_2^2 . \tag{4.36}$$

Note that $c_{0\ldots0}$ always describes the identity and can be subtracted to obtain the above relation. □

Note that equality in 4.33 can be achieved only for pure states. All mixed states can be represented as convex combinations of pure states, meaning that the mixed states have to lie in the interior of the state space. Therefore the inequality 4.33 will be always strict for mixed states.

Before going on we want to comment on the relevance of the above bound 4.33. The purity bound is our first example of an upper bound on the norm of the correlation tensor. We will continue the search for sharper upper bounds later in the Section 4.3.3. Within that section we will provide a series of increasingly sharper upper bounds, inspired by the problem of entanglement detection. In essence all of those upper bounds allow us to construct an entanglement witness, whose noise robustness increases depending on the sharpness of the bound, thus motivating our search for as tight bounds as possible. Section 4.3.2 takes the complementary approach of providing lower bounds to the correlation tensor norm. In addition, to being of technical use in sharpening the already provided upper bounds, the lower bounds will turn out to enforce limits on the sparsity structure of correlation tensors.

2-norms and LU-invariance

Due to the connection between trace and correlation tensor 4.34, we have

$$\mathrm{Tr}((U_1 \otimes \cdots \otimes U_n)\rho(U_1^\dagger \otimes \cdots \otimes U_n^\dagger))^2) = \mathrm{Tr}((U_1 \otimes \cdots \otimes U_n)(U_1^\dagger \otimes \cdots \otimes U_n^\dagger))\rho^2) \tag{4.37}$$

$$= \mathrm{Tr}(\rho^2) = \frac{1}{d^n} \sum_{i_1,i_2,\cdots,i_n=0}^{d-1} c^2_{i_1,i_2,\cdots,i_n}. \tag{4.38}$$

and can thus directly infer the local unitary invariance of the correlation tensor norm $\|\mathcal{C}(\rho)\|_2$ from the cyclicity of the trace as a simple corollary. Even though this observation is mathematically simple, it has a concrete practical implications. Intuitively speaking tensor products of local unitaries model every operation that can be applied locally to a system without destroying its entanglement properties. More concretely these include operation such as local rotations meaning that rotations

of a local reference frame will not change a LU-invariant quantity.

In one of our publications for this thesis [27] we have argued that the property of LU-invariance is of interest to experimentalists working with multipartite systems. In experiments with spatially distributed quantum systems the experimentalist typically has to make sure that measurements on separate systems have to be performed in comparison to the same reference frame. All systems have to be aligned by the experimentalist, often this is done by hand. This task can be handled comfortably for 2 systems by an experienced experimentalist, however if multiple systems are involved the task get's increasingly tedious. Each system has to be aligned pairwise with all other systems, thus we have quickly growing effort for more systems. LU-invariant quantities completely free the experimentalist from every alignment procedure between different systems, one only has to ensure that measurements are locally orthogonal, thus considerably reducing alignment effort as well as removing a possible source of imprecision due to alignment errors. We want to also point out the concurrently published article [40]. Both articles advocate the experimental benefits of LU-invariance, however they differ in the specific quantities used to detect multipartite entanglement. While our contribution [27] proposes norms of correlation tensors, [40] in contrast uses several related Bell-inequalities namely the Mermin, Mermin-Klyshko and Svetlichny inequalities.

The practical relevance of LU-variance was already demonstrated, since recently the 3-partite GHZ state has been successfully detected in an experimental setup without any common reference frame [49] by applying the entanglement witnesses proposed in our publication [27] as well as the Bell inequalities proposed in [40]. In addition to detecting general entanglement, our criterion has been used to certify genuinely multipartite entanglement in that experiment.

Even though we think that LU-invariance is a worthwhile property to have in mind while considering witnesses and measurement procedures for practical use, we have to caution that LU-operations do not cover the important case of losses during an experiment. If losses are to be modeled theoretically, they have to be modeled by the dissipative part of a master equation. For this reason LU-invariant quantities remain sensitive to losses.

It seems that the choice of 2-norms is essential in order to connect the Bloch decomposition to the correlation tensor elements. We are aware of no similar procedure for other norms. That only the 2-norm seems directly connected to the LU-invariant squares of correlation tensor norms follows from the fact that the 2-norm is the only p-norm that exhibits unitary invariance. The elementary fact

$$||x||_2^2 = \sum_i |x_i|^2 \leq \sum_i |x_i|^2 + 2\sum_{i\neq j} |x_i||x_j| = ||x||_1^2, \tag{4.39}$$

tells us that any bounds we derive with 2-norms could be sharper if we had used 1-norms instead. If we want to use 2-norms, this weakening of our bounds is the prize we have to pay. In return we obtained not only the purity bound 4.33 for the norms of correlation tensor elements, but also all bounds inherit the LU-invariance of the correlation tensor norm.

One may of course also consider different norms than the p-norm. For a discussion of entanglement detection via Ky Fan norms (i.e. norms of the k-largest eigenvalues) refer to [10].

4.3.2 Fundamental limits on the form of the correlation tensor

Schmidt decomposition implications

For a pure state $\rho = |\psi\rangle\langle\psi|$ the structure of its correlation tensor is in fact quite restricted. The purity of a system delivers together with the Schmidt decomposition, a convenient relation between the correlation tensor norms of high- and low- order terms.

Theorem 4.3.2. *Consider an n-partite system to be partitioned in two disjoint parts $n = \beta \dot{\cup} \beta'$, w.l.o.g. $|\beta| \geq |\beta'|$. Let the system be described by a pure state $\rho = |\psi\rangle\langle\psi|$.*

$$\|\mathcal{C}(\rho_\beta)\|_2^2 \geq d^{n-2|\beta'|} - 1. \tag{4.40}$$

Proof. Consider the marginals ρ_β,$\rho_{\beta'}$ obtained by tracing out a disjoint partition of the whole system. Since ρ is a pure state, by the properties of the Schmidt decomposition the spectrum of the marginal obtained by tracing out an any set of subsytems such as β, has the same spectrum as the marginal obtained by tracing over its complement β'. Therefore any function, that only depends on the spectrum will be the same on both marginals. Consider the function of the linear entropy

$$S_L(\rho) := 1 - \mathrm{Tr}(\rho^2), \tag{4.41}$$

it is obvious that this is a function of the eigenvalues of ρ only. Therefore we have by the Schmid decomposition that

$$S_L(\rho_\beta) = S_L(\rho_{\beta'}). \tag{4.42}$$

Working with the linear entropy S_L is convenient because it is defined by $\mathrm{Tr}(\rho^2)$. This is a quantity that we can easily handle for pure states. We have already shown a relation between the trace of a squared density matrix and its respective correlation tensor in the derivation of the purity bound (confer 4.34), applying this we get

$$\frac{1}{d^{|\beta|}}(\|\mathcal{C}(\rho_\beta)\|_2^2 + 1) = S_L(\rho_\beta^2) = S_L(\rho_{\beta'}^2) = \frac{1}{d^{|\beta'|}}(\|\mathcal{C}(\rho_{\beta'})\|_2^2 + 1) \tag{4.43}$$

The difference in partition size introduces an asymmetry in the bounding procedure. Without loss of generality we assumed $|\beta| \geq |\beta'|$. We lower bound the terms of the smaller partition trivially by $\|\mathcal{C}(\rho_{\beta'})\| \geq 0$. This choice of the smaller partition terms leaves us with a non trivial expression for the terms contained in the larger partition

$$\|\mathcal{C}(\rho_\beta)\|_2^2 \geq d^{|\beta|-|\beta'|} - 1 = d^{n-2|\beta'|} - 1. \tag{4.44}$$

□

Note that in the proof, we lower bound the terms of the smaller partition β' by $\|\mathcal{C}(\rho_{\beta'})\| \geq 0$. This is a sensible choice because the bound is tight at least for the maximally entangled state.

The theorem above is a recipe for getting lots of relations between sums of correlation tensor norms, each depending on the size of the appearing disjoint partitioning. The most simple relation is of course for a single system setting $\beta = j$ leads to,

$$\frac{1}{d} \leq \mathrm{Tr}(\rho_j^2) = \mathrm{Tr}(\rho_{\beta\setminus j}^2) = \frac{1}{d^{n-1}}\Big(\sum_{\gamma \subseteq \{1,\cdots,n\}/j} ||\mathcal{C}(\rho_\gamma)||_2^2 + 1\Big), \tag{4.45}$$

that shows us already one important consequence: sums of non-full body correlation tensor norms are always lower bounded by a nonzero constant.

To close we sharpen the above relation slightly. We put together the above relations to bound all the sums of the non-full body correlation tensors.

$$\sum_{|\beta|=n-i,i=1}^{\lceil \frac{n}{2} \rceil} \|\mathcal{C}(\rho_\beta)\| \geq \sum_{|\beta|=n-i,i=1}^{\lceil \frac{n}{2} \rceil} d^{|\beta|-|\beta'|} - 1 \tag{4.46}$$

$$= \sum_{\beta=1}^{\lceil \frac{n}{2} \rceil} \binom{n}{|\beta|} d^{n-2|\beta'|} - \underbrace{\sum_{\beta=1}^{\lceil \frac{n}{2} \rceil} \binom{n}{|\beta|}}_{2^{n-1}-1} = d^n \sum_{\beta=1}^{\lceil \frac{n}{2} \rceil} \binom{n}{|\beta|} d^{-2|\beta'|} - 2^{n-1} + 1 \tag{4.47}$$

What types of correlation tensors are possible?

We close this section by a short discussion what form a correlation tensor of a pure state can take in principle. Let us first repeat that the purity bound 4.33 is saturated for pure states and thus fixes the total amount of correlation in the system to be $\|\mathcal{C}(|\psi\rangle\langle\psi|)\| = d^n - 1$. Pure states can thus only differ in the way their correlations are distributed over the different sectors $\mathcal{C}_{|\alpha|}$. For example one might think of the two most extreme cases, a state of completely equidistributed correlations with $\|\mathcal{C}_\alpha(|\psi\rangle\langle\psi|)\| = \|\mathcal{C}_{\alpha'}(|\psi\rangle\langle\psi|)\| \, \forall \alpha, \alpha'$ or a state of completely concentrated correlations $\|\mathcal{C}_j(|\psi\rangle\langle\psi|)\| = d^n - 1$. Is it possible for states to attain these structures of correlations?

Due to 4.45, we know that $\sum_{\gamma \subseteq \{1,\cdots,n\}/j} ||\mathcal{C}(|\psi\rangle\langle\psi|_\gamma)||^2 + 1$ is nonzero, this restricts the form of a correlation tensor severely. For example it is not possible for a state to contain all of its sectors in only one sector j, since by the above also the complement $\beta \setminus j$ needs to contain at least some correlations. Thus we get a first no-go result stating: *in general there is no state of completely concentrated correlations*!

More generally we have the relation 4.43 showing us that for any partition of n into k and $n-k$ this reasoning yields a corresponding relation and thus rules out the possibility of states that have solely correlations between any fixed number k of parties, since the set complement of $n-k$ is forced to be nonzero as well.

4.3.3 Entanglement detection via correlation tensors

Detection via single tensors sectors: The full body approach

Our first result concerns the detection of non-k-separability from the 2-norm of the full body correlations alone.

Theorem 4.3.3. *For all k-separable states ρ_{k-sep} the 2-norm of the full body correlation tensor is bounded from above by*

$$||\mathcal{C}_{|\alpha|=n}(\rho_{k-sep})||_2 \leq \sqrt{(d^2-1)^k (d^{\lceil \frac{n}{k} \rceil - 2})^R (d^{\lfloor \frac{n}{k} \rfloor - 2})^{k-R}} \tag{4.48}$$

with $R = n - k\lfloor \frac{n}{k} \rfloor$.

Whenever considering multipartite entanglement, one is faced with many different possible partitions of the whole systems that could possibly be separable. Usually the arising combinatorial

considerations follow the same scheme. This is why we collect them into the following straightforward but useful lemma. In essence the equipartition lemma describes for any k-partition of a set, which partition maximizes the norm, depending only on a few basic properties of the upper bounds on the individual norms. Generically such multiplicative bounds favor equipartitions, this is why we name it the **equipartition lemma**.

Lemma 4.3.1. (Equipartition lemma) Let ρ be a k-separable state and let us denote the correlation tensors of its marginals by $\mathcal{C}(\rho_{\beta_1}), \ldots, \mathcal{C}(\rho_{\beta_k})$. If we have an upper bound $f(|\beta|)$ on the norm of a marginals correlation tensor $||\mathcal{C}(\rho_\beta||$ depending on the cardinality $|\beta|$ of the multi-index β, $y \in \mathbb{N}$ that satisfies $f(|\beta|)^2 \geq f(|\beta|+y)f(|\beta|-y)$, then we can bound norms of the k partitions correlation tensor β_i by

$$||\mathcal{C}(\rho_{\beta_1})|| \times ||\mathcal{C}(\rho_{\beta_2})|| \times \cdots \times ||\mathcal{C}(\rho_{\beta_k})|| \leq \sqrt{(f(\lceil \frac{n}{k} \rceil))^R (f(\lfloor \frac{n}{k} \rfloor))^{k-R}}, \tag{4.49}$$

with $R = n - k\lfloor \frac{n}{k} \rfloor$.

Proof. Let $|\beta| = x$,$|\beta'| = x + y$,$|\beta''| = x - y$ and γ arbitrary large for $x, y \in \mathbb{N}$.

Then for any function of a tensor norm with $f(x)^2 \geq f(x+y)f(x-y)$ we trivially get

$$\frac{||\mathcal{C}(\rho_{\beta'})|| \times ||\mathcal{C}(\rho_{\beta''})|| \times ||\mathcal{C}(\rho_\gamma)||}{||\mathcal{C}(\rho_\beta)|| \times ||\mathcal{C}(\rho_\beta)|| \times ||\mathcal{C}(\rho_\gamma)||} \leq \frac{f(x+y)f(x-y)||\mathcal{C}(\rho_\gamma)||}{f(x)^2||\mathcal{C}(\rho_\gamma)||} \leq 1.$$

Therefore k partitions of equal size yield the maximal value. However if n/k is not a natural number we may not be able to choose all partitions to be equal. Note that the above case includes all partitions where the difference in size between two partitions $||\beta'| - |\beta''|| = 2y$ is larger than one and we can apply the observation above.

However we have yet to discuss the special case of the partition sizes differing by one. All partitions of the largest partitioning, have to be either of size $\lfloor \frac{n}{k} \rfloor$ or $\lceil \frac{n}{k} \rceil$, else we could apply the above argument. $R = n - k\lfloor \frac{n}{k} \rfloor$ is the remainder of the division of n by k. In order to make sure all partitions sum up to n exactly R can be larger by one element than the remaining ones, proving our claim. □

The purity bound is by no means the only possible function that we can insert above for f, we will provide several different sensible choices later in the text, still we start by using (4.33) as the simplest choice obtaining our first corollary.

Corollary 4.3.1. Let ρ be a k separable state with marginals $\rho_{\beta_1}, \ldots, \rho_{\beta_k}$. Then we can bound the norms of the k marginals correlation tensors by the purity bound (4.33) obtaining

$$||\mathcal{C}(\rho_{\beta_1})|| \times ||\mathcal{C}(\rho_{\beta_2})|| \times \cdots \times ||\mathcal{C}(\rho_{\beta_k})|| \leq \sqrt{(d^{\lceil \frac{n}{k} \rceil} - 1)^R (d^{\lfloor \frac{n}{k} \rfloor} - 1)^{k-R}} \tag{4.50}$$

with $R = n - k\lfloor \frac{n}{k} \rfloor$.

Proof. The corollary follows from (4.50) by using the purity bound (4.33) $f(x) = d^x - 1$. We calculate $f(x)^2 = d^{2x} - d^x + 1 \geq d^{2x} - d^{x+y} - d^{x-y} + 1 = f(x+y)f(x-y)$ thus checking that our condition on f is fulfilled. □

The upper bound above is derived only using the assumption of purity and k-separability. We emphasize that the norms above are convex and the maximum is therefore taken at the boundary of the set of states, the pure states. Any mixed state lies in the interior of the set of states and could only achieve smaller values. This means that if a states full body correlation tensor violates the above inequality, this state can not be k-separable.

Since the 2-norm of the correlation tensor is convex in the space of density matrices it is sufficient to bound the maximal norm for a k-separable pure state. For those the norms are multiplicative under tensor products, so one only needs to find the maximal 2-norm of general states and then find the maximal product among all possible k-partitions of n.

Example 4.3.4. *(Detection of the GHZ-state) We show that the three dimensional GHZ state* $|GHZ_3\rangle = \frac{1}{\sqrt{3}}|000\rangle + |111\rangle + |222\rangle$ *is detectable only measuring its full body correlation tensor. We provide the correlation tensor of the GHZ(3,3) in the generalized Gell-Mann Basis, denoting the diagonal basis elements by* $D(i)$*, the symmetric basis elements by* $S(j)$ *and the antisymmetric basis elements by* $A(k)$*.*

Value	*Basis Elements*		
2-body sector $\mathcal{C}_2(GHZ(3,3))$			
1	*D(1)*	*D(1)*	$\mathbb{1}$
1	*D(1)*	$\mathbb{1}$	*D(1)*
1	$\mathbb{1}$	*D(1)*	*D(1)*
1	*D(2)*	*D(2)*	$\mathbb{1}$
1	*D(2)*	$\mathbb{1}$	*D(2)*
1	$\mathbb{1}$	*D(2)*	*D(2)*
$\|\mathcal{C}_2\|_2^2 = 6$			
full-body sector $\mathcal{C}_3(GHZ(3,3))$			
-1.2247	*A(1)*	*S(1)*	*S(1)*
-1.2247	*S(1)*	*A(1)*	*S(1)*
-1.2247	*A(2)*	*S(2)*	*S(2)*
-1.2247	*S(2)*	*A(2)*	*S(2)*
-1.2247	*S(1)*	*S(1)*	*A(1)*
1.2247	*A(1)*	*A(1)*	*A(1)*
0.7071	*D(2)*	*D(1)*	*D(1)*
0.7071	*D(1)*	*D(2)*	*D(1)*
-1.2247	*A(3)*	*S(3)*	*S(3)*
-1.2247	*S(3)*	*A(3)*	*S(3)*
-1.2247	*S(2)*	*S(2)*	*A(2)*
1.2247	*A(2)*	*A(2)*	*A(2)*
-1.2247	*S(3)*	*S(3)*	*A(3)*
1.2247	*A(3)*	*A(3)*	*A(3)*
0.7071	*D(1)*	*D(1)*	*D(2)*
-0.7071	*D(2)*	*D(2)*	*D(2)*
$\|\mathcal{C}_3\|_2^2 = 20$			

Now let us evaluate the criterion 4.50 for the GHZ(3,3) state. In the 3 dimensional case biseparability is the only possible form of separability thus setting $k = 2$ *is the only relevant case. A*

2-separable three partite qutrit state can only achieve a correlation tensor norm of

$$\|\mathcal{C}(\rho_{2-sep}(3,3))\|_2^2 \leq (3^{\lceil \frac{3}{2} \rceil} - 1)^1 (3^{\lfloor \frac{3}{2} \rfloor} - 1)^{2-1} = (3^2 - 1)(3 - 1) = 16 \tag{4.51}$$

with the remainder term $R = 3 - 2\left\lfloor \frac{3}{2} \right\rfloor = 1$. *In contrast it is evident from the table above that already the full-body correlation tensor norm attains* $\|\mathcal{C}_3(GHZ(3,3))\|_2^2 = 20$, *thus exceeding* $\|\mathcal{C}(\rho_{2-sep}(3,3))\|_2^2 = 16$ *and thereby showing that the GHZ(3,3) is genuinely multi-partite entangled. It is not necessary to measure the 2-body correlations* $\|\mathcal{C}_2(GHZ(3,3))\|_2^2 = 6$.

These bounds enable reliable and robust detection of entanglement for general n and even detection of GME for $n = 3$. However most multipartite entangled states can not be revealed from full-body correlations alone and indeed this theorem fails at providing a detection criterion for verifying GHZ, Cluster or W-state type entanglement for $n > 3$. Hence, when using the correlation tensor norm for the case $k = 2$, i.e. the detection of GME, we need to include lower order correlation terms as well.

Improving the purity-bound: Extending the full body correlation approach to arbitrary single sectors

We try to improve the purity bound (4.33) in this section. We start by splitting the whole correlation tensor norm $\mathcal{C}(\rho)$ of a n-partite state into its sectors $\mathcal{C}_{|\alpha|=i}(\rho)$. Now consider the relation between the full body correlation tensor $\mathcal{C}_{|\alpha|=n}(\rho)$ and the lower sectors

$$||\mathcal{C}_{|\alpha|=n}(\rho)||^2 = d^n - 1 - \sum_{i<n} ||\mathcal{C}_{|\alpha|=i}(\rho)||^2. \tag{4.52}$$

Any entry in the lower order correlation sectors $\mathcal{C}_{|\alpha|=i}(\rho) i < n$ that has nonzero norm will allow us to subtract something from $d^n - 1$ in 4.52. We know that there have to be some nonzero entries in $\sum_{i<n} ||\mathcal{C}_{|\alpha|=i}(\rho)||^2$, since setting all of them zero would require pure states with all correlations concentrated in the full-body correlation tensor. That this is not possible is a simple consequence of the fact that the entropy of different partitions of pure states need to be equal and thus the higher order correlations below n, can generically not all be zero. We have already discussed why this is the case, check 4.43 in the chapter Schmidt decomposition implications for more details.

At least for the $n - 1$ body correlation tensor we can say something non trivial by using the Schmidt decomposition. We establish a first improved bound out of the purity condition.

Theorem 4.3.1. For the full body correlation tensor $\mathcal{C}_{|\alpha|=n}(\rho)$ of a n-partite state ρ,

$$||\mathcal{C}_{|\alpha|=n}(\rho)||^2 \leq d^{n-2}(d^2 - 1). \tag{4.53}$$

Proof. As a consequence of the Schmidt decomposition, any two bipartitions of a pure state have the same entropy, regardless of the entropy we use. In our case it proves to be helpful to consider the special case of linear entropy yielding

$$2(1 + \mathrm{Tr}(\rho_{\alpha/j}^2)) = S_L(\rho_{\alpha/j}) = S_L(\rho_j) = 2(1 + \mathrm{Tr}(\rho_j^2)). \tag{4.54}$$

The choice of the linear entropy is motivated by the fact that a correlation tensor can be expressed in terms of $\mathrm{Tr}(\rho^2) = \frac{1}{d}(1 + ||T||^2)$ due to (4.34), allowing us to write out the condition that ρ is a

pure state

$$\frac{1}{d} \leq \mathrm{Tr}(\rho_j^2) = \mathrm{Tr}(\rho_{\alpha/j}^2) = \frac{1}{d^{n-1}}(\sum_{\beta \subseteq \{1,\cdots,n\}/j} ||c_\beta||^2 + 1), \tag{4.55}$$

or rewriting the above to bound the lower order correlation tensor elements

$$\sum_{\beta \subseteq \{1,\cdots,n\}/j} ||c_\beta||^2 \geq d^{n-2} - 1. \tag{4.56}$$

We insert in (4.52) obtaining $||\mathcal{C}_{|\alpha|=n}(\rho)||^2 \leq d^{n-2}(d^2-1)$. □

This stronger bound can be inserted in the proof of the equipartition bound (4.49) to obtain a stronger result.

Corollary 4.3.2. If we have a k separable n particle state ρ. We can bound the norms of the correlation tensors of the k marginals ρ_{β_i} by (4.33) obtaining

$$||\mathcal{C}(\rho_{\beta_1})|| \times ||\mathcal{C}(\rho_{\beta_2})|| \times \cdots \times ||\mathcal{C}(\rho_{\beta_k})|| \leq \sqrt{(d^2-1)^k (d^{\lceil \frac{n}{k} \rceil - 2} - 1)^R (d^{\lfloor \frac{n}{k} \rfloor - 2} - 1)^{k-R}} \tag{4.57}$$

with $R = n - k\lfloor \frac{n}{k} \rfloor$.

Proof. The corollary follows from (4.50) by using (4.53). We calculate $f(n)^2 = d^{2n-4}(d^2+1)^2 = d^{2n+y-y-4}(d^2+1)^2 = f(n+y)f(n-y)$ thus checking that our condition on f is fulfilled. □

Using some basic combinatorial argument it is possible to further improve the bound by incorporating finite size effects:

Theorem 4.3.2. For a n-partite system ρ, all non full body correlation tensors $\mathcal{C}_{|\beta|}(\rho)$ with $|\beta| < n$ fulfill

$$||\mathcal{C}_{|\beta|}(\rho)||^2 \leq d^n - 1 - \frac{n(d^{n-2}-1)}{(n-1)} \tag{4.58}$$

Proof. Applying (4.56) we get

$$\sum_{j=1}^{n} \sum_{\beta \subseteq \{1,\cdots,n\}/j} ||c_\beta||^2 \geq n(d^{n-2}-1), \tag{4.59}$$

we turn the above to into an improvement by observing

$$\sum_j \sum_{\beta \subseteq \{1,2,\cdots,n\}/j} ||c_\beta||^2 = \sum_{|\beta|=n-1} ||c_\beta||^2 + \sum_{|\beta|=n-2} 2||c_\beta||^2 + (\cdots) + \sum_{|\beta|=1} (n-1)||c_\beta||^2. \tag{4.60}$$

That the equality above holds can be seen by a short combinatorial argument. Some of the terms in $\sum_{\beta \subseteq \{1,2,\cdots,n\}/j} ||c_\beta||^2$ will occur multiple times while summing over j. Let us consider the i-th coefficient c_{β_i} with $|\beta_i| = k$, this index will occur once in every of the inner sums $\sum_{\beta \subseteq \{1,2,\cdots,n\}/j} ||c_\beta||^2$ with $\beta_i \subseteq \beta$. To count the β containing β_i we fix the k components equal

to β_i out of $|\{1,2,\cdots,n\}/j| = n-1$, leaving us to to pick $n-1-k$ elements. The components of β can take n different values, but they may not repeat themselves, by fixing k of them we are left to choose from $n-k$ elements. Clearly this means that the an index of cardinality k will occur $\binom{n-k}{n-k-1} = n-k$ times proving our observation.

Combining (4.59) and (4.60) shows

$$\frac{n(d^{n-2}-1)}{(n-1)} \leq \sum_{|\beta|=n-1} \frac{1}{(n-1)}||c_\beta||^2 + \sum_{|\beta|=n-2} \frac{2}{(n-1)}||c_\beta||^2 + (\cdots) + \sum_{|\beta|=1} ||c_\beta||^2 \leq \sum_{|\beta|<n} ||c_\beta||^2, \tag{4.61}$$

which by inserting in (4.52) proves our theorem. □

This is of course only an improvement for finite n, that is better for small n. Asymptotically it scales equally to the above bound in eq.(4.53).

In any case we have found two stronger bounds (4.53),(4.59) than our initial purity bound (4.33). We can apply these instead of the weaker bound in (4.50).

Detection via sums of correlation tensor sectors

In the first two sections, we have initially derived bounds for the full-body correlation tensor and went on to generalize them to allow bounding any given single sector of the correlation tensor. This serves for developing witnesses of states that have concentrated full-body correlations or at least correlations concentrated in one sector. However concentrating all correlations in only the full body tensor, is a very extreme case of concentrated correlations and not achievable beyond $n = 3$. This no-go result on the form of the correlation tensor is a consequence of the Schmidt decomposition and has been discussed at length in Section 4.3.2. This motivates our second approach, where we derive bounds for states exhibiting most correlations in the "higher" tensors, but not only in a single one. For an example of such a state refer to the case of the AME state 4.3.7, which we discuss at the very end of this section.

Definition 4.3.5. Let ρ be the density matrix of an n-partite system, then we define the **cutoff function** $\mathcal{C}_{(l,m)}$ and the corresponding **cutoff parameters** $l < m \leq n \in \mathbb{N}$ by

$$\mathcal{C}_{(l,m)}(\rho) := \sum_{l=x}^{m} \sum_{|\alpha|\geq m} |c_\alpha|^2 \tag{4.62}$$

Theorem 4.3.6. *For $0 \leq x < n$,the cutoff $x \in \mathbb{N}$,c_α the correlation tensor elements of $\mathcal{C}(\rho)$, α an multi-index, ρ a n partite biseparable (under the partition $\{\beta, \overline{\beta}\}$) qudit system, $|\beta| = k_1$ and $|\overline{\beta}| = k_2$ with $k_1 + k_2 = n$, the inequality*

$$\mathcal{C}_{(n,x)}(\rho) = \sum_{m=x}^{n} \sum_{|\alpha|=m} |c_\alpha|^2 = (d^{k_1}-1)(d^{k_2}-1) + \sum_{k_j \geq x} (d^{k_j}-1) \tag{4.63}$$

holds.

Proof. We split the sum of all correlation tensors into those containing at least an element of β as well as $\overline{\beta}$ and those only containing elements from their respective partition.

$$\sum_{m=x}^{n}\sum_{|\alpha|=m}|c_\alpha|^2=\sum_{m=x}^{n}(\sum_{|\alpha|=m\wedge\alpha\subseteq\beta}|c_\alpha|^2+\sum_{|\alpha|=m\wedge\alpha\subseteq\overline{\beta}}|c_\alpha|^2+\sum_{|\alpha|=m\wedge\alpha\not\subseteq\beta\wedge\alpha\not\subseteq\overline{\beta}}|c_\alpha|^2) \tag{4.64}$$

We use the multiplicativity of the tensor norm, the purity bound (4.33) and the biseparability bounding

$$\sum_{m=x}^{n}\sum_{|\alpha|=m\wedge\alpha\not\subseteq\beta\wedge\alpha\not\subseteq\overline{\beta}}|c_\alpha|^2\leq\mathcal{C}(\rho_\alpha)=\mathcal{C}(\rho_\beta\otimes\rho_{\overline{\beta}})\leq(d^{k_1}-1)(d^{k_2}-1). \tag{4.65}$$

However for the terms not containing an index of β and $\overline{\beta}$, we can not make use of the separability. We can still apply the purity bound though. Take note that (4.33) can be used for sums of tensors, therefore bounding the sum of a multi-index and all other terms contained in the multi-index by a single application of the purity bound. In each bipartition there is one maximal remainder term of cardinality $n-k_i$ containing all other terms. Therefore for each bipartition we get a bound

$$\sum_{m=x}^{n}\sum_{|\alpha|=m\wedge\alpha\subseteq\beta}|c_\alpha|^2\leq\mathcal{C}(\rho_\beta)\leq\begin{cases}(d^{k_1}-1)\,\text{for}\,k_1\geq x\\0\,\text{for}\,k_1<x\end{cases} \tag{4.66}$$

$$\sum_{m=x}^{n}\sum_{|\alpha|=m\wedge\alpha\subseteq\overline{\beta}}|c_\alpha|^2\leq\mathcal{C}(\rho_{\overline{\beta}})\leq\begin{cases}(d^{k_2}-1)\,\text{for}\,k_2\geq x\\0\,\text{for}\,k_2<x\end{cases} \tag{4.67}$$

□

If we use the trivial cut-off $x=0$ this results in

$$\sum_{m=0}^{n}\sum_{|\alpha|}|c_\alpha|^2\leq(d^{k_1}-1)(d^{k_2}-1)+(d^{k_1}-1)+(d^{k_2}-1)=d^n-1$$

that is unfortunately a trivial result for all pure states. To avoid the trivial bound above we have considered the possibility of disregarding some of the lower order correlation tensors. First we note that by (4.33) we get for example $||\mathcal{C}(\rho_{AB})||^2\leq(d^2-1)$ as well as for $||\mathcal{C}(\rho_{AB})||^2+||\mathcal{C}(\rho_A)||^2+||\mathcal{C}(\rho_B)||^2\leq(d^2-1)$. The purity condition bounds a complete sum of correlation tensors, including all multi indices up to a chosen order, therefore the erasure of lower order correlation terms does not affect the purity bound. It is important to omit all terms containing a given index of the remainder term in order to achieve a reduction after application of the purity bound. For k_i smaller than x_i we simply omit the corresponding positive term $(d^{k_i}-1)$ obtaining our statement.

We proceed with a quick discussion on the role of the cutoff parameter x and its relation to the size of the two partitions k_1,k_2. We assume a bipartite state, therefore we only have to consider the three cases

- $x<k_1,k_2$: Since x is smaller than any partition size and the the purity bound is invariant to the omission of lower order correlation tensor elements, we only arrive at the trivial purity bound d^n-1.

- $k_1 \leq x < k_2$: In this case (4.63) reads $d^n - d^{k_1}$. The maximal value of this bound is achieved in the case of $k_1 = 1$ and $k_2 = n - 1$ yielding $d^n - d$.
- $k_1, k_2 \leq x$: (4.63) reads $d^n - d^{k_1} - d^{k_2} + 1$, this value is maximized whenever $|k_1 - k_2|$ is minimal $d^n - d^{\lfloor \frac{n}{2} \rfloor} - d^{\lceil \frac{n}{2} \rceil} + 1$.

Since the general entanglement detection problem is well-known to be NP-hard, no single detection strategy is generally expected to enable detection of all entangled states. Like for all NP-hard problems the usual approach is to consider special cases, that may well be tractable. Thus whenever a new detection procedure is proposed, it has to be analyzed witch class of states are suited for detection. In the case it is easy to give a general heuristic what type of states are well detected, while providing concrete examples is harder.

Clearly the simple idea behind 4.63 is to truncate the lower order sectors of the correlation tensor. Truncation is of course the most convenient, when the truncated quantities are smallest. So intuitively it is clear that 4.63 will work best the less correlations are contained within the lower order correlations. To name this concept, we will call a state that has more sectors of correlation set to zero (or close to zero) than another a state, a *state of more concentrated correlations*. This has to be contrasted with the results found in 4.3.2. There we show that the most extreme example of this class; a state containing only correlations in one sector is not existent for general n! This should not be to discouraging though, consider for example the following example:

Definition 4.3.7. An (n, d) **absolutely maximally entangled (AME(n,d)) state**, is defined as a pure n-party state of local dimension d, with all reduced density matrices equal to the appropriate maximally mixed state for every partition that is smaller or equal than $\frac{n}{2}$.

We give the example of the class of AME(n,d) states, since it is by definition a state of at least some what concentrated correlations and as such it is to be expected that a criterion such as 4.63 can be applied to it. Moreover it is a also a state, that has been known to be of practical importance for parallel teleportation and quantum secret sharing and where existence results are readily available [23]. This makes the AME(n,d) a natural candidate for benchmarking our criterion with a well studied and non trivial state. Briefly summarizing these existence results it can be said that there is a big gap between the qubit and the qudit case.

AME states do not exist for qubits except for very few cases. Existence has been proven for AME(2,2) also known as EPR state, AME(3,2) also known as the GHZ state, AME(5,2) and AME(6,2), while non-existence is known for AME(4,2) as well for all $n \geq 8$. The qubit case is therefore almost completely characterized with the exception of the AME(7,2).

In contrast, even though we are aware of no systematic study for the qudit case, there exists a constructive method that guarantees existence for all n partite system if d can be chosen large enough. (all details on applications or existence results can be found again in [23])

Example 4.3.8. *We know that for $|\alpha| \leq \frac{n}{2}$ the correlation tensor elements satisfy $||\mathcal{C}(\rho_\alpha^{AME(n)})|| = ||\operatorname{Tr}(\lambda_{\alpha_1} \otimes (\cdots) \otimes \lambda_{\alpha_{|\alpha|}} \frac{1}{d^{|\alpha|}} \mathbb{1}_{|\alpha|})|| = 0$ by definition of the AME. The AME is pure, thus*

$$1 = \operatorname{Tr}(\rho_{AME}^2) = \frac{1}{d^n}(1 + \sum_{m > \frac{n}{2}} \sum_{|\alpha|=m} ||c_\alpha||^2 + \sum_{m \leq \frac{n}{2}} \sum_{|\alpha|=m} ||c_\alpha||^2) = \frac{1}{d^n}(1 + \sum_{m > \frac{n}{2}} \sum_{|\alpha|=m} ||c_\alpha)||^2 + 0) \tag{4.68}$$

shows that all nonzero correlations are the correlation tensor elements of size bigger than $\frac{n}{2}$ and $\sum_{m>\frac{n}{2}} \sum_{|\alpha|=m} ||c_\alpha||^2 = d^n - 1$. This illustrates why our cutoff theorem (4.63) is sensible, setting $x = 0$ results as well in the trivial bound $d^n - 1$ in (4.63), since this coincides with the above we can not detect entanglement without an appropriately chosen x. However if x is large enough 4.63 provides a smaller bound than (4.68), allowing us to detect the AME. x is large enough, whenever it is chosen larger than the size of the smallest partition. By choosing $x = \frac{n}{2}$ we can ensure that there exists a partition smaller or equal than x.

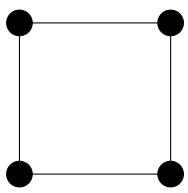
In the discussion following (4.63) we state the possible values this bound can achieve, depending on the size of x, k_1, k_2. For $x = \frac{n}{2}$ the first case is excluded and the maximal remaining value that can be achieved is $d^n - d$ in the case $k_1 = 1 < x < k_{n-1}$. Comparing the correlation tensor norm of the AME state $d^n - 1$ to the maximal correlation tensor norm achievable by a bipartite state $d^n - d$ we get a total separation of $d - 1$, allowing for detection of the AME state. Since the separation involves only the local dimension of the AME state d but not the number of involved parties n, it is immediately apparent that the criterion will be strongest whenever few systems of high dimension are studied.

We conclude by a short remark on the full body correlation tensor of the AME, thereby motivating what led us to derive the above (4.63). The first bounds (4.53),(4.59) we derived are suited for detection of states, whose correlation are concentrated in the full body part. Contrary to our initial hope, the AME is not in general such a state since by the Schmidt decomposition we have

$$\frac{1}{d} = \mathrm{Tr}((\rho_j^{AME})^2) = \mathrm{Tr}((\rho_{\alpha/j}^{AME})^2) = \frac{1}{d^{n-1}}(1 + \sum_{\alpha/j} ||c_{\alpha/j}^2||), \tag{4.69}$$

illustrating $\sum_{\alpha/j} ||c_\alpha^2|| = d^{n-2} - 1$, due to this we know that many of the lower order correlation elements are nonzero. This makes it plausible that AME states are not in general detected by our bounds containing just the full body correlation tensor (4.53),(4.59), but only by bounds on sums of correlation tensor elements such as 4.63.

Example 4.3.9. *(Detection of a Graph-State)*
In this example we will study the case of a 4 qubit graph state on a square lattice.



Graph states are stabilizer states and therefore well suited as example states for studying the correlation tensor norm since every stabilizer forces an entry in the correlation tensor to be equal to one. It suffices to list all stabilizers of a stabilizer state to determine exactly its correlation tensor. We list all nonzero correlation tensor entries of the 4-qubit rectangle graph state, sorted according to sector. Each corresponds exactly to one stabilizer and assumes by definition of the stabilizer state

Sector of Correlation Tensor	*full-body* $\mathcal{C}_4$	*3 body* $\mathcal{C}_3$	*2 body* $\mathcal{C}_2$
	(X,X,X,X)	*(Z,X,Z,𝟙)*	*(X,𝟙,X,𝟙)*
	(Y,Y,Z,Z)	*(𝟙,Z,X,Z)*	*(𝟙,X,𝟙,X)*
	(Z,Y,Y,Z)	*(Z,𝟙,Z,X)*	
	(Z,Z,Y,Y)	*(X,Z,𝟙,Z)*	
a value of 1.	*(Y,Z,Z,Y)*	*(Y,X,Y,𝟙)*	
		(𝟙,Y,X,Y)	
		(Y,𝟙,X,Y)	
		(X,Y,𝟙,Y)	
Sum of Correlations	$\|\mathcal{C}_4\|_2 = 5$	$\|\mathcal{C}_3\|_2 = 8$	$\|\mathcal{C}_2\|_2 = 2$

We calculate the cut-off function $\mathcal{C}_{(n,n-1)}$. *W.l.o.g.* $k_1 \geq k_2$, *let* $x \geq k_2$,

$$\mathcal{C}_{(n,n-1)}(\rho) = \sum_{m=n-1}^{n} \sum_{|\alpha|=m} ||c_\alpha(\rho)||^2 \tag{4.70}$$

Note that in the case of checking 2-separability of a 4-partite systems we have to compare the values of 4.70 for all possible number partitions of $n = 4$, *namely* $k_1 = 2, k_2 = 2$ *and* $k_1 = 3, k_2 = 1$. *In order to certify 2-separability in an experiment, we have to measure correlations that exceed the maximum of both functions.*

- *Case 1* $k_1 = 2, k_2 = 2$ *Since* $k_1, k_2 < x$ *all the correlation within the respective partitions are truncated. The cutoff function lets only the inter-partition correlations contribute.*

  $$\mathcal{C}_{(n,n-1)}(\rho_{2-sep}) = ||c_{|\alpha|=n}(\rho)||^2 = (d^{k_1}-1)(d^{k_2}-1) = d^n - d^{k_1} - d^{k_2} + 1 = 16 - 4 - 4 + 1 = 9.$$

- *Case 2* $k_1 = 3, k_2 = 1$ *When only applying the naive purity bound 4.33 we are left with*

  $$\mathcal{C}_{(n,n-1)}(\rho_{2-sep}) = ||c_{|\alpha|=n}(\rho)||^2 + ||c_{|\alpha|=n-1}(\rho)||^2 \tag{4.71}$$
  $$\leq (d^{k_1}-1)(d^{k_2}-1) + (d^{k_1}-1) = d^n - d^{k_2} = 16 - 2 = 14. \tag{4.72}$$

  Note that $\|\mathcal{C}_4\|_2 + \|\mathcal{C}_3\|_2 = 13$ *and thus measuring fullbody and 3-body correlations will not result in detection of 4-qubit square lattice cluster state, if only the purity bound is used to bound the correlation tensor norms. In contrast let us instead bound the 3-body correlation term* $||c_{|\alpha|=n-1}(\rho)||^2$ *by our sharper bound 4.58. With the help of a single application of our newly derived bound we obtain*

  $$\mathcal{C}_{(n,n-1)}(\rho_{2-sep}) = ||c_{|\alpha|=n}(\rho)||^2 + ||c_{|\alpha|=n-1}(\rho)||^2 \tag{4.73}$$
  $$\leq (d^{k_1}-1)(d^{k_2}-1) + ((d^{k_1}-1) - \frac{n}{n-1}(d^{k_1-2}-1)) = 7 + (7 - \frac{4}{3}(1)) = 12,666. \tag{4.74}$$

This value is below the threshold of 13 *necessary for detection by 4-body and 3-body correlations, although with not a particularly large noise resistance. Still, we summarize that the sharper bounds we developed, make the difference between the impossibility of entanglement detection and possibility. Also slightly better noise resistance would be possible through applying equation (4.58) also to the 4-body correlation terms, for simplicity we bounded them by the standard purity bound.*

Bounds for entanglement dimensionality detection

In the multipartite setting, the definition entanglement has to be subdivided into combinatorially many more sub-concepts including, but not limited to, notions such as k-separability. In fact every number partitioning of the number of parties n gives rise to its corresponding subclass of entanglement. The main part of this chapter has been devoted to the construction of general entanglement witnesses, where we only care whether any kind of entanglement is present within a system. In this paragraph we take a different viewpoint on (4.63) by considering a more specific problem than entanglement detection: so-called Schmidt rank vector witnesses.

If we closely consider inequality 4.63 it obviously depends on the sizes $k_1, \ldots, k_i$ of all involved partitions, that are present in any chosen partitioning of a system. This makes it clear that exceeding the inequality for any choice of partition size $k_1, \ldots, k_i$ also rules out that one can partition the state into that specific choice of partition. Note that frequently the Schmidt rank is also referred to as the dimensionality of entanglement. In that sense, the following bound can be seen as an entanglement dimensionality witness.

Theorem 4.3.3. For n-partite multipartite entangled states that can be decomposed into pure states with 1-party rank [6] of $(r_1, r_2, \cdots, r_n)$

$$\mathcal{C}_{(n,2)}(\rho_{(r_1,r_2,\cdots,r_n)}) \leq d^n + n - 1 - \sum_i (\frac{d}{r_i}) \tag{4.75}$$

Proof. We write out $\mathcal{C}_{(n,2)}$

$$\mathcal{C}_{(n,2)}(\rho_{(r_1,r_2,\cdots,r_n)}) = \underbrace{\sum_{m=2}^{n} \sum_{|\alpha|=m} |c_\alpha|^2 + \sum_i |c_i|^2}_{d^n-1} - \sum_i \underbrace{|c_i|^2}_{\geq \frac{d}{r_i}-1} = d^n + n - 1 - \sum_i (\frac{d}{r_i}), \tag{4.76}$$

where we made use of the purity and the fact that any operator of dimension r with positive eigenvalues and trace equal to one, can only have a squared trace as low as $\frac{1}{r}$. □

4.4 Multipartite entanglement detection: The anti-commuting basis approach

Finally, we come to our third strategy for multipartite entanglement detection: *entanglement detection through anti-commutativity.*

This chapter contains work published in [2]. We motivate & present the main theorem 4.4.1, discuss how it can be applied for entanglement detection and why it performs favorable in relation to other entanglement criteria, but also what obstacles hinder a universal application of this strategy.

4.4.1 Why anti-commutativity?

We once again come back to the question, that we have asked ourselves already in 4.1.1, what is an *entanglement detection criterion?*

This time, we give a slightly different answer. Since separability is expressed through tensor products in essence, an entanglement criteria needs to be a property that is in someway non-invariant under a tensor product. For example we have utilized

$$\|\rho_A \otimes \rho_B\| = \|\rho_A\| * \|\rho_B\| \tag{4.77}$$

in the Section 4.2 to detect entanglement.

Another famous example is the positive partial transpose criteria [25]. It exploits the simple fact that a state is Hermitian, while this is not necessarily true for any submatrix of a state. We have

$$\|(\mathbb{1} \otimes T)(\rho_A \otimes \rho_B)\|_1 = \left\|\rho_A \otimes \rho_B^T\right\|_1 = \|\rho_A \otimes \rho_B^*\|_1 = \|\rho_A \otimes \rho_B\|_1 . \tag{4.78}$$

We note that the a transposed state will again be a physical state and therefore be positive again. In contrast for an entangled ρ_{AB}

$$\|(\mathbb{1} \otimes T)(\rho_{AB})\|_1 =?, \tag{4.79}$$

here transposing only a part of the state may very well map a positive state onto a matrix with negative eigenvalues. Should a state be factorisable, the factors are again states. On these the transposition is a valid operation that can never leave the physical state space, thus any state that is mapped to a negative matrix by partial transposition will not factorize.

Finally anti-commutativity is exactly such a property. Why is this the case? Consider the following sketch of the basic idea. Let $\{A_i\}$ denote a family of anti-commuting operators. Now

$$A_i \tag{4.80}$$

is by definition a family of anti-commuting operators. However as soon as we add another anti-commuting operator into the picture the situation changes. The multi-linearity of the tensor product allow us to pull out all negative factors from the product. Therefore all tensor products containing an even number off anti-commuting factors, like

$$A_i' \otimes A_i \tag{4.81}$$

are not anymore anti-commuting operators. Now this is exactly the kind of behavior we are interested in for entanglement detection. Consider a system $\rho_{A'A}$ expressed in the basis 4.81. 4.81 is commuting by the above argument. Should $\rho_{A'A}$ factorize along the cut between A' and A, then in contrast the basis of for example the reduced density matrix ρ_A would be an anti-commuting basis of the form 4.80. As a closing remark we add that tensor multiplication with commuting operators does not influence the overall anti-commutativity. This allows us to to extend an anticommuting familiy to larger sets by padding with commuting operators.

The following theorem represents an effort to harness anti-commutativity into a concrete bound, that can be used as a detection criterion.

4.4.2 Anti-commutativity bound

Bound

The following theorem is a generalization of a series of known results for the qubit case qubits [46, 35, 51]. It has been known that sums of expectation values of squared dichotomic observables

are always bounded by one. The proof basically uses the positivity as an operator and the fact that the Pauli basis is anti-commuting. We extend this result to the d-dimensional case with the following theorem.

Theorem 4.4.1. *(Anti-commuting basis bound) Let $\{\lambda_i\}_{i\in\mathcal{I}}$ with the index set $\mathcal{I} = \{1, 2, \ldots, d^2\}$ denote an orthonormal self-adjoint basis $\mathcal{B}$ of a d-dimensional Hilbert space $\mathcal{H}$ and $\mathcal{A} \subseteq \mathcal{I} = \mathcal{B}$ be a partition of $\mathcal{B}$ such that the anti commutator $\{\lambda_i, \lambda_j\} = 0 \forall i, j \in \mathcal{A}$. The complement is denoted $\overline{\mathcal{A}}$.*

The corresponding coefficients c_i of any density matrix ρ expressed in $\mathcal{B}$ as $\rho = \sum_{i\in\mathcal{A}} c_i\lambda_i + \sum_{k\in\overline{\mathcal{A}}} c_k\lambda_k$ can be bounded by

$$\sum_{i\in\mathcal{A}} c_i^2 \leq \frac{\max_{i\in\mathcal{A}}\langle\lambda_i^2\rangle}{(\min_{i\in\mathcal{A}} \mathrm{Tr}(\lambda_i^2))^2} \tag{4.82}$$

Proof. Any ρ can be expressed in an orthonormal basis $\mathcal{B}$. The chosen basis $\mathcal{B}$ can always be divided into $\mathcal{A} \cup \overline{\mathcal{A}}$, where the set of anti-commuting basis elements is at worst trivial.

Now consider the observable P

$$P := \sum_{i\in\mathcal{A}} c_i\lambda_i. \tag{4.83}$$

Thus we trivially have

$$(\Delta P)^2 := \langle P^2\rangle - \langle P\rangle^2 \geq 0 \tag{4.84}$$

since the variance of any observable is positive.

Inserting for P we obtain

$$\langle P^2\rangle = \langle \sum_{i\in\mathcal{A}} c_i^2\lambda_i^2 + \underbrace{\sum_{i\neq j\in\mathcal{A}} c_ic_j\lambda_i\lambda_j\rangle}_{=0} = \sum_{i\in\mathcal{A}} c_i^2\langle\lambda_i^2\rangle, \tag{4.85}$$

with application of the anti-commutativity of $\mathcal{A}$, as well as

$$\langle P\rangle^2 = (\sum_{i\in\mathcal{A}} c_i\langle\lambda_i\rangle)^2 = (\sum_{i\in\mathcal{A}} c_i \,\mathrm{Tr}(\rho\lambda_i))^2 \tag{4.86}$$

$$= (\sum_{j\in\mathcal{A}} \mathrm{Tr}(\sum_{i\in\mathcal{A}} c_i\lambda_i)c_j\lambda_j) + \sum_{k\in\overline{\mathcal{A}}} \mathrm{Tr}(\sum_{i\in\mathcal{A}} c_i\lambda_i)c_k\lambda_k))^2 = (\sum_{i\in\mathcal{A}} c_i^2 \,\mathrm{Tr}(\lambda_i^2))^2 \tag{4.87}$$

by use of orthonormality of $\mathcal{B}$. Summing up the above three equations yields

$$0 \leq \langle P^2\rangle - \langle P\rangle^2 = \sum_{i\in\mathcal{A}} c_i^2\langle\lambda_i^2\rangle - (\sum_{i\in\mathcal{A}} c_i^2 \,\mathrm{Tr}(\lambda_i^2))^2 \tag{4.88}$$

$$\leq \sum_{i\in\mathcal{A}} c_i^2\Big(\max_{i\in\mathcal{A}}\langle\lambda_i^2\rangle - (\sum_{i\in\mathcal{A}} c_i^2)\min_{i\in\mathcal{A}} \mathrm{Tr}(\lambda_i^2)^2\Big). \tag{4.89}$$

The positivity enforces

$$0 \leq \max_{i\in\mathcal{A}}\langle\lambda_i^2\rangle - \sum_{i\in\mathcal{A}} c_i^2 \min_{i\in\mathcal{A}} \mathrm{Tr}(\lambda_i^2)^2 \Leftrightarrow \sum_{i\in\mathcal{A}} c_i^2 \leq \frac{\max_{i\in\mathcal{A}}\langle\lambda_i^2\rangle}{\min_{i\in\mathcal{A}} \mathrm{Tr}(\lambda_i^2)^2} \tag{4.90}$$

and thus the claim. □

Entanglement witness

As an corollary, we give a bound that acts as an experimental entanglement criterion.

Corollary 4.4.2. *Theorem (Separability Bound) Consider a product state $\rho = \rho_A \otimes \rho_{A'}$, whose marginals ρ_A and $\rho_{A'}$ are describable by two equally sized anti-commuting bases denoted as $\rho_A = \sum_{i\in\mathcal{A}} c_i\lambda_i$,$\rho_{A'} = \sum_{k\in\mathcal{A}'} c_k\lambda'_k$ such that for all $i,j \in \mathcal{A}$ $\{\lambda_i,\lambda_j\} = 0$ and for all $i,j \in \mathcal{A}'$ $\{\lambda'_i,\lambda'_j\} = 0$. Then we can bound*

$$\sum_{i\in\mathcal{A}} \mathrm{Tr}(\rho\lambda_i \otimes \lambda'_i) \leq \left(\sqrt{\max_{i\in\mathcal{A}}\langle\lambda_i^2\rangle}\frac{\max_{i\in\mathcal{A}}\mathrm{Tr}(\lambda_i^2)}{\min_{i\in\mathcal{A}}\mathrm{Tr}(\lambda_i^2)}\right)\left(\sqrt{\max_{i\in\mathcal{A}'}\langle\lambda_i'^2\rangle}\frac{\max_{i\in\mathcal{A}'}\mathrm{Tr}(\lambda_i'^2)}{\min_{i\in\mathcal{A}'}\mathrm{Tr}(\lambda_i'^2)}\right) \tag{4.91}$$

If this bound is violated, then ρ may not be separable.

Proof. By use of separability and $|\mathcal{A}| = |\mathcal{A}'|$ we can rewrite

$$\sum_{i\in\mathcal{A}} \mathrm{Tr}(\rho\lambda_i \otimes \lambda'_i) = \sum_{i\in\mathcal{A}} \mathrm{Tr}(\rho_A\lambda_i)\,\mathrm{Tr}(\rho_{A'}\lambda'_i) = \langle\vec{v}_A|\vec{v}_{A'}\rangle \leq ||\vec{v_A}|| * ||\vec{v_{A'}}|| \tag{4.92}$$

the expectation value into the vectors $[\vec{v}_A]_k =: \mathrm{Tr}(\rho_A\lambda_k)$,$[\vec{v}_{A'}]_k =: \mathrm{Tr}(\rho_{A'}\lambda'_k)$ and apply Cauchy-Schwarz. To compute the norms of these vectors we make use of our result 4.82 to derive

$$||\vec{v_A}|| = \sqrt{\sum_{i\in\mathcal{A}}\big(\mathrm{Tr}(\rho_A\lambda_i)\big)^2} = \sqrt{\sum_{i\in\mathcal{A}} c_i^2\big(\mathrm{Tr}(\lambda_i^2)\big)^2} \leq \sqrt{\max_{i\in\mathcal{A}}\langle\lambda_i^2\rangle}\frac{\max_{i\in\mathcal{A}}\mathrm{Tr}(\lambda_i^2)}{\min_{i\in\mathcal{A}}\mathrm{Tr}(\lambda_i^2)}. \tag{4.93}$$

The second norm is computed totally analogous proofing our statement. □

A systematic way to detect entanglement using anti-commutativity can accomplished by first identifying a set of non-zero bloch vector entries of a multipartite quantum state with anti-commuting reductions across the partition one is interested in. That is, we are looking for a partition $A|\bar{A}$ and a set $\beta = \beta_A \cup \beta_{\bar{A}}$ with elements $\tau_\beta = \mathrm{Tr}(\rho\lambda_i \otimes \lambda_j)$, where $i \in \beta_A$, $j \in \beta_{\bar{A}}$ and $\lambda_i(\lambda_j)$ are arbitrary observables acting on subsystem $A(\bar{A})$. The sum of moduli of these correlations can be bounded for states, which are product with respect to these partitions as

$$\mathrm{Tr}(\rho_A \otimes \rho_{\bar{A}}\lambda_i \otimes \lambda_j) = \mathrm{Tr}(\rho_A\lambda_i)\mathrm{Tr}(\rho_{\bar{A}}\lambda_j)\,, \tag{4.94}$$

and $|\langle u|v\rangle| \leq ||u||_2||v||_2$. Now let us assume that all observables λ_i are anti-commuting and normalized (i.e. $\mathrm{Tr}(\lambda_i\lambda_{i'}) = d\delta_{i,i'}$). Then we can make direct use of the anti-commutativity bound to assert that

$$\sum_{i\in\beta_A} |\mathrm{Tr}(\rho_A\lambda_i)|^2 \leq \max_{i\in\beta_A}\langle\lambda_i^2\rangle\,, \tag{4.95}$$

and analogously for $\bar{A}$. To finish we only need to point out that the original expression, a sum of moduli of expectation values, is convex in the space of density matrices and thus the validity of the inequality for product states translates to a general validity for separable states. The case of non-normalized or only partially anti-commuting observables works analogously.

For the case of HW-observables the anti-commutativity bound can be simplified to

$$\sum_{i\in\mathcal{A}}\langle\mathcal{Q}(\alpha_i)\rangle^2 \leq q_{\max}^2 \tag{4.96}$$

where $q_{\max}^2 = 1 + \max_{n\in\mathbb{N}} \sin(4\pi n/d)$ is the maximum eigenvalue of a HW observable which is the same for any $\mathcal{Q}(\alpha_j)$ for a given dimension.

In the supplemental material we present the proof that these are already the maximal set of anti-commuting operators in the HW-basis, i.e. that no more than three anti-commuting HW–observables exist. The proof is due to Ottfried Gühne and we received it in private communication. It is published only in the appendix of [2], where we acknowledge his contribution.

Example 4.4.3. *To illustrate this method in an exemplary case, let us turn to qudit systems with the maximally entangled state defined as*

$$|\phi_d\rangle = \frac{1}{\sqrt{d}}\sum_{j=0}^{d-1}|j\rangle\,|j\rangle\,. \tag{4.97}$$

This important class of entangled states in quantum information is a maximal resource for many tasks. Its Bloch decomposition in terms of HWOs is simply given by

$$|\phi_d\rangle = \frac{1}{d^2}\Big(\mathbb{1}\otimes\mathbb{1} + \sum_{\alpha\in\mathcal{S}}\mathcal{Q}(\alpha)\otimes\mathcal{Q}(\alpha)^*\Big) \tag{4.98}$$

where $\mathcal{Q}(\alpha)^* = \mathcal{Q}(-\alpha^*)$ denotes the complex conjugate. From above Bloch decomposition it follows that the expectation value of the correlations are all equal to 1, and this means that measuring only two anti-commuting local observables for each party is sufficient to violate the upper bound and thus detect entanglement. The violation is obviously enhanced with three pairwise anti-commuting observables whose respective amplitudes fulfill the constraint

$$|\alpha_1\times\alpha_2| = |\alpha_2\times\alpha_3| = |\alpha_3\times\alpha_1| = \pi/2(2n+1) \tag{4.99}$$

yielding a general recipe for finding three pairwise anti-commuting observables. In this case, the criterion (4.95) may written as

$$\sum_{i=1}^{3}\langle\mathcal{Q}(\alpha_i)\otimes\mathcal{Q}(\alpha_i)^*\rangle \overset{DV}{\leq} q_{\max}^2 \overset{CV}{\leq} 2, \tag{4.100}$$

with respective upper bounds on separable states for discrete (DV) and continuous variable (CV) cases. In the continuous limit ($d\to\infty$) the above entangled state becomes a perfectly correlated Einstein-Podolski-Rosen(EPR) entangled state [13] which is equal to an infinitely squeezed two-mode squeezed state, i.e. $\frac{1}{\sqrt{\pi}}\int_{\mathbb{R}} dx\,|x\rangle\,|x\rangle = \frac{1}{\sqrt{\pi}}\int_{\mathbb{R}} dx\,|p\rangle\,|-p\rangle$ with continuous Bloch decomposition $\frac{1}{\pi^2}\int d\alpha^2\mathcal{Q}(\alpha)\otimes\mathcal{Q}(\alpha)^*$. The associated set of three pairwise anti-commuting observables in this limit is simply given by a symmetric case of three equiangular amplitudes with equal lengths $|\alpha_j| = \sqrt{\pi/\sqrt{3}} \simeq 1.34$ mutually separated by an angle $2\pi/3$. In comparison, the corresponding correlations

with respect to the generalized Gell-Mann basis are all equal to $2/d$ [3]. Thus, the number of required measurements in order to detect entanglement scales with d making it impractical in high dimensions. This is an example which clearly demonstrates the advantage of the HW observables in high-dimensional entanglement detection.

Example 4.4.4. *Entanglement detection for Gaussian States*
Consider for example the following set of observables

$$\mathcal{Q}(\alpha_1) := \mathcal{Q}(0, \sqrt{\frac{d}{2}}), \tag{4.101}$$

$$\mathcal{Q}(\alpha_2) := \mathcal{Q}(\sqrt{\frac{d}{2}}, 0), \tag{4.102}$$

$$\mathcal{Q}(\alpha_3) := \mathcal{Q}(\sqrt{\frac{d}{2}}, \sqrt{\frac{d}{2}})) \tag{4.103}$$

It is necessary that the complex index α of $\mathcal{Q}$ scales with d to make sure that anti-commutativity is always ensured as well as that all index values stay finite. First check that the choice of indices

$$\alpha_1 = \sqrt{\frac{\pi}{d}}(0 + i\sqrt{\frac{d}{2}}) = i\frac{\pi}{2}, \tag{4.104}$$

$$\alpha_2 = \sqrt{\frac{\pi}{d}}(\sqrt{\frac{d}{2}} + i * 0) = \frac{\pi}{2}, \tag{4.105}$$

$$\alpha_3 = \sqrt{\frac{\pi}{d}}(\sqrt{\frac{d}{2}} + i\sqrt{\frac{d}{2}}) = \sqrt{\frac{\pi}{2}}(1 + i) \tag{4.106}$$

remain finite regardless of the the dimension d and in particular for $d \to \infty$. We compute the pair wise cross products

$$|(l_1, m_1) \times (l_2, m_2)| = \frac{d}{2} \tag{4.107}$$

$$|(l_1, m_1) \times (l_3, m_3)| = \frac{d}{2} \tag{4.108}$$

$$|(l_2, m_2) \times (l_3, m_3)| = \frac{d}{2} \tag{4.109}$$

and find that all of them full fill the anti-commutativity condition $|l'm - lm'| = \frac{d}{2}(2n+1)$ for $n = 0$ and all d. Therefore equation 4.101 is a well defined set of anti-commuting observables.

On the applicability of the anti-commutativity bound

We have to point out what a seemingly impossible task the anti-commutativity bound seems to achieve. Large parts of this thesis (refer to 4.3) have been devoted to entanglement detection through the purity bound and indeed one could argue that this is almost the archetypical approach towards entanglement detection. Now recall the purity bound for a n-qudit system

$$1 \geq \mathrm{Tr}(\rho^2) = \frac{1}{d^n}(1 + r_1^2 + \ldots + r_{d^n-1}^2) = \frac{1}{d^n}(\mathbb{1} + \|\vec{r}\|_2^2).$$

All detection strategies that aim to bound via the purity suffer from the normalization factor $\frac{1}{d^n}$. For high dimension or many particles this factor basically blows up very quickly and is not longer a strong constraint on the system. Somehow this could be seen as the fundamental problem of multipartite entanglement detection. In contrast the anti-commutativity bound is completely independent of both the particle number and n and the local system dimension d. This is an enormous simplification for large systems. At a first glance it may promise to give a general entanglement detection approach, however one has to point out that the bound comes with a caveat.

It holds by definition *only for anti-commuting sets of operators*. Now, if those sets are large enough to contain a basis for a quantum state it is indeed possible to easily detect this state independent of size and dimension. However typically this is not the case beyond the qubit case, in fact it seems a surprisingly hard task to find arbitrary large sets of pair wise non-trivial anti-commuting operators.

With non-trivial anti-commuting set, we refer to sets that are not commuting and anti-commuting at the same time. An example for trivially anti-commuting sets, are basically all basis elements with different support. In the case of the generalized Gell-Mann basis a huge part of the basis is trivially anti-commuting. Excluding trivially anti-commuting sets, has been one of our main motivation to investigate full rank bases like the Heisenberg-Weyl or the symmetrized Heisenberg-Weyl basis.

In fact, it turns out to be even quite hard to find sets of quite modest size. Typically, one can easily find up to 3 pairwise anti-commuting operators within the basis of state, however extending beyond this has eluded us. It seems we have shifted from one hard problem to another.

Clearly, pairwise anti-commutativity is a set of combinatorially growing many constraint, thus the complexity of the condition should grow somewhat fast. For a single system, it is quite puzzling that already such small numbers as a basis of 4 pairwise anti-commuting bases seem impossible to achieve. However, even though this may sound like a problem, the anti-commutativity bound remains very useful for composite systems. The situation here mirrors somehow the complications arising in the generalized Bloch decomposition. Even though extending to high dimension seem hard or even impossible, extending to large particle number is a relatively straightforward procedure. Note that a maybe quite small anti-commuting basis of a single system may be extended via tensor products to quite large bases of a multi-partite system.

We know at least of one result indicating the hardness to find anti-commuting bases and thank Otfried Gühne for the comment, that in the case of Heisenberg-Weyl Observables it is impossible to extend beyond three anti-commuting basis elements. A similar result basis independent result has not been found by us up to date.

4.4.3 The anti-commutativity bound revisited: Extending to almost anti-commuting sets

In the last Section 4.4.2 we have discussed the problem of finding large anti-commuting sets of bases. We have come to the somewhat unsatisfying conclusion, that this at least seems to be a hard task. As a possible solution we proposed the notion of *almost anti-commutativity*.

Definition 4.4.5. Let $\{\lambda_i\}_i$ be a basis of an Hilbert space. We define the quantity

$$\mathcal{K} := \frac{1}{2}\sqrt{\sum_{i\neq j\in\mathcal{A}} \langle\{\lambda_i,\lambda_j\}\rangle^2} \tag{4.110}$$

and call $\mathcal{K}$ the **averaged anti-commutator** of the basis $\{\lambda_i\}_i$.
If $\mathcal{K}$ is small, we call $\{\lambda_i\}_i$ an **almost anti-commuting basis**.

Note, that similar problems have been studied when discussing uncertainty relations [30].

The above notion is a quite intuitive one. It is a standard mathematically technique to relax to restrictive definition to a ϵ-small definition and thus extend the size of too constrained sets. We hope that the almost anti-commuting bases compensate a slightly larger averaged anti-commutator by correspondingly larger sets of almost anti-commuting bases. We propose the second version of 4.4.1, where we simply replace the notion of anti-commutativity to almost anti-commutativity. The proof is of course almost identical.

Corollary 4.4.6. *Let $\{\lambda_i\}_{i\in\mathcal{I}}$ with the index set $\mathcal{I} = \{1, 2, \ldots, d^2\}$ denote an orthonormal self-adjoint basis $\mathcal{B}$ of a d-dimensional Hilbert space $\mathcal{H}$ and $\mathcal{A} \subseteq \mathcal{I}$ refer to a subset of $\mathcal{B}$ such that $\frac{1}{2}\sqrt{\sum_{i\neq j\in\mathcal{A}}\langle\{\lambda_i,\lambda_j\}\rangle^2} \leq \mathcal{K}$. Then the corresponding Bloch vector components c_i of any density matrix ρ expressed in $\mathcal{B}$ as $\rho = \sum_{i\in\mathcal{A}} c_i\lambda_i + \sum_{l\in\overline{\mathcal{A}}} c_l\lambda_l$ can be bounded by*

$$\sum_{i\in\mathcal{A}} c_i^2 \leq \frac{\max_{i\in\mathcal{A}}\langle\lambda_i^2\rangle + \mathcal{K}}{(\min_{i\in\mathcal{A}} \mathrm{Tr}(\lambda_i^2))^2} \tag{4.111}$$

Proof. Any ρ can be expressed in an orthonormal basis $\mathcal{B}$. The chosen basis $\mathcal{B}$ can always be divided into $\mathcal{A}\cup\overline{\mathcal{A}}$, where the set $\mathcal{A}$ is at worst trivial.

Now consider the observable $\mathcal{O}$

$$\mathcal{O} := \sum_{i\in\mathcal{A}} c_i\lambda_i. \tag{4.112}$$

We trivially have

$$(\Delta\mathcal{O})^2 := \langle\mathcal{O}^2\rangle - \langle\mathcal{O}\rangle^2 \geq 0 \tag{4.113}$$

since the variance of any observable is positive.

Inserting for $\mathcal{O}$ we obtain

$$\begin{aligned}\langle\mathcal{O}^2\rangle = \langle\sum_{i\in\mathcal{A}} c_i^2\lambda_i^2 + \sum_{i\neq j\in\mathcal{A}} c_ic_j\lambda_i\lambda_j\rangle = \langle\sum_{i\in\mathcal{A}} c_i^2\lambda_i^2 + \frac{1}{2}\sum_{i\neq j\in\mathcal{A}} c_ic_j \underbrace{(\lambda_i\lambda_j + \lambda_j\lambda_i)}_{\{\lambda_i,\lambda_j\}}\rangle \leq \\ \sum_{i\in\mathcal{A}} c_i^2\langle\lambda_i^2\rangle + \underbrace{\sqrt{\sum_{i\neq j\in\mathcal{A}} c_i^2c_j^2}}_{\leq\sum_i c_i^2}\underbrace{\frac{1}{2}\sqrt{\sum_{i\neq j\in\mathcal{A}}\langle\{\lambda_i,\lambda_j\}\rangle^2}}_{\leq\mathcal{K}},\end{aligned} \tag{4.114}$$

with application of the boundedness of the anti commutator of $\mathcal{A}$, as well as

$$\langle\mathcal{O}\rangle^2 = (\sum_{i\in\mathcal{A}} c_i\,\mathrm{Tr}(\rho\lambda_i))^2 = (\sum_{j\in\mathcal{A}}\mathrm{Tr}(\sum_{i\in\mathcal{A}} c_i\lambda_i)c_j\lambda_j) + \sum_{l\in\overline{\mathcal{A}}}\mathrm{Tr}(\sum_{i\in\mathcal{A}} c_i\lambda_i)c_l\lambda_l))^2 = (\sum_{i\in\mathcal{A}} c_i^2\,\mathrm{Tr}(\lambda_i^2))^2 \tag{4.115}$$

by use of orthonormality of $\mathcal{B}$. Summing up the above three equations yields

$$0 \leq \langle \mathcal{O}^2 \rangle - \langle \mathcal{O} \rangle^2 \leq \sum_{i \in \mathcal{A}} (c_i^2 \langle \lambda_i^2 \rangle + c_i^2 \mathcal{K}) - (\sum_{i \in \mathcal{A}} c_i^2 \operatorname{Tr}(\lambda_i^2))^2 \leq \sum_{i \in \mathcal{A}} c_i^2 \Big(\max_{i \in \mathcal{A}} \langle \lambda_i^2 \rangle - (\sum_{i \in \mathcal{A}} c_i^2) \min_{i \in \mathcal{A}} \operatorname{Tr}(\lambda_i^2)^2 \Big) + \sum_{i \in \mathcal{A}} c_i^2 \mathcal{K}. \tag{4.116}$$

thus

$$\sum_{i \in \mathcal{A}} c_i^2 \Big(\max_{i \in \mathcal{A}} \langle \lambda_i^2 \rangle - (\sum_{i \in \mathcal{A}} c_i^2) \min_{i \in \mathcal{A}} \operatorname{Tr}(\lambda_i^2)^2 \Big) + \sum_{i \in \mathcal{A}} c_i^2 \mathcal{K} \leq \sum_{i \in \mathcal{A}} c_i^2 \Big(\max_{i \in \mathcal{A}} \langle \lambda_i^2 \rangle - (\sum_{i \in \mathcal{A}} c_i^2) \min_{i \in \mathcal{A}} \operatorname{Tr}(\lambda_i^2)^2 \Big) \tag{4.117}$$

The positivity enforces

$$0 \leq \max_{i \in \mathcal{A}} \langle \lambda_i^2 \rangle - (\sum_{i \in \mathcal{A}} c_i^2) \min_{i \in \mathcal{A}} \operatorname{Tr}(\lambda_i^2)^2 + \mathcal{K} \Longleftrightarrow \sum_{i \in \mathcal{A}} c_i^2 \leq \frac{\max_{i \in \mathcal{A}} \langle \lambda_i^2 \rangle + \mathcal{K}}{\min_{i \in \mathcal{A}} \operatorname{Tr}(\lambda_i^2)^2} \tag{4.118}$$

□

As it may still be cumbersome to compute $\mathcal{K}$ we point out that $\mathcal{K} = \frac{1}{2} \sqrt{\sum_{i \neq j \in \mathcal{A}} \langle \{\lambda_i, \lambda_j\} \rangle^2} \leq \frac{1}{2} \sqrt{\sum_{i \neq j \in \mathcal{A}} \|\{\lambda_i, \lambda_j\}\|_\infty^2}$. As the HW-Basis obeys $\operatorname{Tr}(\lambda_i^2)^2 = d^2 \, \forall \, \lambda_i$ and the coefficients c_i are related to the Bloch vector components by $c_i = \frac{1}{d} \langle \mathcal{Q}(\alpha) \rangle$ we can conclude with the simple bound

$$\sum_{i \in \mathcal{A}} \langle \mathcal{Q}(\alpha_i) \rangle^2 \leq 1 + \max_n \left(\sin \frac{4\pi n}{d} \right) + \frac{1}{2} \sqrt{\sum_{i \neq j \in \mathcal{A}} \|\{\mathcal{Q}(\alpha_i), \mathcal{Q}(\alpha_j)\}\|_\infty^2} \tag{4.119}$$

Chapter 5

Application: Thermodynamics

5.1 Energy cost of correlations & work extraction

5.1.1 The formulation of thermodynamic notions within quantum mechanical context

What is (microscopic) work?

In the context of quantum thermodynamics the energy difference between a certain initial state ρ_i and a final state ρ_f has been proposed. (c.f. [29], where the authors analyze the variation of correlations on the unitary orbit).

Definition 5.1.1. Let us call the following quantity the **untouched work**

$$W := \mathrm{Tr}(H_{tot}(\rho_f - \rho_i)) = \mathrm{Tr}(H_{tot}(U\rho_i U^\dagger - \rho_i)). \tag{5.1}$$

It is an interesting quantity, since it surely captures the cost of energy of the operation U. We hesitate though to call it work, because we are not going to introduce any microscopic version of the second law of thermodynamics in this section. Thus we have no notion of heat. More precisely, we lack any notion of separation of energy into heat and work. This is why we differentiate by calling the quantity above untouched work.

It should be added as a warning though, that the very definition of work in quantum thermodynamics is a disputed subject with many different proposed definitions (see [17] or [45] for two differing approaches). The above one seems sensible, since it simply calculates the difference in energy of a state before and after a time evolution. If the initial state has a higher energy than the final state, the difference in energy could be utilized in a different system and thus called "work". Apart from this rather intuitive justification, it turns out that this quantity coincides with the free energy difference of a quantum system. Please refer for more details to our introduction of quantum thermodynamics in Section 2.3 and the relation in equation (2.57) given therein. There we further justify, why we believe the above quantity to be analogous to the classical definition.

However this way to define "microscopic work" is by no means the only one. The above framework can be employed to study different classes of states with respect to the work cost or gain by transforming between them.

On the relevance of thermal states

Thermal states always play an important role in thermodynamics, since thermal states at ambient temperature are considered as free states in a resource theoretic approach to thermodynamics. In [28] we propose the initial state $\rho_i = \bigotimes_{i=1}^{n} \tau_\beta$. Considering this state, we have studied what kind of states are accessible from a set of uncorrelated thermal states at a given temperature β, while we are limited to a certain amount of work W. This setting seems to be a good starting point, since by the validity of statistical physics a state without any preprocessing would most likely be a thermal state at the temperature of our lab. Moreover, in comparison to a quantum mechanical system at study the environment of the lab is generally considered to be so large that statistical arguments can be applied with idealized certainty. Thus it is described as an infinite ensemble of thermal states. We are not limited to singling out one or just a few thermal state, but have an arbitrary supply of thermal states. For this reason, we consider a tensor product of thermal states as the prototypical input, for a quantum mechanical experiment. Even though most well-known quantum information algorithms are usually described as starting with entangled states or at least some kind of correlated states, these have to be generated with a considerable amount of beforehand effort. In essence, these have to be generated for example by state distillation from less entangled states such as thermal states during preprocessing. Typically state preparation is simply omitted in the description of quantum information algorithms. In reality we might want to start a given experiment with a larger set of uncorrelated thermal states that is then manipulated by some preparation resulting in only a hand full of entangled states. Those states will then be the basis for further protocols, but if the whole preparation stage of the algorithm is not to be disregarded, we should always start a protocol from thermal states.

Summing up the thermal state captures the notion of being the state we "normally" find around us. In addition all temperatures, that we can attain by macroscopic heating and cooling, can also be considered to be generally easily accessible as additional heat baths. Their corresponding thermal states are also readily available.

Work cost of creating correlation from thermal states I: Formulating an optimization problem

Now for any heat bath at temperature β

$$W(\beta, U) := \mathrm{Tr}(H_{tot}(U \bigotimes_{i=1}^{n} \tau_\beta U^\dagger - \bigotimes_{i=1}^{n} \tau_\beta) \tag{5.2}$$

describes the work cost of U. Work is such a fundamental concept in thermodynamics, that it has to show up in some way in all thermodynamic models. The only exception may be a setting that purely considers heat flows, but as soon as we are considering any kind of operation that has to be performed work has to be modeled. Now depending on the circumstances we can either choose the above 5.2 to be a constraint in an optimization problem for some kind of task, that we are about to study. The exact phrasing of the optimization problem determines the type of thermodynamic question we are posing.

We can study problems formulated as

$$\begin{aligned} &\max \quad f(\rho_f, \rho_i, \ldots) \\ &\text{s.t.} \quad W(\beta, U) = \mathrm{Tr}(H_{tot}(U \bigotimes_{i=1}^{n} \tau_\beta U^\dagger - \bigotimes_{i=1}^{n} \tau_\beta)) \leq c. \end{aligned} \tag{5.3}$$

where c is a constant limiting the available work, β is fixed and optimality is considered with respect to some other function f. f represents either some quantum mechanical task or any kind of information processing, while the equation (5.2) serves as a constraint for the optimization problem, that characterizes the work cost of the task. Interesting quantities for the objective function f could be all types of entanglement, correlation, information (all preceding examples are covered in [28]), coherence (as covered in [37]), the overlap with a target state or even discord. The following Subsection 5.1.2 studies for example the energy cost of creating correlations.

5.1.2 Creating correlations at optimal energy cost

Work cost of creating correlation II: General conditions for optimally correlation creating unitaries

Again, as usually in the quantum information framework, the bipartite case is easier. We find in [28] that for example the problem of quantifying how much correlation can be created between a number of uncorrelated thermal states when a certain amount of work W is invested can be solved for some classes of states analytically. For the bipartite case we find the optimal solution. The generalization of the above approach yields a viable but non-optimal general protocol. The mutual information $I(A_1 : A_2) = S(A_1) + S(A_2) - S(A_1A_2)$ is a well-known measure for correlations between two systems and thus a suitable objective function for analyzing correlations. Considering 5.2 as an energy constraint and choosing $I(A_1 : A_2)$ is thus our choice for the optimization problem we study in [28]. Apart from quantifying the correlations, $I(A_1 : A_2)$ consists of von-Neumann entropies that simplify sufficiently when applied to thermal states to allow for an analytical treatment of the qubit case.

Considering the bipartite case we have for $\rho_i = \tau_\beta \otimes \tau_\beta$,

$$I(A_1 : A_2) = S(\tau_\beta) + S(\tau_\beta) - \underbrace{S(\tau_\beta \otimes \tau_\beta)}_{S(\tau_\beta)+S(\tau_\beta)} = 0, \tag{5.4}$$

since the joint entropy is additive on tensor product states. This means that in our setting we always start with no information, where we hope to find the optimal unitary $\hat{U}$ that yields a maximal mutual information. It is clear from the start that globally we can not do much, because any unitary will leave the entropy $S(A_1A_2)$ invariant. The action of U on the whole system can therefore be disregarded. The situation is different however the subsystems, since a global unitary does not necessarily act unitarily on a smaller local system, it may very well influence the local entropies $S(\tau_\beta)$. In fact we know already what kind of state maximizes entropy for a given average energy: again it is a thermal state.

This argument determines the general form the marginals of ρ_f by requiring the action of the optimal unitary $\hat{U}$ to be

$$\mathrm{Tr}_{A_1}(\hat{U}\rho\hat{U}^\dagger) = \mathrm{Tr}_{A_2}(\hat{U}\rho\hat{U}^\dagger) = \tau_{\beta_f}. \tag{5.5}$$

The form of the marginals are almost completely fixed up to the free parameter in the definition of thermal states: the temperature β_f. The attainable temperature β_f in turn, needs to be matched to the energy at our disposal by

$$W \stackrel{!}{=} \mathrm{Tr}(H_{tot}(\hat{U}\bigotimes_{i=1}^{2}\tau_\beta\hat{U}^\dagger - \bigotimes_{i=1}^{2}\tau_\beta)). \tag{5.6}$$

It is noteworthy that the whole argument about invariance of the global entropy and the maximization procedure of the local entropies, that appear in the definition of the mutual information can be extended without any problems to a multipartite setting. The necessary condition from equation (5.5) remains completely analogous for any number of involved subsystems.

Even though we have now, a much clearer idea how the optimal unitary has to be acting, we are still lacking a constructive example how this unitary might look like. In the next section, we will find another equivalent condition for optimality: a decomposition into circulant matrices.

Work cost of creating correlation III: Constructing optimal correlation creating unitaries

In this section we concern ourselves with finding unitaries that accomplish the task of optimal correlation creation under energy constraints. We note however that this will not be a constructive result. Still, this is sufficient to analyze possible optimality conditions on unitaries for optimal work extraction.

We have already derived the necessary condition in equation (5.5), how can we ensure that this condition is fulfilled?

Theorem 5.1.2. *Let τ be a thermal state and τ' another thermal state. Optimal conversion of untouched work into mutual information is possible if the thermal distribution τ' is cyclically majorized by τ.*
More specifically, if U is a unitary, that transforms one marginal of $\tau\otimes\tau$ like

$$Tr_A(U\tau\otimes\tau U^\dagger) = \tau', \tag{5.7}$$

and additionally has an action expressible as a circulant matrix C on the diagonal part of the marginal by

$$\mathrm{diag}(Tr_A(U\tau\otimes\tau U^\dagger)) = C\,\mathrm{diag}(Tr_A(\tau\otimes\tau)), \tag{5.8}$$

then both marginals transform equally. Thus we have

$$Tr_A(U\tau\otimes\tau U^\dagger) = Tr_B(U\tau\otimes\tau U^\dagger) = \tau', \tag{5.9}$$

coinciding with the optimality condition for U given in equation (5.5). Thus U is converting all work optimally into mutual information.

Proof. We outline our proof structure:

1. To start we show that the action of different partial traces, either Tr_A or Tr_B, on diagonal operators D are expressible in terms of each other. Apart from allowing us to rewrite one marginal into the other, we show that the partial traces action on the diagonal part of the marginals can be decomposed into the action of d disjoint d-dimensional subspaces. This is the content of Lemma 5.1.3.

2. We argue that a due to the properties of tensor products of thermal states $\mathrm{Tr}_A(\rho) \stackrel{!}{=} \mathrm{Tr}_B(\rho)$ can be asserted. This shows that initially both marginals are the same.

3. We show that unitary transformations induce stochastic transformations on the diagonal part of an operator they transform in Lemma 5.1.4.

4. The first Lemma 5.1.3 allows us to rewrite the different marginals into each other. Now we show that if a unitary exists, that fulfills our assumption of cyclic action on the diagonal, it commutes with the cyclic permutations arising when rewriting Tr_A into Tr_B. Thus we can apply the action to the initial marginals, that we have argue to be equal in the second step. This forces both marginals to transform equally, thus satisfying the optimality condition 5.5.

Lemma 5.1.3. *Let D be a diagonal bipartite state. Assuming marginals of equal dimension n, D's partial trace's diagonal part* $\mathrm{diag}(\mathrm{Tr}_A(D))$*, can be represented as a sum of vectors by*

$$\mathrm{diag}(\mathrm{Tr}_A(D)) = \sum_{i=0}^{\dim(B)=n} \vec{p}_i. \tag{5.10}$$

The other partial trace $\mathrm{Tr}_B(D)$ *, can be rewritten terms of $\vec{p}_i$ belonging to the partial trace* $\mathrm{Tr}_A(D)$*, by applying suitable circulant matrices* Π_c*. Thus we have the description*

$$\mathrm{diag}(\mathrm{Tr}_B(D)) = \sum_{i=0}^{\dim(B)=n} \Pi_c^i \vec{p}_i. \tag{5.11}$$

It is possible to describe A in terms of B and vice versa.

Proof. We start by noting that the partial trace of a diagonal operator is again a diagonal operator. More specifically the action of the partial trace on a diagonal operator reduces to a sum of diagonal operators allowing us to write

$$\mathrm{Tr}_A(D) = \begin{pmatrix} \sum_i d_{i0} & 0 & \dots & 0 \\ 0 & \sum_i d_{i1} & 0 & 0 \\ \vdots & 0 & \ddots & 0 \\ 0 & 0 \dots & 0 & \sum_i d_{in} \end{pmatrix}, \mathrm{Tr}_B(D) = \begin{pmatrix} \sum_j d_{0j} & 0 & \dots & 0 \\ 0 & \sum_j d_{1j} & 0 & 0 \\ \vdots & 0 & \ddots & 0 \\ 0 & 0 \dots & 0 & \sum_j d_{nj} \end{pmatrix}. \tag{5.12}$$

Now let us only consider the diagonal part of $\mathrm{Tr}_A(D)$. It can be treated in terms of certain vectors $\vec{p}_i$. We can rewrite it as

$$\mathrm{diag}(\mathrm{Tr}_A(D)) = \begin{pmatrix} \sum_i d_{i,0} \\ \sum_i d_{i,1} \\ \vdots \\ \sum_i d_{i,n} \end{pmatrix} = \sum_i \begin{pmatrix} d_{i,0} \\ d_{i+1,1} \\ \vdots \\ d_{i+n,n} \end{pmatrix} := \sum_i \vec{p}_i, \tag{5.13}$$

where all indices larger than n should be taken in the modulo n (i.e.: identifying $d_{i+n,n} = d_{i,n}$). We make use of the fact that the addition is commutative in the second equality. This allows us

to arbitrary rearrange the summands into any permutation, but we choose to apply a right shift cyclic permutation on the summands. Now lets turn our attention to the remaining partial trace

$$\mathrm{diag}(\mathrm{Tr}_B(D)) = \begin{pmatrix} \sum_j d_{0,j} \\ \sum_j d_{1,j} \\ \vdots \\ \sum_j d_{n,j} \end{pmatrix} = \sum_j \begin{pmatrix} d_{0,j} \\ d_{1,j} \\ \vdots \\ d_{n,j} \end{pmatrix} = \sum_j \begin{pmatrix} d_{0,j} \\ d_{1,j+1} \\ \vdots \\ d_{n,j+n} \end{pmatrix} \tag{5.14}$$

$$= \sum_j \Pi_c^{-j} \begin{pmatrix} d_{-j,0} \\ d_{-j+1,1} \\ \vdots \\ d_{-j+n,n} \end{pmatrix} = \sum_{i=-j} \Pi_c^i \begin{pmatrix} d_{i,0} \\ d_{i+1,1} \\ \vdots \\ d_{i+n,n} \end{pmatrix} = \sum_{i=-j} \Pi_c^i \vec{p_i}, \tag{5.15}$$

showing that we can rewrite one partial trace diagonal part in terms of the cyclically permuted others vector description $\vec{p_i}$. □

Clearly, a tensor product of a thermal state $\tau \otimes \tau$ is representable diagonally in energy basis, thus the above lemma applies. Giving us the following equality of its marginals

$$\sum_i \vec{p_i} = \mathrm{Tr}_A(\tau \otimes \tau) = \mathrm{Tr}_B(\tau \otimes \tau) = \sum_i \Pi_c^i \vec{p_i}. \tag{5.16}$$

Furthermore, the above lemma can be seen as a decomposition of the marginals ρ_A, ρ_B into d-dimensional subspaces represented by the $\vec{p_i}$ and respectively $\Pi_c^i \vec{p_i}$. For each i these vectors are defined on a different d-dimensional support. Let U^d denote a unitary with d-dimensional support, (but possibly global!). For example a unitary U_d acting on $span(p_1)$, will be acting exclusively on p_1,as

$$\mathrm{Tr}_A(U_i^d \rho (U_i^d)^\dagger) = A p_1 + \sum_{i=2}^{d} \vec{p_i}. \tag{5.17}$$

Now if we assume that we can transform $\tau = \mathrm{Tr}\, A(\tau \otimes \tau)$ into the optimal thermal state $\tau' = Tr_A(U^{d_1} \dots U^{d_{d-1}} \tau \otimes \tau (U^{d_1} \dots U^{d_{d-1}})^\dagger)$, by application of d different U_i^d with circulant action C, then we show that the optimality condition 5.5 is met.
This would mean $U^{d_1} \dots U_{d-1}^d \overset{!}{=} \hat{U}$ and that finding such a set of circulant U_i^d already fixes an optimal unitary.

Observe that the action of U_i^d is necessary doubly stochastic. This is clear due to the following lemma:

Lemma 5.1.4. *Let U_i^d be a unitary with d dimensional support on $span(\vec{p_i})$, for diagonal operators D.*

$$U_i^d D_A (U_i^d)^\dagger = M_i \vec{p_i} + \sum_{i \neq j = 1}^{d} \vec{p_i}, \tag{5.18}$$

with M_i doubly stochastic.

Proof. Generally the action of a unitary on the diagonal is described by

$$\langle j|D|j\rangle = \sum_i U_{ji} \langle i|D|i\rangle (U_{ji})^\dagger = \sum_i |U_{ij}|^2 \underbrace{\langle i|D|i\rangle}_{=(\vec{p_j})_i}. \tag{5.19}$$

Take care not to confuse the sub-indices. $\langle i|D|i\rangle = (\vec{p_j})_i$ is to be red as the i-th component of the j-th vector $\vec{p_j}$.

The expression above is familiar from the theory of uni-stochasticity. The problem is well-known and not thought to be easy in general. For literature on the connection check of unitaries and doubly stochastic refer either to [36] for a textbook introduction, or to [44][42] for more recent papers with a more quantum theoretic flavor.

A doubly stochastic matrix M is called uni-stochastic, if it can decomposed into unitaries U as

$$M = |U|^2. \tag{5.20}$$

Even though it is still an open question to completely characterize uni-stochasticity of a matrix M, there is an easy and a hard direction in the above equality.

Starting from an arbitrary doubly stochastic matrix M is the hard direction. It is not always possible and generally hard to find a unitary decomposition for a given M. We believe, that contributions from uni-stochasticity may be applied to ensure the existence, that we have to require up until now.

The other way, starting from a unitary U is clear however. we know that every unitary U induces a doubly stochastic $M = |U|^2$ through the mapping of the Schur product $\circ$ by $U \circ U^* = |U|^2$. Since we postulated the existence of suitable unitaries ΠU^d, this case is sufficient, thus allowing us to write

$$\langle j|D|j\rangle = \sum_i |U_{ij}|^2 \langle i|D|i\rangle = \sum_i M_{ij}\langle i|D|i\rangle \tag{5.21}$$

and thus proving the lemma. □

Now compare the action of ΠU_j^d on the different marginals. Remember, that in our assumption we have required U_j^d to transform the first marginal to the optimal temperature thermal state τ' as well as being circulant. For the first marginal, we have by assumption on U_j^d that $\tau' = \rho'_A$. Now recall, that the state that maximizes entropy with minimal energy is always a thermal state and can be expressed again as a diagonal operator $D_{\tau'}$ in energy eigenbasis. This assures, that the prerequisites of 5.1.4 are met. We can thus write

$$\mathrm{diag}(\tau') = \mathrm{diag}(\rho'_A) = \sum_j M_j \vec{p_j} = \mathrm{Tr}_A(U_j^d \rho (U_j^d)^\dagger). \tag{5.22}$$

Now the action on the remaining marginal is the interesting part, because if we manage to show that it transforms to the same thermal state τ', we have shown that the optimality condition 5.5 is fulfilled. Check, that by using the assumption that M_j are in fact one cyclic matrix C and thus $Pi_c M_j = C Pi_c$

$$\mathrm{diag}(\rho'_B) = \sum_j \Pi_c^i M_j \vec{p_j} = C \sum_j \Pi_c^i \vec{p_j} = C\,\mathrm{diag}\,\rho_B = C\,\mathrm{diag}\,\rho_A = \mathrm{diag}(\tau'), \tag{5.23}$$

that both marginals transform the same and are thus optimal. □

The above theorem establishes the utility of the concept of circulant majorization. In contrast to the classical notion of majorization the topic of circulant majorization has received far less attention. Nontheless, it has already been studied in the mathematical community (c.f. [43]). We believe that the interplay between circulant matrices, doubly-stochastic matrices and majorization merits further investigation.

5.2 Temperature limits to correlation creation

This section is devoted to the study of the maximum amount of correlations that can be generated in tensor products of thermal states. Again our main tool to quantify correlations will be the mutual information in either the bipartite or multipartite formulation. We directly treat the more general multipartite case by considering

$$I(A_1 : \ldots : A_n) = \sum_i S(A_i) - S(A_1 \ldots A_n) \tag{5.24}$$

Finding out the maximum of achievable correlation possible from thermal states turns out to be a considerable easier task than to quantify the work cost of creating correlations. This is because we can easily construct upper bounds on the mutual information that can be saturated by the right choice of state.

$$I(A_1 : \ldots : A_n) \leq n(\log d - S(\tau(\beta))) \tag{5.25}$$

Before demonstrating the saturation of the above bound, we take some time to comment this result. Already this very simple bound has a quite satisfying behavior. Clearly, we have $0 \leq S(\tau(\beta)) \leq \log d$, where the maximum is of-course attained at infinite temperature. Observe that S grows monotonously with β. This basically tells us that increasing the temperature directly lowers the correlations present within the system, giving us exactly the kind of relation between temperature and correlations that we would expect to find.

When looking for the maximizing state, our choice is already almost fully determined, since we want to have a state with marginals of maximal entropy. Considering

$$S(Tr_X(|\psi_+\rangle\langle\psi_+|)) = S(\frac{1}{2}\operatorname{Tr}_X(\sum_{i,j=0}^{1}|ii\rangle\langle jj|)) = S(\frac{1}{2}\mathbb{1}) = \log(2) \tag{5.26}$$

we see that for example already a qubit Bell state provides the probably archetypical example of a so-called maximally entangled state. Maximally entangled states reduce to the identity whenever being traced over, thus they posses by definition marginals of maximal entropy. The Bell state is only defined for bipartite qubit systems, so it is not very general. Luckily, the generalized GHZ-state provide a readily available generalization of the Bell states to arbitrary dimensions and particle numbers. This allows us to extend the above argument to the multi-partite scenario without any problems by writing

$$S(Tr_{X^c}(|GHZ(n,d)\rangle\langle GHZ(n,d)|)) = S(\frac{1}{d}\operatorname{Tr}_{X^c}(\sum_{i,j=0}^{d-1}|i\ldots i\rangle\langle j\ldots j|)) = S(\frac{1}{d}\mathbb{1}) = \log(d), \tag{5.27}$$

where X^c denotes the complement of X, meaning we trace out all systems except X. This is important due to different behavior of the GHZ marginals under less than n-1 partial traces. Tracing out

a smaller number of systems of the GHZ(n,d) leads to some kind of sub-identity as shown before 2.38. As opposed to whenever tracing over n-1 systems the GHZ(n,d) and Bell state marginals coincide.

Now that we have identified suitable maximizer state, we employ a simple construction method to form a basis of the whole state space from these candidate states. With this in mind we construct

$$Z^i \otimes X^{j_1} \otimes \ldots \otimes X^{j_1} |GHZ(n,d)\rangle\langle GHZ(n,d)|. \tag{5.28}$$

Why does this construction make sense? Recall that tenor products of local unitaries can not change the local entropies. Any state with maximally entropic marginals retains this property under local unitaries. When introducing the clock Z and shift X in chapter on Heisenberg-Weyl operators we have already remarked before 3.2.3, that their defining property is the unitarity of the clock Z and shift X operator. Therefore the local entropy is invariant under tensor products of Heisenberg-Weyl operators.

What a local unitary can do however is to rotate into orthogonal bases. Now as shown in Observation 3.2.3 the Heisenberg-Weyl operators are orthogonal and span the whole space. Thus rotating in all possible directions provided by the Heisenberg-Weyl operators will map the generalized GHZ to a whole basis of GHZ states. This allows a covering of all the state spaces dimensions starting from the candidate state, where all basis elements have maximally mixed single particle marginals.

Chapter 6

Conclusion

6.1 Conclusion & outlook

We have set out to explore the treatment of correlations within quantum mechanics. We have approached this fundamental concept in physics from two different angles: *entanglement theory* and *quantum thermodynamics*.

We begin with wrapping up our treatment of *quantum thermodynamics*.

In Chapter 5 we have discussed the influence of temperature and limited amount of available energy for the creation of correlations. We have derived conditions for unitaries that optimally create correlations under energy limitations. We propose to employ the concept of circulant majorization. Circulant matrices seem to simplify the search for such unitaries, but it may still be the case that other classes of matrices are even better suited. It may be worthwhile to get a deeper understanding of circulant matrices. For example understanding what type of transformations induce circulant matrices may be a good starting point to generate optimally correlation creating classes of transformations. Majorization is a well studied subject, that has been related to many topics such as unistochasticity and the theory of doubly stochastic matrices. It seems interesting to study our concept of circulant majorization in that context.

The largest part of this thesis was a discussion of how to best describe correlations when addressing problems in *entanglement theory*.

Multipartite entanglement detection is still not well understood. This thesis hopes to be a step towards tackling this challenging issue. In Chapter 4 we have introduced three separate criteria for multipartite entanglement detection: one based on the Hölder inequality and the multiplicativity of norms under the tensor product, one based on the correlation tensor 2-norm and one based on an anti-commutating basis. The performance of these criteria is basis dependent. For example, in Section 4.2 the repeated application of the Hölder inequality results in the appearance of many infinity norms in the entanglement detection criterion. If a state has a flat spectra in any given basis the infinity norms do not perform worse than other norms improving the criterion. The criterion in section 4.4 exhibits an even more extreme case of basis dependance. Within its region of

applicability the anti-commutativity criterion promises dimension and number of particle independent detection of entanglement. This is a tremendous simplification in the multipartite scenario. It comes at the price of an apparently restricted applicability to observables that are coverable by few sets of anti-commuting basis elements. The question how to find such sets remains open. The generalized Gell-Mann basis only has very small anti-commuting sets. The Heisenberg-Weyl type bases (including classical & symmetrized) have slightly larger sets, allowing for easier covering of observables. How to construct bases with large anti-commuting sets is an interesting question that we have to leave open.

We have proposed a Hermitian, non-unitary generalization of the Pauli matrices: the symmetrized Heisenberg-Weyl basis. Due to their properties they should be rather compared to the GGM than the similarly named HW operators. The symmetrized Heisenberg-Weyl basis outperforms the GGM in terms of the size of their non trivially anti-commuting sets. Furthermore, they exhibit a well behaved continuous limit and we hope to continue the work on this basis or related constructions. For example, Bloch vectors have yielded a notion of strict monogamy equalities for qubits [14], maybe similar notions can be derived for our new basis. Are there other full rank generalizations of Pauli basis besides the HWO? It is appealing that the symmetrized Heisenberg-Weyl matrices consist of observables, but is Hermiticity enough to ensure easy experimental implementability or should such a basis meet further specifications?

The availability of the generalized Gell-Mann and the Heisenberg-Weyl operators may suggest that the search for SU(d) generators is already satisfyingly settled. We are of the opinion that multipartite entanglement detection could benefit tremendously from continued investigation into new basis constructions for qudits. It may very well be the case that bases tailored towards the detection of specific states outperform the canonical choices of SU(d) generators. Even outside of entanglement detection it may be of interest to generalize different properties of the Pauli operators than unitarity and Hermiticity. We ask what kind of set of operators would (best) retain their pairwise anti-commutativity?

We end this thesis by stating that we would be very pleased to see that the questions we pose, result in the community reexamining the established theory of SU(d) generators. The same holds true for the long standing issue of the generalized Bloch decomposition. We believe that the Bloch picture should be further developed for qudit systems. Our use of the Bloch decomposition as a multipartite entanglement detection tool contributes toward this goal.

Chapter 7

Appendix

7.0.1 An upper bound on the number of exactly anticommuting observables

Please note that this appendix is no original work of the author of this thesis. It is the work of Otfried Gühne and has come up during discussions of the paper [2]. We restate it in the appendix, since it is of relevance to the question of finding anti-commuting basis sets. It provides a no-go result for one of the basis constructions proposed in this thesis.
In this section, we show that using the phase space-approach one can generate at most sets of three observables, which are mutually anti-commuting. In other words, if we consider a set of four observables, then they can not be pairwise anti-commuting.
To see this, consider a set of four observables, which are pairwise anti-commuting. We parametrize the four observables by two-dimensional real vectors $\vec{A}, \vec{B}, \vec{C}$ and $\vec{D}$, describing the displacement in phase space. The anti-commutativity leads to six conditions on the vectors, the first one reading

$$|\vec{A} \times \vec{B}| = k_1 \frac{\pi}{2} \tag{7.1}$$

with k_1 being an odd integer. The other five conditions are analogous, with odd numbers $k_2, \ldots, k_6$. Solving now the first five equations for $\vec{A}, \vec{B}, \vec{C}$ and $\vec{D}$ and inserting the resulting conditions into the sixth one, leads to the insight that the k_i have to obey:

$$k_1 k_6 + k_2 k_5 - k_3 k_4 = 0. \tag{7.2}$$

This condition, however, can not be fulfilled if all the k_i are odd.

Bibliography

[1] A. Acín, D. Bruß, M. Lewenstein, and A. Sanpera. Classification of mixed three-qubit states. *Phys. Rev. Lett.*, 87:040401, Jul 2001. https://arxiv.org/abs/quant-ph/0103025.

[2] A. Asadian, P. Erker, M. Huber, and C. Klöckl. Heisenberg-Weyl observables: Bloch vectors in phase space. *Phys. Rev. A*, 94:010301 (R), Jul 2016. http://arxiv.org/abs/1512.05640.

[3] R. A. Bertlmann and P. Krammer. Bloch vector for qudits. *J. Phys. A: Math.Theor. 41 235303*, 2008. http://arxiv.org/abs/0806.1174.

[4] F. Bloch. Nuclear induction. *Phys. Rev.*, 70:460–474, Oct 1946. http://link.aps.org/doi/10.1103/PhysRev.70.460.

[5] N. Brunner, D. Cavalcanti, S. Pironio, V. Scarani, and S. Wehner. Bell nonlocality. *Rev. Mod. Phys.*, 86:419–478, Apr 2014. http://arxiv.org/abs/1303.2849.

[6] J. Cadney, M. Huber, N. Linden, and A. Winter. Inequalities for the ranks of quantum states. *Linear Algebra and Applications, vol. 452, pp. 153-171*, 2014. https://arxiv.org/abs/1308.0539.

[7] E. Chitambar, D. Leung, L. Mancinska, M. Ozols, and A. Winter. Everything you always wanted to know about LOCC (but were afraid to ask). *Commun. Math. Phys., vol. 328, no. 1, pp. 303-326*, 2014. http://arxiv.org/abs/1210.4583.

[8] J. I. De Vicente. Separability criteria based on the Bloch representation of density matrices. *Quantum Info. Comput.*, 7(7):624–638, Sept. 2007. http://arxiv.org/abs/quant-ph/0607195.

[9] J. I. de Vicente. Further results on entanglement detection and quantification from the correlation matrix criterion. *Journal of Physics A: Mathematical and Theoretical*, 41(6):065309, 2008. http://arxiv.org/abs/0705.2583.

[10] J. I. de Vicente and M. Huber. Multipartite entanglement detection from correlation tensors. *Phys. Rev. A*, 84:062306, Dec 2011. http://arxiv.org/abs/1106.5756.

[11] J. I. de Vicente, C. Spee, and B. Kraus. Maximally entangled set of multipartite quantum states. *Phys. Rev. Lett.*, 111:110502, Sep 2013. http://arxiv.org/abs/1305.7398.

[12] W. Dür, G. Vidal, and J. I. Cirac. Three qubits can be entangled in two inequivalent ways. *Phys. Rev. A*, 62:062314, Nov 2000. http://arxiv.org/abs/quant-ph/0005115.

[13] A. Einstein, B. Podolsky, and N. Rosen. Can quantum-mechanical description of physical reality be considered complete? *Phys. Rev.*, 47:777–780, May 1935. http://link.aps.org/doi/10.1103/PhysRev.47.777.

[14] C. Eltschka and J. Siewert. Monogamy equalities for qubit entanglement from Lorentz invariance. *Phys. Rev. Lett.*, 114:140402, Apr 2015. http://arxiv.org/abs/1407.8195.

[15] C. Eltschka and J. Siewert. Quantifying entanglement resources. *J. Phys. A, Volume 47, Number 42*, 2016. http://arxiv.org/abs/1402.6710.

[16] U. Fano. Pairs of two-level systems. *Rev. Mod. Phys.*, 55:855–874, Oct 1983. http://link.aps.org/doi/10.1103/RevModPhys.55.855.

[17] R. Gallego, J. Eisert, and H. Wilming. Defining work from operational principles. *arXiv preprint arXiv:1504.05056*, 2015. http://arxiv.org/abs/1504.05056.

[18] J. Goold, M. Huber, A. Riera, L. del Rio, and P. Skrzypczyk. The role of quantum information in thermodynamics — a topical review. *Journal of Physics A: Mathematical and Theoretical, Volume 49, Number 14*, 2015. http://arxiv.org/abs/1505.07835.

[19] S. K. Goyal, B. N. Simon, R. Singh, and S. Simon. Geometry of the generalized Bloch sphere for qutrits. 2016. http://arxiv.org/pdf/1111.4427v2.pdf.

[20] O. Gühne and G. Toth. Entanglement detection. *Physics Reports 474, 1*, 2009. http://arxiv.org/pdf/0811.2803.pdf.

[21] L. Gurvits. Classical deterministic complexity of Edmond's problem and quantum entanglement. In *Proceedings of the Thirty-fifth Annual ACM Symposium on Theory of Computing*, STOC '03, pages 10–19, New York, NY, USA, 2003. ACM. https://arxiv.org/abs/quant-ph/0303055.

[22] A. S. M. Hassan and P. S. Joag. Separability criterion for multipartite quantum states based on the Bloch representation of density matrices. *Quantum Information and Computation, vol. 8, 0773-0790*, 2008. http://arxiv.org/abs/0704.3942.

[23] W. Helwig, W. Cui, J. I. Latorre, A. Riera, and H.-K. Lo. Absolute maximal entanglement and quantum secret sharing. *Phys. Rev. A, vol. 86, 052335*, 2012. https://arxiv.org/abs/1204.2289.

[24] F. T. Hioe and J. H. Eberly. n-level coherence vector and higher conservation laws in quantum optics and quantum mechanics. *Phys. Rev. Lett.*, 47:838–841, Sep 1981. http://link.aps.org/doi/10.1103/PhysRevLett.47.838.

[25] M. Horodecki, P. Horodecki, and R. Horodecki. Separability of mixed states: necessary and sufficient conditions. *Physics Letters A*, 223, 1996. http://arxiv.org/abs/quant-ph/9605038.

[26] R. Horodecki, P. Horodecki, M. Horodecki, and K. Horodecki. Quantum entanglement. *Rev. Mod. Phys.*, 81:865–942, Jun 2009. https://arxiv.org/abs/quant-ph/0702225.

[27] M. Huber and C. Klöckl. Characterizing multipartite entanglement without shared reference frames. *Phys. Rev. A 91, 042339*, 2015. http://arxiv.org/abs/1411.5399.

[28] M. Huber, M. Perarnau-Llobet, K. V. Hovhannisyan, P. Skrzypczyk, C. Klöckl, N. Brunner, and A. Acín. Thermodynamic cost of creating correlations. *New J. Phys. 17, 065008*, 2015. http://arxiv.org/abs/1404.2169.

[29] S. Jevtic, D. Jennings, and T. Rudolph. Maximally and minimally correlated states attainable within a closed evolving system. *Phys. Rev. Lett.*, 108:110403, Mar 2012. http://arxiv.org/abs/1110.2371.

[30] J. Kaniewski, M. Tomamichel, and S. Wehner. Entropic uncertainty from effective anticommutators. *Phys. Rev. A*, 90:012332, Jul 2014. https://arxiv.org/abs/1402.5722.

[31] G. Kimura. The Bloch vector for n-level systems. *Phys. Lett. A 314, 339*, 2003. http://arxiv.org/abs/quant-ph/0301152.

[32] G. Kimura and A. Kossakowski. The Bloch vector space for n-level systems – the spherical-coordinate point of view. *Open Sys. Information Dyn. 12, 207*, 2005. http://arxiv.org/abs/quant-ph/0408014.

[33] P. Kurzynski. Multi-Bloch vector representation of the qutrit. *Quantum Inf. Comp. 11, 361*, 2011. http://arxiv.org/pdf/0912.3155v1.pdf.

[34] P. Kurzynski, A. Kolodziejski, W. Laskowski, and M. Markiewiczi. Three-dimensional visualisation of a qutrit. 2016. http://arxiv.org/pdf/1601.07361v1.pdf.

[35] P. Kurzynski, T. Paterek, R. Ramanathan, W. Laskowski, and D. Kaszlikowski. Correlation complementarity yields Bell monogamy relations. *Phys. Rev. Lett.*, 106:180402, May 2011. https://arxiv.org/abs/1010.2012.

[36] A. W. Marshall and I. Olkin. *Inequalities: Theory of Majorization and Its Applications*. Springer, New York, 1979.

[37] A. Misra, U. Singh, S. Bhattacharya, and A. K. Pati. Energy cost of creating quantum coherence. *Phys. Rev. A 93, 052335*, 2016. http://arxiv.org/abs/1602.08437.

[38] M. A. Nielsen. Conditions for a class of entanglement transformations. *Physical Review Letters*, 83:436, 1999. http://arxiv.org/abs/quant-ph/9811053.

[39] M. A. Nielsen and I. L. Chuang. *Quantum Computation and Quantum Information*. Cambridge University Press, New York, NY, USA, 10th edition, 2010. http://www.michaelnielsen.org/qcqi/QINFO-book-nielsen-and-chuang-toc-and-chapter1-nov00.pdf.

[40] C. F. Senel, T. Lawson, M. Kaplan, D. Markham, and E. Diamanti. Demonstrating genuine multipartite entanglement and nonseparability without shared reference frames. *Phys. Rev. A 91, 052118*, 2015. http://arxiv.org/pdf/1408.1837.pdf.

[41] G. Sergiol, E. Santucci, L. Didac, J. law A. Miszczak, and R. Giuntini. Pattern recognition on the quantum Bloch sphere. 2016. http://arxiv.org/pdf/1603.00173.pdf.

[42] S. Severini. On the digraph of a unitary matrix. *SIAM Journal on Matrix Analysis and Applications, Volume 25, Number 1, pp. 295-300*, 2003. http://arxiv.org/abs/math/0205187.

[43] M. Soleymani and A. Armandnejad. Linear preservers of circulant majorization on. *Linear Algebra and its Applications*, 440:286 – 292, 2014.

[44] W. Tadej and K. Zyczkowski. Defect of a unitary matrix. *Linear Algebra and its Applications, vol. 429 issues 2-3, pp. 447-481*, 2008. http://arxiv.org/abs/math/0702510.

[45] P. Talkner and P. Hänggi. Aspects of quantum work. *Phys. Rev. E*, 93:022131, Feb 2016. http://arxiv.org/abs/1512.02516.

[46] G. Tóth and O. Gühne. Detecting genuine multipartite entanglement with two local measurements. *Phys. Rev. Lett.*, 94:060501, Feb 2005. https://arxiv.org/abs/quant-ph/0405165.

[47] M. C. Tran, B. Dakic, F. Arnault, W. Laskowski, and T. Paterek. Quantum entanglement from random measurements. *Phys. Rev. A 92, 050301(R)*, 2015. https://arxiv.org/abs/1411.4755.

[48] M. C. Tran, B. Dakic, W. Laskowski, and T. Paterek. Correlations between outcomes of random observables. 2016. https://arxiv.org/abs/1605.08529.

[49] Z. Wang, C. Zhang, Y.-F. Huang, B.-H. Liu, C.-F. Li, and G.-C. Guo. Experimental verification of genuine multipartite entanglement without shared reference frames. *Sci. Bull. 61.9 714-719*, 2016. http://link.springer.com/article/10.1007/s11434-016-1063-5.

[50] H. Weyl. Quantenmechanik und gruppentheorie. *Zeitschrift für Physik*, 46(1):1–46, 1927. http://dx.doi.org/10.1007/BF02055756.

[51] M. Wieśniak and K. Maruyama. Package of facts and theorems for efficiently generating entanglement criteria for many qubits. *Phys. Rev. A*, 85:062315, Jun 2012. http://arxiv.org/abs/1204.2722.

[52] M. M. Wilde. *Quantum Information Theory.* Cambridge University Press, 2013. https://arxiv.org/abs/1106.1445.