Ludwig-Maximilians-Universität München

Master Thesis

Reconstruction and analysis of multi-photon entangled quantum states

Author:
Lukas Knips

Supervisor:
Prof. Dr. Harald Weinfurter

A thesis submitted in fulfilment of the requirements
for the Master of Science

in the

Working group of Prof. Dr. Weinfurter
Chair Prof. Dr. Hänsch

August 2013

LUDWIG-MAXIMILIANS-UNIVERSITÄT MÜNCHEN

MASTER THESIS

Rekonstruktion und Analyse von verschränkten Multiphotonen-Quantenzuständen

---

Author:
Lukas KNIPS

Supervisor:
Prof. Dr. Harald WEINFURTER

A thesis submitted in fulfilment of the requirements
for the Master of Science

in the

Working group of Prof. Dr. Weinfurter
Chair Prof. Dr. Hänsch

August 2013

Declaration of Authorship

Erklärung:

Hiermit erkläre ich, Lukas KNIPS, die vorliegende Arbeit selbstständig verfasst zu haben und keine anderen als die in der Arbeit angegebenen Quellen und Hilfsmittel benutzt zu haben.

München, ______________________________

Unterschrift: ______________________________

LUDWIG-MAXIMILIANS-UNIVERSITÄT

Abstract

Faculty of Physics
Working group of Prof. Dr. Weinfurter

Master of Science

Reconstruction and analysis of multi-photon entangled quantum states

by Lukas Knips

States with genuine entanglement are not necessarily exhibiting correlations between the measurement outcomes of all involved parties what might seem paradox at first. This thesis studies an exemplary three qubit state with these properties - being genuinely tripartite entangled and showing no correlations between the outcomes of all parties (vanishing *full correlations*). This property is not restricted to the one considered state, but will be generalised to a class of states. Furthermore, this thesis considers different quantum state reconstruction methods. Comparing the quality of those estimators, the usually employed methods are shown to be biased, i.e. showing a systematical deviation between the theoretical prediction and the estimated value. A procedure will be introduced how non-linear, but convex or concave functions can be bounded by a linear approximation. Finally, another quantum state estimator based on modifying the eigenvalues of a matrix will be introduced.

Acknowledgements

First of all, I want to thank Prof. Dr. Harald Weinfurter for the opportunity to write this thesis in his group. Thank you for all your support. Furthermore, I am very grateful for Christian Schwemmer's continous assistance in every concern of my work. I am also much obliged for Daniel Richart's aid. My thanks also go to Dr. Wiesław Laskowski, Dr. Tomasz Paterek, Dr. Marcin Wiesniak, Dr. Tobias Moroder, Dr. Matthias Kleinmann, Prof. Dr. Otfried Gühne and Prof. Dr. Géza Tóth for our successful cooperations. Besides, I additionally thank the whole working group as well as all the people in Garching. And - of course - thank you, Alex! Moreover, I give props to my parents. Additionally, I am deeply grateful for your long-lasting support, Heinrich, which afforded a lot. Thank you. Furthermore, I acknowledge the support of the German National Academic Foundation for the financial and ideological sponsorship.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Quantum entanglement [1, 2] as a fundamental concept of quantum information processing establishes a widespread field of research. The counter-intuitive behaviour of entangled states, with classical thinking not comprehensible, empoweres ways for improved metrology [3, 4], computation [5–7] and communication and cryptography [8, 9].

The main topics of this thesis will be discussed after a short introduction of some fundamental concepts. In the next chapter, chapter 2 - "Fundamentals", an overview of the used notation and the concepts of entanglement will be given. Furthermore, important quantities will be discussed that are used to estimate for example the quality of a prepared quantum state with respect to its theoretical state. Moreover, a short survey of different methods to detect entanglement will be made [10]. After summing up some quantum states of relevance for this thesis, experimental requirements will be treated such as the creation of entangled photon pairs.

Intuitively, one would assume that an entangled state exhibits correlations between its measurement outcomes. Chapter 3 - "No Correlations" - will treat the question if it is possible to find a genuinely tripartite entangled state where no full correlations between the measurement outcomes of all parties are present. Consider a state with three qubits and the respective measuring parties denoted by Alice, Bob and Charlie. If Alice and Bob perform measurements onto their respective qubits, can Charlie infer his measurement outcome by the knowledge of the others' outcomes? As will be further illustrated in that chapter, Charlie is able to infer his measurement result if both, Alice and Bob, tell him their concrete outcomes. Instead, if Alice and Bob talk to each other about their results and tell Charlie only if they obtained the same or opposite results, Charlie cannot necessarily infer anything about his respective outcome. The discrimination between full knowledge of all other outcomes and knowing only the parity of these outcomes will be clarified by an exemplary state. While this effect was already purely theoretically studied [11], this thesis deepens the analysis for an exemplary state. A more general class of states without any n-partite correlations is discussed in [12]. Here, the theoretical considerations are generalised to another class of states with genuine n-partite entanglement, but without any n-partite correlations [13].

Another topic of this thesis is the investigation of quantum state reconstruction methods [14–20]. All measurements performed in the laboratory come intrinsically only with finite statistics [20]. Even if the experiment would be operated with perfect devices and accurate alignments, some of the reconstructed states would not represent a physical state [15]. Because the eigenvalues of the density matrix of a physical state correspond to probabilities, these eigenvalues are expected to be bounded between zero and unity. This interpretation fails for non-physical results. Furthermore, some non-linear functions cannot be evaluated if the physicality constraint is not fulfilled. To be able to call the obtained result a density matrix and use it for further analysis, one is interested in finding the physical quantum state that is in best agreement with the measured data [16]. Chapter 4 - "Bias of Estimators" - discusses different methods [15–18] to find the (physical) quantum state that, e.g., most likely produced the measured data. It furthermore inspects the quality of the obtained quantum state. To fulfil this quality research, one needs a large data set for investigation. For that purpose, this chapter is based on numerical simulations [21], bringing the advantage that one knows the underlying quantum state and allowing to compare the simulated data with the expectation values. The quality of commonly used quantum state estimators will be studied as well as the widespread bootstrapping methods [22] to infer statistical information. This chapter will present a method to evaluate non-linear, but convex or concave functions without estimating the underlying quantum state [23]. Using a linear approximation delivers unbiased results. Furthermore, the calculation of confidence regions is based on the Hoeffding calculus [24] and does not rely on further assumptions. Finally, another state estimator will be introduced that is able to provide promising results for a subset of states. This study will show the gain of quality of the state estimation if a priori knowledge is used.

After a short conclusion of some of the insights of this thesis (chapter 5) calculations closely related to the topics of the main text, but not absolutely necessary for understanding it, are done in appendix A - "Additional calculations". More detailed information concerning the used convex optimisation for quantum state reconstruction is given in appendix B - "Convex optimisation". Finally, this thesis concludes with a list of the used references.

Chapter 2

Fundamentals

2.1 Qubits

While classical computation is based on *bits* [25] - a short term for binary digit - , quantum computation makes use of so called *qubits* (quantum bits) [26]. A good starting point for a suited quantum system is to think of two states of a system: a ground state and an excited state, which will be denoted in Dirac's Bra-Ket notation [27] as $|0\rangle$ and $|1\rangle$, respectively. A classical bit can either be in the ground state or in the excited state while in contrast the state of the qubit can be any superposition of the states $|0\rangle$ and $|1\rangle$:

$$|\psi\rangle = \alpha|0\rangle + \beta|1\rangle, \tag{2.1}$$

with $\alpha, \beta \in \mathbb{C}$ and $\langle\psi|\psi\rangle = 1$. Since $|\alpha|^2 + |\beta|^2 = 1$ must hold, $|\psi\rangle$ can also be written as $|\psi\rangle = e^{i\gamma}\left(\cos\left(\frac{\theta}{2}\right)|0\rangle + \sin\left(\frac{\theta}{2}\right)e^{i\phi}|0\rangle\right)$ with $\gamma, \theta, \phi \in \mathbb{R}$ [5]. The global phase $e^{i\gamma}$ cannot be observed whereby the state can be expressed with the help of the transformation $|\psi\rangle \to e^{-i\gamma}|\psi\rangle$ as [28]

$$|\psi\rangle = \cos\left(\frac{\theta}{2}\right)|0\rangle + \sin\left(\frac{\theta}{2}\right)e^{i\phi}|0\rangle. \tag{2.2}$$

The parametrisation by means of the angles $\theta \in [0, \pi]$ and $\phi \in [0, 2\pi]$ enables a spherical illustration of the state. This geometric representation is called *Bloch sphere* [29], which is shown in Fig. 2.1 with an exemplary Bloch vector $|\psi\rangle$. The states $|0\rangle$ and $|1\rangle$ correspond to the poles of the sphere. In contrast, states of the form $\frac{|0\rangle + e^{i\phi}|1\rangle}{\sqrt{2}}$ (i.e. $\theta = \pi/2$) lie on its equator.

Similar to the basis $\{|0\rangle; |1\rangle\}$ used above, all sets of two orthonormal states can be used as a basis. The three most important ones are given by the eigenvectors of the frequently used Pauli matrices which span up the corresponding basis. Throughout this thesis, the Pauli matrices are defined by

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \text{ and } \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{2.3}$$

It is also useful to define another matrix σ_0 which equals the 2×2-unity matrix:

$$\sigma_0 = \mathbb{I} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \tag{2.4}$$

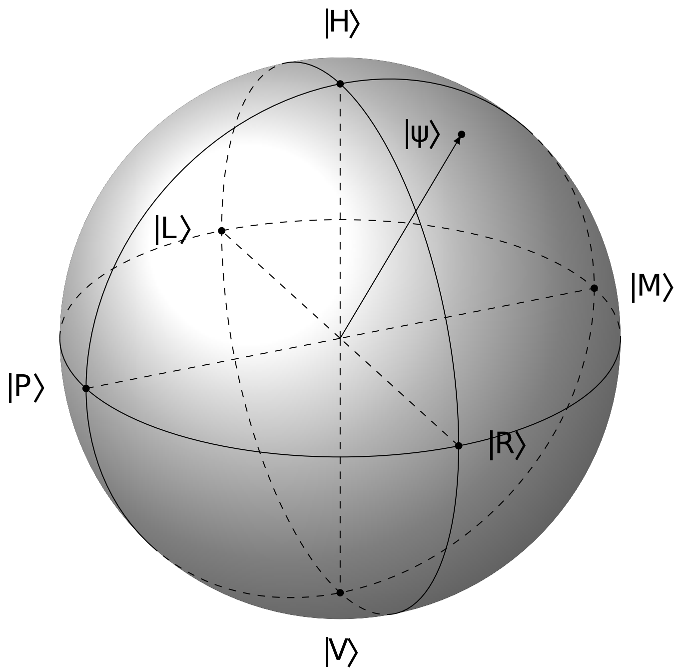

FIGURE 2.1: The bloch sphere illustrating a one qubit state $|\psi\rangle$. Pure states lie on the sphere's surface as the shown exemplary state. In general, the radius is a measure for the purity of the state.

To improve the readability of some expressions, the matrices $\{\sigma_0, \sigma_x, \sigma_y, \sigma_z\}$ may also be addressed via a scalar index[1] such that $\sigma_1 \equiv \sigma_x, \sigma_2 \equiv \sigma_y, \sigma_3 \equiv \sigma_z$.

The qubit formalism can be used for instance to describe the polarisation degree of freedom of photons [30]. In this case, the horizontal polarisation $|H\rangle$ can be identified with the ground state $|0\rangle$ while the vertical polarisation $|V\rangle$ corresponds to the excited state $|1\rangle$. $|H\rangle$ and $|V\rangle$ are the eigenstates to the Pauli matrix σ_z:

$$\sigma_z|H\rangle = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} = +|H\rangle \tag{2.5}$$

$$\sigma_z|V\rangle = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = - \begin{pmatrix} 0 \\ 1 \end{pmatrix} = -|V\rangle. \tag{2.6}$$

According to Eq. 2.5 and 2.6, the eigenstates of σ_x and σ_y can be determined:

$$\sigma_x|P\rangle = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = +|P\rangle \tag{2.7}$$

$$\sigma_x|M\rangle = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} -1 \\ 1 \end{pmatrix} = -|M\rangle \tag{2.8}$$

$$\sigma_y|R\rangle = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ i \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ i \end{pmatrix} = |R\rangle \tag{2.9}$$

$$\sigma_y|L\rangle = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -i \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} -1 \\ i \end{pmatrix} = -|L\rangle, \tag{2.10}$$

---

[1] Depending on the context, either the numerical ("0", "1", "2", "3") or the alphanumerical ("0", "x", "y", "z") indexing will be preferred. Both ways are fully equivalent.

where the *plus* state $|P\rangle$ and the *minus* state $|M\rangle$ are the eigenstates corresponding to a linear diagonal polarisation. In contrast, the *right* state $|R\rangle$ and the *left* state $|L\rangle$ are used to describe right and left circular polarisation, respectively. The set of bases $\{\sigma_x, \sigma_y, \sigma_z\}$ corresponds to the basis vectors $\{\vec{e}_x, \vec{e}_y, \vec{e}_z\}$ (as for example used by the illustration of the Bloch sphere) since there is a homomorphism from the special unitary group $SU(2)$ to the special orthogonal group $SO(3)$ [31].

2.2 Mixed states

Since the previous discussion treated only pure states, a more general approach to describe states is required. The *density operator* or *density matrix* formalism helps to describe incoherent summations of pure states. If some pure states $|\psi_i\rangle$ are mixed, each with a corresponding probability p_i, the mixed state can be expressed as [5]

$$\rho = \sum_i p_i |\psi_i\rangle\langle\psi_i|, \tag{2.11}$$

where the probabilities $p_i \geq 0$ have to sum up to 1. Because of $|\psi_i\rangle$ being normalised the trace of ρ equals

$$\mathrm{tr}(\rho) = \mathrm{tr}\left(\sum_i p_i |\psi_i\rangle\langle\psi_i|\right) = \sum_i p_i \,\mathrm{tr}\left(|\psi_i\rangle\langle\psi_i|\right) = \sum_i p_i = 1. \tag{2.12}$$

While pure states $\rho = |\psi\rangle\langle\psi|$ are idempotent, i.e. $\rho^2 = \rho$ holds[2], in general the condition $\mathrm{tr}\,\rho^2 \leq 1$ is fulfilled [28]. Because $\langle\phi_i|\rho|\phi_i\rangle$ corresponds to the probability to find ρ in the state $|\phi_i\rangle$, $\langle\phi_i|\rho|\phi_i\rangle \in [0,1]$ must hold $\forall|\phi_i\rangle$. Therefore, it is obvious that ρ has to be positive semi-definite, i.e. $\rho \geq 0$.

It is also possible to parametrise ρ in terms of a Bloch vector $\vec{r} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$:

$$\rho = \frac{\mathbb{I}}{2} + \frac{1}{2}\vec{r}\vec{\sigma} = \frac{1}{2}\begin{pmatrix} 1+z & x-iy \\ x+iy & 1-z \end{pmatrix}, \text{ where } \vec{\sigma} = \begin{pmatrix} \sigma_x \\ \sigma_y \\ \sigma_z \end{pmatrix} \tag{2.13}$$

is composed by the Pauli matrices. While for pure states $\rho = |\psi\rangle\langle\psi|$ the Bloch vector is of unit length, mixed states lie on the inside of the Bloch sphere. If instead of the Bloch vector $\vec{r}$ a first order tensor $\hat{T}$ is used, where $T_i \in [-1,1]$ with $i \in \{0,1,2,3\}$ and $T_0 = 1$, Eq. 2.13 becomes

$$\rho = \frac{1}{2}\sum_{i=0}^{3} T_i \sigma_i. \tag{2.14}$$

Since $\mathrm{tr}\,\rho^2 \leq 1$ must hold for valid density operators, there are also constraints to the elements of the tensor $\hat{T}$ such that Eq. 2.14 still describes a physical state. Since $T_0 = 1$ holds, it can directly deduced that $\sum_{i=0}^{3} T_i^2 \leq 2$ must be valid for all single qubit density matrices. The derivation for this constraint will follow for the general multiqubit case in Eq. 2.20.

---

[2] since $\rho^2 = |\psi\rangle\langle\psi||\psi\rangle\langle\psi| = |\psi\rangle\langle\psi|$

2.3 Multiqubits and Entanglement

States that are composed by more than one qubit are called multiqubit states. In contrast to the basis of a pure single qubit state, that is according to Eq. 2.1 spanned by only two orthonormal states, the basis of states with n qubits is now composed of 2^n orthonormal elements. If this composite quantum system only consists of n distinct physical systems $\mathcal{S}_i$, its state can be written as [5]

$$|\psi\rangle = \bigotimes_{i=1}^{n} |\phi_i\rangle \tag{2.15}$$

with $|\phi_i\rangle \in \mathcal{H}_{\mathcal{S}_i}$, where $\mathcal{H}_{\mathcal{S}_i}$ denotes the Hilbert space of the system $\mathcal{S}_i$. States that can be decomposed into all subsystems according to Eq. 2.15 are called *fully separable* or *product states* [28, 32]. If the given pure state is not of the form of Eq. 2.15, but instead of the more general form and composed of at least two terms

$$|\psi\rangle = \sum_{i_1,\ldots,i_n} \alpha_{i_1,\ldots,i_n} \bigotimes_{j=1}^{n} |\phi_{i_j}\rangle, \tag{2.16}$$

this state is *entangled* [33]. If it is possible to find a basis such that all but one coefficients $\alpha_{i_1,\ldots,i_n}$ vanish, the left over coefficient is normalised to 1, Eq. 2.16 simplifies to Eq. 2.15 and the state is fully separable. To illustrate this non-separability of entangled states, it is useful to look at an exemplary state like the $|\phi^+\rangle$ (*Bell state* or *EPR pair*) state [5] (Sec. 2.5.1):

$$|\phi^+\rangle = \frac{|0\rangle_A \otimes |0\rangle_B + |1\rangle_A \otimes |1\rangle_B}{\sqrt{2}}, \tag{2.17}$$

where the indices A and B denote the first and the second qubit, which shall be given to the parties called Alice and Bob. Assuming Alice wants to measure the state of her qubit, she projects it onto the possible local results and obtains that her qubit is with the same probability of $\frac{1}{2}$ in each state $|0\rangle_A$, $|1\rangle_A$. If for example she measured her qubit being in state $|0\rangle_A$, she directly knows the state of Bob's qubit to be $|0\rangle_B$. Therefore, it is manifest to call the measurement outcomes of Alice and Bob *correlated* [5].

Entanglement ("Verschränkung" [34]) is the canonical element to observe such correlations. It is also the main resource of quantum information processing [2, 5]. Knowing the measurement outcome of remote qubits - presumably instantaneously, but at least orders of magnitude faster than the speed of light [35] - just by measuring on another qubit, gives rise to fundamental questions. Is the correlation, that can occur in quantum systems, in any mean different from classical correlations [36]? Can the correlation of the measurement outcomes of a quantum state be explained by classical models, i.e. does the experimentalist only investigate inherent attributes of the particles, that exist independently of the measurement process [37]? Since it is possible to determine a bound that is fulfilled by local-realistic theories, but beaten by theories that are not both local and realistic [38], Einstein's explanation of local hidden variables can be falsified. These inequalities are called *Bell inequalities* [39]. The correlation of Alice's and Bob's measurement outcomes cannot be explained classically, i.e. with local-realistic theories. Finding vivid descriptions for entanglement is intrinsically difficult due to its non-classicality. Enlightening publications concerning this topic are for example [40–42]. The at first glance paradox situation of the "quantum kitchen" [42] will be briefly recapitulated here. Consider two ovens producing each one soufflé. The only possibility to inspect the cake is by opening the oven making further baking impossible. Assume that the following statements hold: if both bakers, A and B, open their ovens midway, their soufflés already rose in 9% of the cases. Furthermore, assume that if one cook opens his oven during baking and sees his cake to be risen early, the other one's cake is delicious after baking. Classically, one would infer that in at least 9% of the cases both soufflés succeed and taste good. Instead, this lower bound can be beaten by an appropriately constructed quantum state up to the case that both cakes never taste good [42].

Up to now, with Eq. 2.16 only pure multiqubit states could be described. Instead, a multiqubit expression for the density operator formalism is needed, that can describe multiqubit mixed states. The usage of Pauli matrices as a basis for density operators as given in Eq. 2.14 leads directly to a general expression for n qubit density operators [43]:

$$\rho = \frac{1}{2^n} \sum_{\mu_1,\ldots,\mu_n=0}^{3} T_{\mu_1\ldots\mu_n} \bigotimes_{i=1}^{n} \sigma_{\mu_i}. \tag{2.18}$$

The tensor $\hat{T}$ before of first order, is now of order n with real entries $T_{\mu_1\ldots\mu_n} \in [-1,1]$ with indices $\mu_i \in \{0,1,2,3\}$ for $i = 1,\ldots,n$. Because of the properties of Pauli matrices[3] the entries of the tensor $\hat{T}$ can be obtained easily from Eq. 2.18 by calculating

$$T_{\mu_1\ldots\mu_n} = \mathrm{tr}\left[\rho\left(\bigotimes_{i=1}^{n} \sigma_{\mu_i}\right)\right]. \tag{2.19}$$

Since $T_{\mu_1\ldots\mu_n}$ represents the correlation of the measurement outcomes in the particular basis $\bigotimes_{i=1}^{n} \sigma_{\mu_i}$, the tensor $\hat{T}$ is called *correlation tensor* [44].
According to the constraint of $\sum_{i=0}^{3} T_i^2 \leq 2$ for the $n = 1$ qubit case, which was already mentioned earlier, this bound for the correlation tensor elements can be calculated for arbitrary number of qubits. The sum of squared correlation tensor elements must fulfil

$$\sum_{\mu_1,\ldots,\mu_n=0}^{3} T^2_{\mu_1\ldots\mu_n} \leq 2^n, \tag{2.20}$$

which will be derived in A.1. This bound is saturated for instance by the already mentioned Bell state $|\phi^+\rangle$ since in this case $T_{00} = 1$, $T_{XX} = 1$, $T_{YY} = -1$, $T_{ZZ} = 1$ and $T_{ij} = 0$ for all other cases.
Note that - exactly as the Pauli matrices - the correlation tensor elements will be addressed by numerical as well as by alphanumerical indexing, for instance $T_{11} \equiv T_{XX}$.

2.3.1 Types of entanglement

Previously, states were described to be either entangled or separable. While two qubit states can only be on one hand entangled or on the other hand separable, for states with higher numbers of qubits a more diverse classification of entanglement is possible. Here, this will be exemplary discussed for the case of states with three qubits, where the qubits will be denoted by "A", "B" and "C". The state is a fully-separable product state if the state is composed of a tensorial product of $A \otimes B \otimes C \hat{=} A|B|C$ according to Eq. 2.15. Furthermore, the state is *bi-separable* if it is of the form $A|BC$ (or $B|AC$ or $C|AB$) [28]. While these bi-product states are already entangled, finally, another class of states exist.

---

[3] The properties $\mathrm{tr}(\sigma_x) = \mathrm{tr}(\sigma_y) = \mathrm{tr}(\sigma_z) = 0$, $\mathrm{tr}(\sigma_0) = 2$ and $\sigma_i\sigma_j = \delta_{ij}\sigma_0 + i\epsilon_{ijk}\sigma_k$ are needed here.

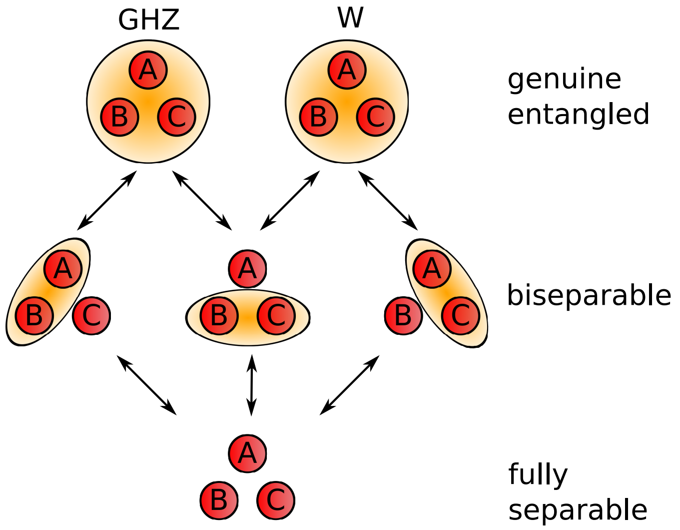

FIGURE 2.2: Different classes of entanglement for three qubit states. States can be either fully separable, bi-separable or genuinely tripartite entangled. The entanglement of the qubits, which are symbolised by the red circles, is depicted by the orange groupings. Note that genuinely tripartite entangled states can also be split into two different classes of states [45].

The state may also be of the form ABC, where the state cannot be decomposed into two or more states at all. Thus, these states are genuinely tripartite entangled [28]. Please note that there are also different classes of genuine tripartite entangled states, as one can still distinguish between the entanglement of the class of $|\text{GHZ}\rangle$ states (see Sec. 2.5.2) and of the $|\text{W}\rangle$ states (see Sec. 2.5.4) [45, 46]. This classification is sketched in Fig. 2.2.

2.4 Quantities of interest

After preparing and measuring a quantum state, it is often desired to study the state with regard to certain measures. In the following an overview of important functions to be evaluated on quantum states will be presented. This section starts with different ways how to detect entanglement of quantum states.

2.4.1 Detecting entanglement

Although discriminating entangled and separable states may seem to be a simple task for a low number of qubits and ideal systems like the Bell state in Eq. 2.17, entanglement detection and quantification can indeed be rather challenging. Since there is a multitude of detection schemes, only a short excerpt over some methods will be given. First, a list of qualitative detection methods will be presented while later on also quantitative measures will be introduced when discussing interesting functions of quantum states. For a more detailed overview, please see [1, 10, 47].

Scalar product on correlation tensor

Analogously to Eq. 2.15, separable mixed states can be decomposed into their separate density matrices

$$\rho = \sum_i p_i \bigotimes_{j=1}^{n} \rho_i^{(j)}, \tag{2.21}$$

with probabilities $p_i \geq 0$ and $\sum_i p_i = 1$ [44]. For these product states the correlation tensor $\hat{T}$ factorises:

$$\hat{T}^{\text{sep}} = \sum_i p_i \hat{T}_i^{\text{prod}} \tag{2.22}$$

with $\hat{T}_i^{\text{prod}} = \bigotimes_{j=1}^{n} \hat{T}_i^{(j)}$. Additionally, this factorisation of product states can be used to detect entanglement. For that purpose, an inner product (scalar product) is defined as

$$\left(\hat{X}, \hat{Y}\right)_G = \sum_{\mu_1,\ldots,\mu_n,\nu_1,\ldots,\nu_n=0}^{3} X_{\mu_1,\ldots,\mu_n} Y_{\nu_1,\ldots,\nu_n} G^{\mu_1,\ldots,\mu_n,\nu_1,\ldots,\nu_n} \tag{2.23}$$

with the positive semi-definite metric G [44]. Here, the metric

$$G^{\mu_1,\ldots,\mu_n,\nu_1,\ldots,\nu_n} = \delta^{\mu_1,\nu_1} \ldots \delta^{\mu_n,\nu_n} \tag{2.24}$$

may be used. Consider a state ρ described by a correlation tensor $\hat{T}$. If ρ is a product state, there exists a state ρ_{prod} with a correlation tensor $\hat{T}^{\text{prod}}$ such that $(\hat{T}, \hat{T}^{\text{prod}})_G = (\hat{T}, \hat{T})_G$. Thus, if the expression

$$\max_{\hat{T}^{\text{prod}}} \left(\hat{T}, \hat{T}^{\text{prod}}\right)_G < \left(\hat{T}, \hat{T}\right)_G \tag{2.25}$$

holds, the state ρ with correlation tensor $\hat{T}$ is not a product state. If instead of ρ_{prod} the state is compared with a bi-separable state $\rho_{\text{bi-sep}}$, genuine entanglement can be detected. Instead of the here used metric G it is already sufficient to find any metric such that $\max_{\hat{T}^{\text{prod}}} \left(\hat{T}, \hat{T}^{\text{prod}}\right)_G < \left(\hat{T}, \hat{T}\right)_G$ holds. For details, please see [44, 48].

Schmidt decomposition

Since pure states $|\psi\rangle$ written in the form of Eq. 2.16 are entangled if $|\psi\rangle$ is composed of more than one $\bigotimes_{j=1}^{n} |\phi_{i_j}\rangle$ for orthonormal basis states, also the *Schmidt decomposition* [32, 49] can be used for detecting entanglement. For two qubits, a pure state may be decomposed into [50]

$$|\psi_S\rangle = \cos(\theta)|a\rangle \otimes |b\rangle + \sin(\theta)|a_\perp\rangle \otimes |b_\perp\rangle \tag{2.26}$$

with the local bases $\{|a\rangle, |a_\perp\rangle\}$ for Alice and $\{|b\rangle, |b_\perp\rangle\}$ for Bob with $\theta \in [0, \pi/4]$ [50]. As soon as $\theta > 0$, $|\psi_S\rangle$ is entangled.

Linear Entanglement Witness

A different type of entanglement criterion can be established by *entanglement witnesses*. An observable $\hat{W}$ is called (entanglement) witness if [10]

- all separable states ρ_{sep} give $\text{tr}(\hat{W}\rho_{\text{sep}}) > 0$ and
- at least one entangled state ρ_{ent} fulfils $\text{tr}(\hat{W}\rho_{\text{ent}}) < 0$.

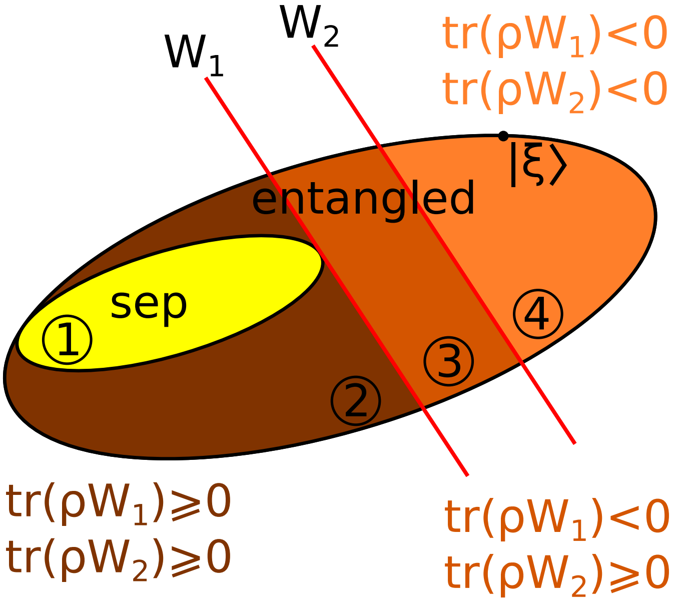

FIGURE 2.3: Illustration of an entanglement witness. The space of separable states is a subspace of the entire state space. Both witnesses W_1 and W_2 can distinguish some entangled states from all separable ones. Obviously, also entangled states can deliver $\text{tr}(W\rho) \geq 0$ and are consequently not detected to be entangled. Separable states belong to the space of states denoted with (1). As well as some entangled states (those of group (2)) the separable states give $\text{tr}(W\rho) \geq 0$ for both shown witnesses. The states of (3) are already detected to be entangled by the witness W_1, but not by W_2. Therefore, W_1 is a finer witness. (4) corresponds to the states detected by both, W_1 and W_2, to be entangled.

Therefore, $\text{tr}(\hat{W}\rho) < 0$ is a sufficient, but not necessary condition to detect entanglement, which is apparent from Fig. 2.3. An often used approach to construct a witness $\hat{W}$ is by projecting ρ onto a pure genuine n-partite entangled target state $|\xi\rangle$. Then, the maximal overlap of $|\xi\rangle$ with separable states is considered to detect entanglement for arbitrary states. If the overlap of ρ with $|\xi\rangle$ is larger than $|\xi\rangle$ with all separable states, ρ is entangled. Therefore, $\hat{W}$ is constructed according to [51] such that

$$\hat{W}_{|\xi\rangle} = \alpha\mathbb{I} - |\xi\rangle\langle\xi| \tag{2.27}$$

with $\alpha = \max_{|\phi_{\text{sep}}\rangle} |\langle\phi_{\text{sep}}|\xi\rangle|^2$.

Partial transpose

Another computationally simple way to detect entanglement is to investigate the eigenvalues of a partially transposed state ρ [52]. A density matrix ρ that is composed out of a sum of direct products (see Eq. 2.21) is separable. To examine the separability, it is possible to construct a partial transpose of ρ and inspect its eigenvalues. While all valid density matrices are positive operators and all eigenvalues λ_i are $\in [0,1]$ and $\sum_{i=1}^{2^n} \lambda_i = 1$, these constraints do not hold for partially transposed non-separable states. Using proper indices $\{m, n, \mu, \nu\}$ for the density matrix, Eq. 2.21 can be written as [10]

$$\rho_{m\mu,n\nu} = \sum_i p_i \left(\rho_i^A\right)_{mn} \left(\rho_i^{BCD}\right)_{\mu\nu}. \tag{2.28}$$

Since $\{m, n\}$ are indexing the first (A) and $\{\mu, \nu\}$ the second subsystem (BCD) according to the cut $A|BCD$, the partial transpose on ρ's first subsystem A results in an exchange of the Latin indices [52] m and n, such that

$$\sigma_{m\mu,n\nu} \equiv \rho_{n\mu,m\nu} \quad \Rightarrow \rho^{T_A} = \sigma. \tag{2.29}$$

If now ρ^{T_A} is again a valid density matrix, i.e. $\rho^{T_A} \geq 0$, ρ is not entangled according to the separation of subsystem A to all other subsystems. According to Sec. 2.3.1, a four qubit state may be for example a bi-product state of the form $A|BCD$. Enquiring the entanglement with respect to the cut $A|BCD$, one does not detect any entanglement, while other choices like $AB|CD$ or $B|ACD$ will prove the state to be entangled.

The partial transposition of a density matrix cannot only be used to detect entanglement, but gives also a measure for the amount that ρ^{T_A} fails to be a positive operator. This measure is called *bipartite negativity* [53] and can be used as a quantitative measure of entanglement [47].
While a two qubit state can only be splitted into $A|B$, three qubits can for example be cut into $A|BC$, $AB|C$ or $AC|B$ as described in Sec. 2.3.1. Formally, the split $A|BC$ would then result in $\rho^{T_A} = (T \otimes \mathbb{I} \otimes \mathbb{I})\rho$, where T denotes the transpose, which here only acts on the first qubit. Thus, the negativity is given by

$$N(\rho) = \frac{||\rho^{T_A}||_1 - 1}{2}, \tag{2.30}$$

where $||\rho^{T_A}||_1$ corresponds to the trace norm of ρ^{T_A}, i.e. $||\rho^{T_A}||_1 = \sum_{i=1}^{2^n} |\lambda_i|$ with ρ^{T_A}'s eigenvalues λ_i [53]. Therefore, the negativity directly corresponds to the sum of negative eigenvalues [54]. Furthermore, the logarithmic negativity $E_N(\rho) = \log_2(||\rho^{T_A}||) = \log_2(1 + 2N(\rho))$ is often considered, too [53].

2.4.2 Trace distance

It is often needed to evaluate the distance between two quantum states. One basic measure to do so is the *trace distance* D between the states ρ and σ [5]:

$$D(\rho, \sigma) \equiv \frac{1}{2} \operatorname{tr}\left(|\rho - \sigma|\right) \tag{2.31}$$

with $|\rho - \sigma| = \sqrt{(\rho - \sigma)^\dagger(\rho - \sigma)}$. Since ρ and σ are Hermitian matrices, $D(\rho, \sigma) = \frac{1}{2}\sum_{i=1}^{2^n} |\lambda_i|$ where the λ_i are the eigenvalues of $(\rho - \sigma)$.
The trace distance fulfils the properties

- $D(\rho, \sigma) \geq 0 \quad \forall \rho, \sigma$ (non-negativity)
- $D(\rho, \sigma) = 0 \iff \rho = \sigma$ (identity of indiscernibles, Leibniz' law)
- $D(\rho, \sigma) = D(\sigma, \rho) \quad \forall \rho, \sigma$ (symmetric in arguments)
- $D(\rho, \sigma) \leq D(\rho, \mu) + D(\mu, \sigma) \quad \forall \rho, \sigma, \mu$ (triangle inequality)

and is therefore a metric. The trace distance is a convex measure in its input arguments, such that

$$D\left(\sum_i p_i \rho_i, \sigma\right) \leq \sum_i p_i D\left(\rho_i, \sigma\right) \tag{2.32}$$

and analogously for the second argument [5].

2.4.3 Fidelity

While the trace distance was already introduced as a distance measure, another prominent measure is given by the *fidelity*. While in some cases the square root fidelity $\tilde{F}(\rho, \sigma) = \sqrt{F(\rho, \sigma)} = \operatorname{tr}\sqrt{\sqrt{\rho}\sigma\sqrt{\rho}}$ is used [5], the so called *Uhlmann fidelity* may be more comprehensible - at least for pure states - if it is interpreted as a probability of success of a quantum computation $|\langle\psi|\phi\rangle|^2$, where $|\psi\rangle$ and $|\phi\rangle$ represent the measured and the target state, respectively. The Uhlmann fidelity [55], in the following just called fidelity, is defined as [56]

$$F\left(\rho, \sigma\right) \equiv \left(\operatorname{tr}\sqrt{\sqrt{\sigma}\rho\sqrt{\sigma}}\right)^2. \tag{2.33}$$

If the target state is pure, i.e. $\sigma = |\psi\rangle\langle\psi|$, the fidelity simplifies to $F(\rho, |\psi\rangle) = \langle\psi|\rho|\psi\rangle$, which directly corresponds to the overlap of ρ and $|\psi\rangle$ [56]. While the interpretation of the fidelity is convincing if one of the states is pure, it is less intuitive for two mixed states [57].
The fidelity is symmetric under exchange of its arguments and is non-negative for all density operators, but since it does not fulfil the *identity of indiscernibles* and the *triangle inequality*, it is not a metric. Two possible ways to use the fidelity to build up a metric are on the one hand the *angle* between the states ρ and σ with $D_\mathrm{A}(\rho, \sigma) \equiv \arccos\sqrt{F(\rho, \sigma)}$ [5] and on the other hand construct the so called *Bures metric* $D_\mathrm{B}(\rho, \sigma) \equiv \sqrt{2 - \sqrt{F(\rho, \sigma)}}$.
Furthermore, the trace distance and the fidelity are closely related, as it holds [56]

$$1 - \sqrt{F(\rho, \sigma)} \leq D(\rho, \sigma) \leq \sqrt{1 - F(\rho, \sigma)}. \tag{2.34}$$

2.4.4 Hilbert-Schmidt distance

Besides the trace distance $D(\rho, \sigma)$, the Bures metric $D_\mathrm{B}(\rho, \sigma)$ and the angle between states $D_\mathrm{A}(\rho, \sigma)$, there are various other metrics and distance measures defined. One of those is the *Hilbert-Schmidt distance*. First, one defines an inner product (scalar product) - the so called *Hilbert-Schmidt inner product* or equivalently the *trace inner product* - to the space of density operators for two operators A and B as [58]

$$\langle A, B\rangle \equiv \frac{1}{2}\operatorname{tr}(A^\dagger B). \tag{2.35}$$

While [5] uses the inner product without the coefficient of $\frac{1}{2}$, here the definition of [58] will be followed. Defining a distance with this inner product is possible. The *Hilbert-Schmidt distance* is

$$D_{\text{HS}}(\rho,\sigma) = \sqrt{\operatorname{tr}(\rho-\sigma)^2}. \tag{2.36}$$

Using the parametrisation of ρ and σ in its correlations, Eq. 2.36 becomes

$$D_{\text{HS}}(\rho,\sigma) = \sqrt{\sum_{i,j}\left(T_{ij}^{(A)} - T_{ij}^{(B)}\right)^2}. \tag{2.37}$$

The trace distance $D(\rho,\sigma)$ and the Hilbert-Schmidt distance are related according to [59]

$$0 \leq D_{\text{HS}}(\rho,\sigma) \leq 2D(\rho,\sigma). \tag{2.38}$$

2.4.5 Purity

The purity $P(\rho) = \operatorname{tr}\left(\rho^2\right)$ of a state is a measure that is related to the degree of mixedness of a quantum state. For pure states $\rho = |\psi\rangle\langle\psi|$, the purity is $P(\rho) = \operatorname{tr}\left(\rho^2\right) = \operatorname{tr}\left(|\psi\rangle\langle\psi||\psi\rangle\langle\psi|\right) = \operatorname{tr}\left(|\psi\rangle\langle\psi|\right) = 1$, while it decreases for states with increasing admixtures of other states. The lowest possible purity of $1/2^n$ is obtained by the n qubit *maximally mixed state* ρ_{WN}, see 2.5.8.

2.4.6 Entropy

According to Shannon's definition of the entropy as a measure of the amount of information [25] with $H(X) = -\sum_i p(x_i)\log(p(x_i))$ there is the quantum mechanical counterpart

$$S(\rho) = -\operatorname{tr}(\rho\log(\rho)) = -\sum_{i=1}^{2^n}\lambda_i\log(\lambda_i) \tag{2.39}$$

where S is the *von Neumann entropy* [60] and log denotes the logarithm to base two [5][4]. The von Neumann entropy can be seen as the Shannon entropy of the spectrum of the density matrix ρ [58]. For pure states the von Neumann entropy directly vanishes $S(|\psi\rangle\langle\psi|) = 0$ while it reaches its maximum value for the maximally mixed state ρ_{WN}, i.e. $S(\rho_{\text{WN}}) = -\sum_{i=1}^{2^n}\frac{1}{2^n}\log(\frac{1}{2^n}) = \log(2^n) = n$ such that the maximally mixed state carries the largest uncertainty. The entropy is a concave measure with $S\left(\sum_i p_i\rho_i\right) \geq \sum_i p_i S(\rho_i)$ with $\sum_i p_i = 1$ [5].

---

[4] $0\log 0 \equiv 0$ is used for evaluating Eq. 2.39, which is justified by $\lim_{x\to 0} x\log x = 0$.

2.5 Prominent Quantum States

For investigating different features of quantum systems various quantum states have to be considered. In the following, an overview of important quantum states, that are used during this thesis, will given given.

2.5.1 Bell states

Four prominent two qubit states that feature maximal entanglement are the *Bell states*. While the EPR state [61] $|\phi^+\rangle$ was already mentioned before, three other Bell states [5] shall be defined:

$$\begin{aligned} |\phi^+\rangle &= \frac{|0\rangle_A \otimes |0\rangle_B + |1\rangle_A \otimes |1\rangle_B}{\sqrt{2}}, \quad (2.40)\\ |\phi^-\rangle &= \frac{|0\rangle_A \otimes |0\rangle_B - |1\rangle_A \otimes |1\rangle_B}{\sqrt{2}}, \quad (2.41)\\ |\psi^+\rangle &= \frac{|0\rangle_A \otimes |1\rangle_B + |1\rangle_A \otimes |0\rangle_B}{\sqrt{2}}, \quad (2.42)\\ |\psi^-\rangle &= \frac{|0\rangle_A \otimes |1\rangle_B - |1\rangle_A \otimes |0\rangle_B}{\sqrt{2}}. \quad (2.43) \end{aligned}$$

The singlet state $|\psi^-\rangle$ is invariant under rotations of equal $SU(2)$ matrices [62]. The four Bell states $\{|\phi^+\rangle, |\phi^-\rangle, |\psi^+\rangle, |\psi^-\rangle\}$ together form an orthonormal basis of the two qubit Hilbert space. The Bell states can be transformed into each other by unitary transformations, such that [28]

$$|\psi^-\rangle \propto (\sigma_0 \otimes \sigma_z)|\psi^+\rangle \propto (\sigma_0 \otimes \sigma_x)|\phi^-\rangle \propto (\sigma_0 \otimes \sigma_y)|\phi^+\rangle, \quad (2.44)$$

where the proportionality is due to some irrelevant global phases.

2.5.2 Greenberger-Horne-Zeilinger states

It was argued with the Bell state $|\phi^{(+)}\rangle$ in Eq. 2.17 that the measurement of $|0\rangle_A$ directly implies the measurement outcome of Bob to be $|0\rangle_B$ and analogously for $|1\rangle_A$ and $|1\rangle_B$. While in this case only two parties, Alice and Bob, are involved, the state can be extended to more qubits such that the measurement outcome on one qubit directly determines the outcome of all other qubits to be the same outcome. For n qubits, a possible *GHZ* (Greenberger-Horne-Zeilinger) state reads [63]

$$|\mathrm{GHZ}\rangle = \frac{|0\rangle^{\otimes n} + |1\rangle^{\otimes n}}{\sqrt{2}}. \quad (2.45)$$

Interestingly, partially tracing out one qubit causes the loss of the entanglement of the state. For $n = 3$ qubits, the GHZ state reads $|\mathrm{GHZ}\rangle = \frac{|000\rangle+|111\rangle}{\sqrt{2}}$. Tracing out the third qubit gives

$$\mathrm{tr}_3\left[\frac{1}{2}(|000\rangle + |111\rangle)(\langle 000| + \langle 111|)\right] = \frac{1}{2}(|00\rangle\langle 00| + |11\rangle\langle 11|) = \rho_{\mathrm{WN},2} \quad (2.46)$$

where ρ_{WN} denotes the maximally mixed state, which will be further explained in 2.5.8.

2.5.3 Dicke states

One important class of permutationally invariant states [64] are the *Dicke states* [65]. The Dicke state $|D_n^{(e)}\rangle$ is the symmetric state with e qubits in the excited state and $(n-e)$ qubits in their respective ground state. Therefore, the Dicke state is defined as [66]

$$|D_n^{(e)}\rangle = \binom{n}{e}^{-\frac{1}{2}} \sum_k \Pi_k \left(|1\rangle^{\otimes e} \otimes |0\rangle^{\otimes n-e}\right), \tag{2.47}$$

where $\Pi_k(|\psi\rangle)$ denotes the k-th permutation of the qubits of $|\psi\rangle$. Since it is summed over all permutations, the state is fully symmetric under exchange of any qubits. The prefactor $\binom{n}{e}^{-\frac{1}{2}}$ is just due to normalisation.

2.5.4 W states

The symmetric $|\mathrm{W}\rangle$ state [45] and its counterpart, the $|\overline{\mathrm{W}}\rangle$ state, are the special cases of the Dicke state with $e = 1$ excitations and $e = n-1$ excitations, respectively, i.e.

$$|\mathrm{W}\rangle = |\mathrm{D}_\mathrm{n}^{(1)}\rangle, \tag{2.48}$$
$$|\overline{\mathrm{W}}\rangle = |\mathrm{D}_\mathrm{n}^{(\mathrm{n}-1)}\rangle. \tag{2.49}$$

The entanglement of the $|\mathrm{W}\rangle$ state can for example be shown by the negativity of the partial transpose[5] or by using an entanglement witness [67].

Generalised W states

Motivated by these symmetric superpositions, a generalised form of the $|\mathrm{W}\rangle$ state can be defined[6]. $|\mathrm{W}^g\rangle$ and $|\overline{\mathrm{W}}^g\rangle$ fulfil for the case of $n = 3$ qubits

$$|\mathrm{W}^g\rangle = \lambda|001\rangle + \mu|010\rangle + \nu|100\rangle, \tag{2.50}$$
$$|\overline{\mathrm{W}}^g\rangle = \lambda|110\rangle + \mu|101\rangle + \nu|011\rangle \tag{2.51}$$

with $|\lambda|^2 + |\mu|^2 + |\nu|^2 = 1$. With this normalisation constraint the definition can be reformulated as $|\mathrm{W}_3^g\rangle = \cos(\alpha)\sin(\beta)|001\rangle + \sin(\alpha)\sin(\beta)|010\rangle + \cos(\beta)|100\rangle$ (and for $|\overline{\mathrm{W}}^g\rangle$ analogously). Therefore, the *generalised* $|\mathrm{W}^g\rangle$ state is obtained by a superposition of the same states as the *symmetric* $|\mathrm{W}\rangle$ state, but with coefficients freely chosen.

2.5.5 Graph states

Pure n qubit states that obey a vertex-and-edge structure are called *Graph states* [10]. Let $G = (V, E)$ be a graph with the vertices V and the edges E. Each vertex $a \in V$

---

[5] The eigendecomposition of each possible partial transposition of $|\mathrm{W}_3\rangle$ delivers amongst others the eigenvalue $-\frac{\sqrt{2}}{3}$ and is therefore according to Sec. 2.4.1 entangled.

[6] Please note that these generalised W states are, strictly speaking, no W states.

corresponds to a qubit of the state, while each edge represents an interaction between two qubits. When all vertices b that are connected to the vertex a via an edge are grouped by the set of neighbouring vertices N_a, to each vertex a Hermitian operator

$$K_G^{(a)} = \sigma_x^{(a)} \prod_{b \in N_a} \sigma_z^{(b)} \tag{2.52}$$

may be assigned, where $\sigma_i^{(j)}$ denotes the ith Pauli matrix acting on the jth qubit [68]. A state is a Graph state $|G\rangle$ if it is an eigenstate to all $K_G^{(a)}$ $\forall a \in V$ operators with the eigenvalue 1, i.e.

$$K_G^{(a)}|G\rangle = |G\rangle. \tag{2.53}$$

Because of the property to correspond to a specific graph $G = (V, E)$, the Graph states can easily be drawn schematically. Different graphs are shown in Fig. 2.4. The graph "B" in Fig. 2.4 corresponds to a three qubit Greenberger-Horne-Zeilinger state [10]. The graph "C" depicts a four qubit state where all qubits have interaction with one single qubit.This state corresponds (up to local unitary transformation) also to the already introduced $|\text{GHZ}\rangle$ state [69]. While there is only one possibility to create a graph for two ("A") and three qubits ("B"), respectively, a four qubit graph can be built in different ways. Besides the $|\text{GHZ}\rangle$ state in "C", "E" depicts also a possible connection of four qubits. This graph belongs to the four qubit *Cluster state* [70].

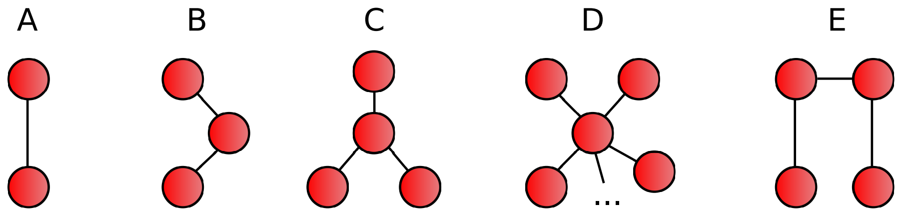

FIGURE 2.4: Examples for graph states. Each red dot represents a vertex (qubit), while the lines show the edges (interactions) in the graph formalism. "A" represents the Bell state $|\phi^+\rangle$. "B" and "C" correspond to the states $|\text{GHZ}_3\rangle$ and $|\text{GHZ}_4\rangle$, respectively [69]. The general n qubit Greenberger-Horne-Zeilinger state $|\text{GHZ}_n\rangle$ can be represented by the graph "D". The graph "E", depicting a four qubit state with linear next-neighbour interactions, belongs to the Cluster state, see Sec. 2.5.6.

2.5.6 Cluster states

Cluster states [70] are special cases of the Graph states where the vertices are located in a rectangular lattice structure [68]. The cluster state is of special interest for one-way quantum computing [71], where it has been shown to be a resource for [72]. For the case of one-dimensional ordered n qubits, i.e. the qubits build up a chain with next-neighbour interactions, the cluster states for 2, 3 and 4 are defined as

$$|\psi_2\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle), \tag{2.54}$$

$$|\psi_3\rangle = \frac{1}{\sqrt{2}}(|000\rangle + |111\rangle), \tag{2.55}$$

$$|\psi_4\rangle = \frac{1}{2}(|0000\rangle + |0011\rangle + |1100\rangle - |1111\rangle). \tag{2.56}$$

The states $|\psi_2\rangle$ and $|\psi_3\rangle$ equal the $|\phi^+\rangle$ and the $|\mathrm{GHZ}_3\rangle$ states, respectively, and are therefore depicted in Fig. 2.4 by the graphs "A" and "B". $|\psi_4\rangle$ is represented by the graph "E" in Fig. 2.4. In contrast to the four qubit $|\mathrm{GHZ}\rangle$ state, where losing one qubit destroys the entanglement (Eq. 2.46), the cluster state $|\psi_4\rangle$ is more persistent against losing photons [70]. While the entanglement still lingers on after one photon is lost, losing two photons can ruin the entanglement [28].

2.5.7 Smolin states

Not only pure states like the Bell states, the Greenberger-Horne-Zeilinger states $|\mathrm{GHZ}\rangle$ or the Dicke states $|D_n^{(e)}\rangle$ exhibit entanglement, but instead this may also hold for mixed states. One interesting mixed state is the *Smolin* [73] state, because it can be for example used as a resource for multiparty communication schemes [74]. The generalised $n = 2k$ qubit form of the Smolin state is [75]

$$\rho_{\mathrm{Smolin}}^{(n)} = \frac{1}{2^n}\left(\mathbb{I}_{n\times n} + (-1)^k \sum_{i=1}^{3} \sigma_i^{\otimes n}\right), \tag{2.57}$$

i.e. the elements of the correlation tensor are easily determined with $T_{0\ldots0} = 1$, $T_{1\ldots1} = T_{2\ldots2} = T_{3\ldots3} = (-1)^k$. The purity of the Smolin state decreases with the number of qubits, that is $\mathrm{tr}((\rho_{\mathrm{Smolin}}^{(n)})^2) = \frac{1}{2^{n-2}}$. Corresponding to the decreasing purity, the rank of the Smolin state increases with the number of qubits $n = 2k$ according to $\mathrm{rank}(\rho_{\mathrm{Smolin}}^{(n)}) = 2^{(n-2)}$.

2.5.8 Maximally mixed states

Another important state that shall be discussed here is the n qubit *maximally mixed state* or *white noise*. In contrast to the earlier mentioned pure states, this state is of full rank: $\mathrm{rank}(\rho_{\mathrm{WN}}) = 2^n$, where ρ_{WN} denotes the white noise

$$\rho_{\mathrm{WN}} = \frac{1}{2^n}\mathbb{I}. \tag{2.58}$$

The purity of the maximally mixed states is the lowest possible purity value of all states with

$$\mathrm{tr}\left(\rho_{\mathrm{WN}}^2\right) = \mathrm{tr}\left(\frac{1}{2^n}\mathbb{I}\frac{1}{2^n}\mathbb{I}\right) = \frac{2^n}{4^n} = \frac{1}{2^n}. \tag{2.59}$$

The eigendecomposition of the maximally mixed state gives 2^n degenerate eigenvalues of $1/2^n$. The corresponding eigenstates are any set of 2^n mutually orthonormal states forming a basis.

2.5.9 Random quantum state according to Haar measure

It can be helpful to simulate random states for numerical simulations. Therefore, a short recipe will be given how equally distributed random states can be obtained. A possible parametrisation for a single qubit was already given in Eq. 2.2 by $|\psi\rangle = \cos\left(\frac{\theta}{2}\right)|0\rangle + \sin\left(\frac{\theta}{2}\right)e^{i\phi}|0\rangle$. It is obvious that choosing the parameters θ and ϕ uniformly

$[0, \pi]$ and $[0, 2\pi]$, respectively, leads to a non-uniform distribution of states. In this case, a randomly chosen state would lie with higher probability on one of the poles while the density of an ensemble of these states would be lowest on the equator $\theta = \frac{\pi}{2}$. To overcome this skewed distribution, θ may not be chosen uniformly, but instead according to the inverse sine function $\sin^{-1}$. Thus, a u is uniformly chosen with $u \in [-1, 1]$, while $\theta = \sin^{-1}(u) + \frac{\pi}{2}$. Using this recipe, the random single qubit states are now uniformly distributed around the Bloch sphere.

This technique can be generalised to find uniformly distributed n qubit states in their state space. A good way to do so is to start with any n qubit state, e.g. $|0\rangle^{\otimes n}$. Afterwards, this state shall undergo an unitary transformation to deliver all pure n qubit states with the same probability. A complete recipe to create random unitary matrices whose distribution is in agreement with the *Haar measure* is according to [58, 76]

- Create a complex random $2^n \times 2^n$ matrix Z with normally distributed matrix elements.
- Find a QR decomposition of the random matrix with $Z = QR$.
- Build a diagonal matrix Λ with $\Lambda = \mathrm{diag}\left(\frac{r_{1,1}}{|r_{1,1}|}, \dots, \frac{r_{2^n,2^n}}{|r_{2^n,2^n}|}\right)$, where $(r_{i,i})_i$ denote the diagonal entries of the R matrix.
- Finally, $U = Q\Lambda$ is an unitary matrix whose density of eigenvalues is uniformly distributed.

2.6 Optical components

For experimental preparation and application of entangled quantum states, it is crucial to have a versatile toolbox. Since the experiments that this thesis is founded onto are dealing with photonic systems, this toolbox is composed of optical devices that allow to modify the polarisation of light. First of all, one requires a device to create entanglement, i.e. a pair or a group of entangled photons. In the following section, an overview of the most important utilities for photonic manipulation will be given.

2.6.1 Spontaneous parametric down conversion

Entangled states can be realised with different systems. One of the most important methods - and the one used throughout this thesis - is based on employment of non-linear crystals. Isotropic and linear materials are polarised parallel and proportional to the incident electric field, i.e. [77]

$$\vec{P} = \epsilon_0 \chi \vec{E}, \tag{2.60}$$

where ϵ_0 denotes the vacuum permittivity and χ the electric susceptibility. Since Eq. 2.60 is only a linear approximation, while the polarisation indeed does not increase linearly for all electrical fields $\vec{E}$, the polarisation is better approximated by a power series in $\vec{E}$ [78]

$$P = \epsilon_0 \left(\chi E + \chi^{(2)} E^2 + \chi^{(3)} E^3 + \dots\right) \tag{2.61}$$

now for scalar P and E and with $\chi^{(i)} \gg \chi^{(i+1)}$. To drive the higher orders, appropriate pump powers are needed, since the susceptibilities $\chi^{(i>1)}$ are diminutive compared to

χ. Consequently, only for high intensity of the electric field their corresponding term becomes influential. Since isotropic substances have vanishing susceptibilities for even powers in E due to the symmetry, that is $\chi^{(2i)} = 0$, it is crucial to study the anisotropic case. For materials without inversion-centres the susceptibilities are not scalar, but tensorial [77]. Therefore, the polarisation P_i along a direction i can be written as [79]

$$P_i(\vec{E}) = \epsilon_0 \left(\sum_{j\in\{x,y,z\}} \chi_{i,j} E_j + \sum_{j,k\in\{x,y,z\}} \chi^{(2)}_{i,j,k} E_j E_k + \sum_{j,k,l\in\{x,y,z\}} \chi^{(3)}_{i,j,k,l} E_j E_k E_l + \dots \right). \tag{2.62}$$

The used crystal for *spontaneous parametric down conversion* (SPDC) allows to convert *pump* photons into *signal* and *idler* photons in the three-wave mixing process [79], which $\chi^{(2)}$ has to be non-vanishing for. Since energy and momentum have to be conserved during this process, the frequencies ω_{Pump}, ω_{Signal} and ω_{Idler} as well as the respective wavevectors $\vec{k}_{Pump}$, $\vec{k}_{Signal}$ and $\vec{k}_{Idler}$ are subject to the conditions [78]

$$\hbar\omega_{Pump} = \hbar\omega_{Signal} + \hbar\omega_{Idler} \tag{2.63}$$

and

$$\hbar\vec{k}_{Pump} = \hbar\vec{k}_{Signal} + \hbar\vec{k}_{Idler}. \tag{2.64}$$

While Eq. 2.63 and Eq. 2.64 were derived in the photonic picture as energy and momentum conservation, respectively, in the picture of incident and emitted waves, the conditions $\omega_{Pump} = \omega_{Signal} + \omega_{Idler}$ (frequency-matching condition) and $\vec{k}_{Pump} = \vec{k}_{Signal} + \vec{k}_{Idler}$ (phase-matching condition) have to fulfilled in the exact same manner [78]. The crystal used for the experimental preparations of this thesis is made of *BBO* (beta barium borate). Depending on the orientation of the optical axis of the used crystal, the phase-matching conditions are fulfilled by different constellations of frequencies and wave-vectors. Therefore, different types of SPDC have to be distinguished, which will be presented here in brief.

Type I

One possible solution for the phase-matching condition is matched if the incoming pump photons' polarisation is along the so called optical axis of the crystal [28]. These *extraordinary* polarised photons are affected by a refraction index n_e. The emitted photons (signal and idler) are then polarised perpendicular to the crystal's axis. For this *ordinary* polarisation of the signal and the idler photons, the index of refraction is n_o with $n_e < n_o$ for the case of negative, uni-axial crystals [79]. The incoming pump photons can produce signal and idler photons, which are distributed along the surface of a cone. The corresponding Hamiltonian reads as

$$\hat{H}_{type\ I} \propto \hat{a}^\dagger_H \hat{b}^\dagger_H + H.c., \tag{2.65}$$

where $\hat{a}^\dagger_H$ is the creation operator of a photon in output mode a with the horizontal polarisation. $H.c.$ denotes the Hermitian conjugate. Depending on the orientation of the crystal, the output can also be a pair of vertically polarised photons. Since signal and idler photon are polarised in the same way, two crossed crystals have to be used with indistinguishable cones in order to create entangled states. One crystal produces for example $|HH\rangle$ photons, the other one $|VV\rangle$. Their overlapping emission corresponds to the entangled state $\propto |HH\rangle + e^{i\varphi}|VV\rangle$.

Type II

For type II phase-matching, the pump beam is polarised parallel to the optical axis of the crystal. One of the signal and idler photons is polarised parallel, as well, while the other one's polarisation is perpendicular to the crystals optical axis [28]. The Hamiltonian of type II sources for the non-collinear case along two modes where the two cones of signal and idler emission intersect is found to be [80]

$$\hat{H}^{(ncol)}_{type\ II} \propto \left(\hat{a}^{\dagger}_{H}\hat{b}^{\dagger}_{V} - \hat{a}^{\dagger}_{V}\hat{b}^{\dagger}_{H}\right) + H.c.. \tag{2.66}$$

The proportionality hides the coupling between crystal and pump field and depends on $\chi^{(2)}$ [79]. Acting on the vacuum, in first order a state like $\frac{1}{\sqrt{2}}(|HV\rangle - |VH\rangle)$ is created while the full expansion reads as $|\psi\rangle \propto \exp(-i\alpha(\hat{a}^{\dagger}_{H}\hat{b}^{\dagger}_{V} - \hat{a}^{\dagger}_{V}\hat{b}^{\dagger}_{H}))|0\rangle$ [81]. In the collinear case, the entangled photons cannot be distinguished by spatial modes a and b. Therefore, the Hamiltonian for the collinear case becomes [79]

$$\hat{H}^{(col)}_{type\ II} \propto \hat{a}^{\dagger}_{H}\hat{a}^{\dagger}_{V} + H.c.. \tag{2.67}$$

In first order, the collinear SPDC type II source produces a pair with one horizontally and one vertically polarised photon. The full term again reads $|\psi\rangle \propto \exp(-i\alpha\hat{H}^{(col)}_{type\ II})$ [28]. Therefore, the emission of ith order corresponds to i horizontally and i vertically polarised photons while the probability for emission in ith order is $\frac{2^{i}\alpha^{2i}}{2\alpha^{2}}$ normalised to the probability of emission in first order [28]. For the full derivation and higher order terms of the states, see [79].

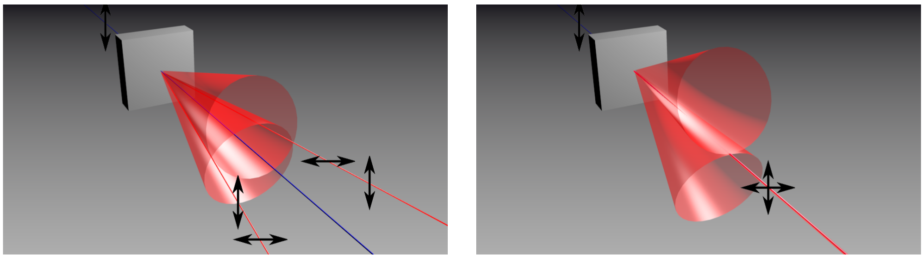

FIGURE 2.5: SPDC Type II sources. The left figure shows the general non-collinear case, while the right one depicts the special case of a collinear source. The red beams illustrate the entangled photons, that are used for the experiments. In the collinear case, the red beam is degenerate and overlaps with the pump beam. The incoming pump beam is partially transmitted through the non-linear crystal. Probabilistically, signal and idler photons are produced. By collecting the photons and the spots of intersection of the cones one obtains an entangled state.

2.6.2 Wave plate

For the experimental realisation of state preparation and analysis, it is crucial to be also able to rotate the polarisation of photons. The method of choice is to use *wave plates* [77]. The general property of these retarding elements is that the birefringent material causes different propagation of the so called ordinary and the extraordinary beams. Different refraction indices for the corresponding polarisations cause a phase shift between those two beams. The phase difference between the ordinary beam, which is polarised perpendicular to the optical axis of the birefringent crystal, and the extraordinary beam, whose polarisation is parallel to the optical axis, after the propagation through the wave plate is given by

$$\Delta\varphi = \frac{2\pi}{\lambda}(n_{\text{slow}} - n_{\text{fast}}), \tag{2.68}$$

where λ denotes the wavelength of the incident light, d the thickness of the wave plate and n corresponds to the refraction index along the slow and the fast axis, respectively. Note that in general the indices of refraction can also be wavelength dependent, i.e. $n_{\text{slow}} \equiv n_{\text{slow}}(\lambda)$ and $n_{\text{fast}} \equiv n_{\text{fast}}(\lambda)$ [78]. The two special cases with the most importance are the half-wave plate (HWP or $\lambda/2$ plate) and the quarter-wave plate (QWP or $\lambda/4$ plate). While the half-wave plate is able to rotate any linear polarisation to any other linear polarisation, the quarter-wave plates can change the polarisation from linear polarisation directions to circular and vice versa. For suitably chosen thickness d it is possible to achieve a phaseshift of $\varphi = \frac{\pi}{2}$ in the case of the quarter-wave plate and a phaseshift of $\varphi = \pi$ for the half-wave plate. The transformation matrices of the HWP and the QWP are given by [79]

$$T_{\text{HWP}}(\theta) = \begin{pmatrix} \cos(2\theta) & \sin(2\theta) \\ \sin(2\theta) & -\cos(2\theta) \end{pmatrix} = \sin(2\theta)\sigma_x + \cos(2\theta)\sigma_z \tag{2.69}$$

for the half-wave plate and

$$\begin{aligned} T_{\text{QWP}}(\theta) = &\begin{pmatrix} \cos^2(\theta) - i\sin^2(\theta) & (1+i)\sin(\theta)\cos(\theta) \\ (1+i)\sin(\theta)\cos(\theta) & -i\cos^2(\theta) + \sin^2(\theta) \end{pmatrix} = \\ &\frac{1}{2}((1-i)\sigma_0 + 2(1+i)\cos(\theta)\sin(\theta)\sigma_x + (1+i)\cos(2\theta)\sigma_z). \end{aligned} \tag{2.70}$$

for the quarter-wave plate.

2.6.3 Yttrium-vanadate crystals

Although it is possible by a combination of HWP and QWP plates to transform to all desired polarisations, it is often useful to modify solely the phase between the horizontal and vertical polarisation, i.e. to set the phase γ within $\cos(\frac{\theta}{2})|0\rangle + \sin(\frac{\theta}{2})e^{i\gamma}|0\rangle$. This can be done with birefringent Yttrium-vanadate crystals YVO_4 [28] which are rotated around their optical axes such that the phase of the perpendicular orientated horizontal polarisation is shifted relative to the vertical polarisation. The corresponding operation can be written as [79]

$$T_{\text{YVO}}(\phi) = \begin{pmatrix} i\cos\left(\frac{\phi}{2}\right) + \sin\left(\frac{\phi}{2}\right) & 0 \\ 0 & i\cos\left(\frac{\phi}{2}\right) - \sin\left(\frac{\phi}{2}\right) \end{pmatrix} = i\cos\left(\frac{\phi}{2}\right)\sigma_0 + \sin\left(\frac{\phi}{2}\right)\sigma_z. \tag{2.71}$$

2.6.4 Beam splitters

It is furthermore often needed to either overlap different beams or to split up a single beam into two. A practical way to do so is given by *beam splitters*. According to the

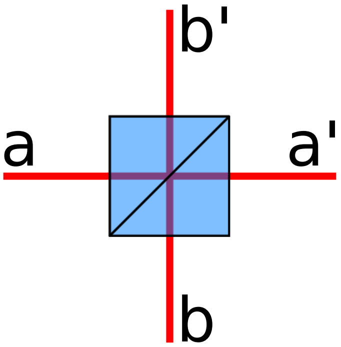

Figure 2.6: A schematic picture of a beam splitter. The input modes a and b are transformed into the output modes a' and b', depending on the transmittivity and the reflectivity.

naming in Fig. 2.6 a beam splitter has two input modes a and b, as well as two output modes a' and b'. Light that is radiated onto mode a can be transmitted to a' or reflected and pass by b'. According to [28] the transformations of an ideal beam splitter read as

$$\hat{a}_H^\dagger \rightarrow \frac{1}{\sqrt{2}}\left(T_H\hat{a}_H'^\dagger + iR_H\hat{b}_H'^\dagger\right), \tag{2.72}$$

$$\hat{a}_V^\dagger \rightarrow \frac{1}{\sqrt{2}}\left(T_V\hat{a}_V'^\dagger + iR_V\hat{b}_V'^\dagger\right), \tag{2.73}$$

$$\hat{b}_H^\dagger \rightarrow \frac{1}{\sqrt{2}}\left(T_H\hat{b}_H'^\dagger + iR_H\hat{a}_H'^\dagger\right), \tag{2.74}$$

$$\hat{b}_V^\dagger \rightarrow \frac{1}{\sqrt{2}}\left(T_V\hat{b}_V'^\dagger + iR_V\hat{a}_V'^\dagger\right), \tag{2.75}$$

where T_H, T_V, R_H and R_V denote the transmittivity and reflectivity for horizontally and vertically polarised light, respectively. One important special case is the $50:50$ beam splitter, that transmits and reflects light with the same probability, i.e. ideally $|T_H|^2 = |T_V|^2 = |R_H|^2 = |R_V|^2 = \frac{1}{2}$ [28]. Of high practical importance is also the *polarising beam splitter* (PBS). While horizontally polarised light is fully transmitted through the PBS, vertically polarised photons are reflected. Therefore, $|T_H|^2 = |R_V|^2 = 1$ and $|T_V|^2 = |R_H|^2 = 0$. PBS are in general used for the polarisation analysis, see 2.6.5.

2.6.5 Polarisation Analysis

After preparation of quantum states, their polarisation degree of freedom has to be investigated. The usual way to enquire the photons' polarisation is depicted in Fig. 2.7. Incoming photons pass through a half-wave plate and a quarter-wave plate before propagating through a polarising beam splitter. Since the PBS is only able to distinguish between two orthogonal polarisations, usually chosen as horizontal and vertical polarisation, one uses the wave plates to be also able to measure in σ_x and in σ_y basis. If the measurement protocol for example requires a measurement in σ_x basis, the HWP is rotated to $\theta = \pi/8$ since $T_{\text{HWP}}(\pi/8) = \frac{1}{\sqrt{2}}(\sigma_x + \sigma_z)$. Therefore $T_{\text{HWP}}(\pi/8)|P\rangle = |H\rangle$ and $T_{\text{HWP}}(\pi/8)|M\rangle = |V\rangle$ and thus, a measurement in the σ_x basis can be performed.

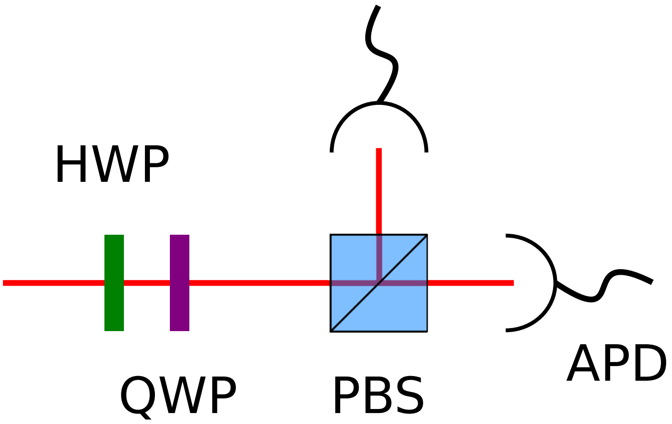

FIGURE 2.7: The polarisation analysis is done with half- and quarter-wave plates, which determine the measurement basis. A polarising beam splitter finally splits up the incident light according to its polarisation. Photonic events are noticed with avalanche photo diodes.

The HWP and QWP allow measurements in any direction, although the standard measurement bases are the σ_x, σ_y and σ_z bases.

2.7 State tomography

After all, the step of assembling a density matrix by means of measurement data will be illuminated here. Since the entries of the correlation tensor are given by $T_{\mu_1\ldots\mu_n} = \mathrm{tr}[\rho(\bigotimes_{i=1}^n \sigma_{\mu_i})]$ (see Eq. 2.19), it is essentially sufficient to measure in the $\bigotimes_{i=1}^n \sigma_{\mu_i}$ basis to determine $T_{\mu_1\ldots\mu_n}$. Naively, one would assume that measurements in 4^n bases would be required since $\mu_i \in \{0,1,2,3\}$ for all n indices μ_i. In contrast, measurements of only σ_x, σ_y, σ_z are adequate to describe the state. The leftover elements of the correlation tensor containing $\mu_i = 0$ are not needed to be determined experimentally, but can be obtained from the 3^n directions. Therefore, one may now measure $\{\sigma_x \otimes \cdots \otimes \sigma_x \otimes \sigma_x, \sigma_x \otimes \cdots \otimes \sigma_x \otimes \sigma_y, \sigma_x \otimes \cdots \otimes \sigma_x \otimes \sigma_z, \ldots, \sigma_z \otimes \cdots \otimes \sigma_z \otimes \sigma_z\}$. According to Sec. 2.6.5, it is possible to detect both the events of $|\uparrow\rangle_i$ as well as the $|\downarrow\rangle_i$ at the same time, such that 2^n different outcomes are recorded simultaneously. These detection events will be called c_i^S, where S denotes the corresponding measurement setting and i reports to which outcome this number of events belongs, i.e. $c_{00\ldots0}^{XX\ldots X}$ indicates the number of detected events by projection onto $|PP\ldots P\rangle$. For the state reconstruction, one is interested in the relative number of counts instead of the absolute number c_i^S. Thus,

$$f_i^S = \frac{c_i^S}{\sum_j c_j^S} = \frac{c_i^S}{N_S}, \tag{2.76}$$

where j labels all possible outcomes in basis setting S. N_S denotes the total number of events per basis setting S.

Since one detects the relative number of counts of, for example, the projection onto $|PP\ldots P\rangle$ and onto $|MP\ldots P\rangle$, it is easy to obtain out of this $\sigma_x \otimes \sigma_x \otimes \cdots \otimes \sigma_x$ measurement also the $\sigma_0 \otimes \sigma_x \otimes \cdots \otimes \sigma_x$ measurement. Because this holds for all bases, the correlations containing one or more σ_0 measurements are determined implicitly by a multitude of other measurements. Thus, it is self-evident to use all contributing measurements to reduce the statistical error of these non-full correlations. Finally, the reconstructed density matrix for n qubits can be written as [28]

$$\rho = \frac{1}{2^n} \sum_{\mu_1,\dots,\mu_n=1}^{3} \sum_{i_1,\dots,i_n=0}^{1} f_i^{S(\mu_1,\dots,\mu_n)} \bigotimes_{j=1}^{n} \left(\frac{1}{3}\sigma_0 + (-1)^{i_j} \sigma_{\mu_j} \right) \tag{2.77}$$

with $S = S(\mu_1, \dots, \mu_n)$ denoting the measurement basis. A set of possible measurement operators to determine the relative number of counts per basis setting are given by

$$M_i^S \equiv M_i^{S(\mu_1,\dots,\mu_n)} = \bigotimes_{j=1}^{n} \frac{1}{2} \left(\sigma_0 + (-1)^{i_j} \sigma_{\mu_j} \right) . \tag{2.78}$$

With these, the frequencies f_i^S can be determined according to

$$f_i^S = \mathrm{tr}\left(\rho M_i^S \right) . \tag{2.79}$$

For infinite number of repetitions, the relative numbers of counts f_i^S as given in Eq. 2.76 converge to these probabilities $P_\rho(i|S) = \mathrm{tr}(\rho^{\mathrm{theo}} M_i^S)$ of a theoretical state. A detailed derivation of an equivalent description to Eq. 2.77 can be found in [28].
As given in Eq. 2.19, the density matrix can be parametrised with its correlations. Each correlation can be written in terms of the number of counts, such that

$$T_{\mu_1\dots\mu_n} = \frac{\sum_{i_1,\dots,i_n=0}^{1} g(\mu_1,\dots,\mu_n; i_1,\dots,i_n) c_{i_1\dots i_n}^{\mu_1\dots\mu_n}}{\sum_{i_1,\dots,i_n=0}^{1} c_{i_1\dots i_n}^{\mu_1\dots\mu_n}} , \tag{2.80}$$

where the function $g(\mu_1, \dots, \mu_n; i_1, \dots, i_n)$ determines the parity of the corresponding outcome, see also A.2.1 for a definition of $g(\mu_1, \dots, \mu_n; i_1, \dots, i_n)$. While Eq. 2.80 seems rather bulky, the expression becomes more comprehensible by the $n = 2$ qubit example of the correlation T_{11}, which is given by

$$T_{11} = \frac{c_{00}^{11} - c_{01}^{11} - c_{10}^{11} + c_{11}^{11}}{c_{00}^{11} + c_{01}^{11} + c_{10}^{11} + c_{11}^{11}} . \tag{2.81}$$

The different signatures in the numerator are due to the parity function g. The parity of c_{00}^{11} and of c_{11}^{11} is even while c_{01}^{11} and c_{10}^{11} have odd parity.
According to Eq. 2.77 the non-full correlations like $\sigma_0 \otimes \sigma_x$ can be computed with reduced statistical variance. Possible ways to compute T_{01} out of the measured data are $T_{01} = \frac{c_{00}^{11} - c_{01}^{11} - c_{10}^{11} + c_{11}^{11}}{c_{00}^{11} + c_{01}^{11} + c_{10}^{11} + c_{11}^{11}}$, $T_{01} = \frac{c_{00}^{21} - c_{01}^{21} - c_{10}^{21} + c_{11}^{21}}{c_{00}^{21} + c_{01}^{21} + c_{10}^{21} + c_{11}^{21}}$ as well as $T_{01} = \frac{c_{00}^{31} - c_{01}^{31} - c_{10}^{31} + c_{11}^{31}}{c_{00}^{31} + c_{01}^{31} + c_{10}^{31} + c_{11}^{31}}$. The error of T_{01} can now be reduced by using the average over all three possible ways to compute the correlation. With Gaussian error propagating the error of $T_{01} = \frac{1}{3}(T_{11} + T_{21} + T_{31})$ is reduced:

$$\begin{aligned} \Delta T_{01} &= \sqrt{\left(\frac{\partial T_{01}}{\partial T_{11}} \right)^2 (\Delta T_{11})^2 + \left(\frac{\partial T_{01}}{\partial T_{21}} \right)^2 (\Delta T_{21})^2 + \left(\frac{\partial T_{01}}{\partial T_{31}} \right)^2 (\Delta T_{31})^2} \\ &= \frac{1}{3} \sqrt{(\Delta T_{11})^2 + (\Delta T_{21})^2 + (\Delta T_{31})^2} . \end{aligned} \tag{2.82}$$

Therefore, full correlations come with largest statistical deviations since they can only be computed in a single way while non-full correlation tensor elements can be decomposed into more contributing measured outcomes which reduces the statistical scatter.

Chapter 3

No Correlations

3.1 Introduction

In Sec. 2.3, the way how entanglement was introduced suggested a close relationship between the entanglement of a state and the correlations of the measurement outcomes. Naively, one would expect genuinely n-partite entangled states to exhibit n-particle correlations. Accordingly, a tripartite state would be expected to be separable if no correlations between the measurement outcomes of the three parties can be observed. Rather, in this chapter *genuinely tripartite entangled* states will be studied, where *no full correlations* can be observed. This at first glance unexpected and strange behaviour will be discussed in detail. Firstly, a particular state with the aforementioned properties will be theoretically presented. Secondly, this phenomenon will be generalised to a class of states. A recipe will be given how one can obtain a mixture of two states with odd number of qubits that show no n-partite correlations. And thirdly, experimental preparation of an exemplary state and its analysis will be discussed. This state will be proven to be entangled while having vanishing full correlations.

3.2 State with vanishing full correlations

According to Sec. 2.5.4 the (symmetric) $|\mathrm{W}_3\rangle$ state is defined as

$$|\mathrm{W}_3\rangle = |\mathrm{D}_3^{(1)}\rangle = \frac{1}{\sqrt{3}}(|001\rangle + |010\rangle + |100\rangle). \tag{3.1}$$

The $|\mathrm{W}_3\rangle$ state is entangled (see Sec. 2.3.1) and shows correlations between all qubits. Consider for example the case that the measurements on the first and the second qubit give the results $|0\rangle_A$ and $|0\rangle_B$, respectively. Then, by knowing these measurement outcomes, the measurement result on the third qubit can easily be determined to be $|1\rangle_C$. In this case it is trivially clear that not even the knowledge of the exact measurement results is needed, but only their correlation to determine the third outcome. Knowing the first two qubits in the same state, which in this case is equivalent to knowing both in state $|0\rangle$, lets one directly infer that the third qubit is in state $|1\rangle$. If the third party

Correlation	Measurement	$\vert\mathrm{W}_3\rangle$	$\vert\overline{\mathrm{W}}_3\rangle$
T_{000}	$\sigma_0 \otimes \sigma_0 \otimes \sigma_0$	1	1
T_{003}	$\sigma_0 \otimes \sigma_0 \otimes \sigma_z$	$1/3$	$-1/3$
T_{011}	$\sigma_0 \otimes \sigma_x \otimes \sigma_x$	$2/3$	$2/3$
T_{022}	$\sigma_0 \otimes \sigma_y \otimes \sigma_y$	$2/3$	$2/3$
T_{030}	$\sigma_0 \otimes \sigma_z \otimes \sigma_0$	$1/3$	$-1/3$
T_{033}	$\sigma_0 \otimes \sigma_z \otimes \sigma_z$	$-1/3$	$-1/3$
T_{101}	$\sigma_x \otimes \sigma_0 \otimes \sigma_x$	$2/3$	$2/3$
T_{110}	$\sigma_x \otimes \sigma_x \otimes \sigma_0$	$2/3$	$2/3$
T_{113}	$\sigma_x \otimes \sigma_x \otimes \sigma_z$	$2/3$	$-2/3$
T_{131}	$\sigma_x \otimes \sigma_z \otimes \sigma_x$	$2/3$	$-2/3$
T_{202}	$\sigma_y \otimes \sigma_0 \otimes \sigma_y$	$2/3$	$2/3$
T_{220}	$\sigma_y \otimes \sigma_y \otimes \sigma_0$	$2/3$	$2/3$
T_{223}	$\sigma_y \otimes \sigma_y \otimes \sigma_z$	$2/3$	$-2/3$
T_{232}	$\sigma_y \otimes \sigma_z \otimes \sigma_y$	$2/3$	$-2/3$
T_{300}	$\sigma_z \otimes \sigma_0 \otimes \sigma_0$	$1/3$	$-1/3$
T_{303}	$\sigma_z \otimes \sigma_0 \otimes \sigma_z$	$-1/3$	$-1/3$
T_{311}	$\sigma_z \otimes \sigma_x \otimes \sigma_x$	$2/3$	$-2/3$
T_{322}	$\sigma_z \otimes \sigma_y \otimes \sigma_y$	$2/3$	$-2/3$
T_{330}	$\sigma_z \otimes \sigma_z \otimes \sigma_0$	$-1/3$	$-1/3$
T_{333}	$\sigma_z \otimes \sigma_z \otimes \sigma_z$	-1	1

TABLE 3.1: Correlation values of the $|\mathrm{W}_3\rangle$ and $|\overline{\mathrm{W}}_3\rangle$ states. All full correlations are opposite.

on the other hand knows that the first two outcomes were anticorrelated, i.e., one measured $|0\rangle$ and the other measurement resulted in $|1\rangle$, the third outcome is surely $|0\rangle$. So, please, note that already the knowledge of the *correlation* of two qubits is sufficient for the $|\mathrm{W}_3\rangle$ state to infer the third measurement outcome. One does not need the information of the exact outcomes. Of course, this holds equivalently for the $|\overline{\mathrm{W}}_3\rangle$ state where the occurrences of the ground and the excited states are exchanged such that $|0\rangle \leftrightarrow |1\rangle$. In this case, knowing two qubits being correlated allows one to infer the third measurement outcome $|0\rangle$ while anticorrelation points to $|1\rangle$ for the third qubit. Tab. 3.1 lists all possible, non-vanishing correlations, full and non-full[1], for the states $|\mathrm{W}_3\rangle$ and $|\overline{\mathrm{W}}_3\rangle$.

Note that the $|\mathrm{W}_3\rangle$ and the $|\overline{\mathrm{W}}_3\rangle$ state show opposite full correlations, as listed in Tab. 3.1. Consequently, this poses the question what the state might look like if these two states are *combined* in that way that their full correlations add up to zero. Is the resulting state still entangled? Can the state be prepared such that no full correlations remain? The full correlations of $|\mathrm{W}_3\rangle$ and $|\overline{\mathrm{W}}_3\rangle$ shall vanish by just adding them. According to Eq. 2.19, the two states have to be incoherently summed, not superposed. Therefore, the state of interest is

$$\rho_{\mathrm{NC},3} = 0.5(\rho_{\mathrm{W}_3} + \rho_{\overline{\mathrm{W}}_3}) = 0.5\left(|\mathrm{W}_3\rangle\langle\mathrm{W}_3| + |\overline{\mathrm{W}}_3\rangle\langle\overline{\mathrm{W}}_3|\right). \tag{3.2}$$

---

[1] The discrimination between *full* and *non-full* correlations is of high importance here. The term "full correlation" (n-partite correlation) references to a correlation where on each qubit a measurement in σ_x, σ_y or σ_z basis is performed. In contrast, for example $T_{110} \equiv T_{XX0}$ labels the "non-full correlation" of the measurement of $\sigma_x \otimes \sigma_x \otimes \sigma_0$. Evidently, a state without any correlations is the maximally mixed state (Sec. 2.5.8).

This state will be called *no correlation* state and will be labelled with $\rho_{\mathrm{NC},3}$. Writing $\rho_{\mathrm{NC},3}$ in terms of its correlations, it reads as

$$\rho_{\mathrm{NC},3} = \frac{1}{2^3}\left(\sigma_0^{\otimes 3} + \frac{2}{3}\Pi\left(\sigma_0 \otimes \sigma_x \otimes \sigma_x\right) + \frac{2}{3}\Pi\left(\sigma_0 \otimes \sigma_y \otimes \sigma_y\right) - \frac{1}{3}\Pi\left(\sigma_0 \otimes \sigma_z \otimes \sigma_z\right)\right), \tag{3.3}$$

where Π corresponds to summing over all permutations of the respective Pauli matrices, i.e. $\Pi(\sigma_0 \otimes \sigma_x \otimes \sigma_x) = \sigma_0 \otimes \sigma_x \otimes \sigma_x + \sigma_x \otimes \sigma_0 \otimes \sigma_x + \sigma_x \otimes \sigma_x \otimes \sigma_0$. One directly sees that there is no full correlation and all remaining correlations of $\rho_{\mathrm{NC},3}$ are non-full correlations. The difference in the predictability by either full knowledge or only the correlations will be further illumined in Sec. A.3.

The effect of being genuine n-partite entangled with vanishing n-partite correlations is not restricted to this particular mixture of equal amounts of the symmetric $|\mathrm{W}_3\rangle$ and $|\overline{\mathrm{W}}_3\rangle$ states. Instead, this phenomenon can be generalised which will be shown Sec. 3.3.1. On the other hand, the $\rho_{\mathrm{NC},3}$ state as given in Eq. 3.2 is suitable for an experimental study.

3.3 Entanglement with generalised class of states

The result of the previous section, where the state $\rho_{\mathrm{NC},3} = \frac{1}{2}(|\mathrm{W}_3\rangle\langle \mathrm{W}_3| + |\overline{\mathrm{W}}_3\rangle\langle\overline{\mathrm{W}}_3|)$ was prepared, can be generalised. Not only the mixture of the symmetric $|\mathrm{W}_3\rangle$ state together with the state $|\overline{\mathrm{W}}_3\rangle$ fulfils the requirement of showing tripartite entanglement while having vanishing full correlations. In this section, the general form of $|\mathrm{W}_3^g\rangle$ states is considered, which was defined as

$$|\mathrm{W}_3^g\rangle = \lambda|001\rangle + \mu|010\rangle + \nu|100\rangle. \tag{3.4}$$

First, an approach will be presented, how the "antistate" $|\overline{\psi}\rangle$ can found. Furthermore, an investigation concerning the entanglement of the mixture will be shown.

The term "antistate" refers to the state that one has to mix with the given state $|\psi\rangle$ to obtain a genuinely entangled state without n-partite correlations, i.e. the state $\rho = \frac{1}{2}(|\psi\rangle\langle\psi| + |\overline{\psi}\rangle\langle\overline{\psi}|)$ is considered. For example, $|\mathrm{W}_3\rangle$ and $|\overline{\mathrm{W}}_3\rangle$ are, using this terminology, antistates with respect to each other. Using $|\mathrm{W}_3^g\rangle$ as $|\psi\rangle$ results in a series of states with not only vanishing full correlations, but also with genuine tripartite entanglement.

3.3.1 Antistates

Here, the case of states with odd number of qubits will be treated, only. Consider the state $|\psi\rangle$ parameterised in the form

$$|\psi\rangle = \sum_{\mu_1,\ldots,\mu_n=0}^{1} \alpha_{\mu_1,\ldots,\mu_n}|\mu_1 \ldots \mu_n\rangle. \tag{3.5}$$

The corresponding antistate can be found by applying the operator $(\sigma_z\sigma_x\hat{K})^{\otimes n}$ onto the state $|\psi\rangle$ where the operator $\hat{K}$ denotes complex conjugation [13]. Therefore, the antistate of $|\psi\rangle$ is obtained to be

$$|\overline{\psi}\rangle = (\sigma_z \sigma_x \hat{K})^{\otimes n} |\psi\rangle = \sum_{\mu_1,\ldots,\mu_n=0}^{1} (-1)^{(\mu_1+\cdots+\mu_n)} \alpha^*_{1-\mu_1,\ldots,1-\mu_n} |\mu_1 \ldots \mu_n\rangle. \tag{3.6}$$

This scheme can be tested easily for the already known state $|\psi\rangle = |\mathrm{W}_3\rangle$, where $\alpha_{1,0,0} = \alpha_{0,1,0} = \alpha_{0,0,1} = \sqrt{\frac{1}{3}}$. This directly leads to the antistate $|\overline{\psi}\rangle = \alpha^*_{1,0,0}|011\rangle + \alpha^*_{0,1,0}|101\rangle + \alpha^*_{0,0,1}|110\rangle = \sqrt{\frac{1}{3}}(|011\rangle + |101\rangle + |110\rangle) = |\overline{\mathrm{W}}_3\rangle$, which is in agreement with the discussed state $\rho_{\mathrm{NC},3}$.

Note that this scheme results in a mixed state with vanishing full correlations, but up to now no statement about the entanglement of this state is made. Consider for example the Greenberger-Horne-Zeilinger state with $n = 3$ qubits, which can be chosen to be $|\psi\rangle = |\mathrm{GHZ}\rangle = \sqrt{\frac{1}{2}}(|000\rangle + |111\rangle)$. Writing this state in terms of Eq. 3.5, one obtains $\alpha^{\mathrm{GHZ}}_{0,0,0} = \alpha^{\mathrm{GHZ}}_{1,1,1} = \sqrt{\frac{1}{2}}$. Applying the presented method to find the antistate $|\overline{\psi}\rangle = |\overline{\mathrm{GHZ}}\rangle$, one obtains $\alpha^{\overline{\mathrm{GHZ}}}_{0,0,0} = \sqrt{\frac{1}{2}}$ and $\alpha^{\overline{\mathrm{GHZ}}}_{1,1,1} = -\sqrt{\frac{1}{2}}$. Therefore,

$$|\overline{\psi}\rangle = |\overline{\mathrm{GHZ}}\rangle = \frac{|000\rangle - |111\rangle}{\sqrt{2}}. \tag{3.7}$$

Mixing these states according to $\rho_{\mathrm{GHZ},\overline{\mathrm{GHZ}}} = \frac{1}{2}(|\mathrm{GHZ}\rangle\langle\mathrm{GHZ}| + |\overline{\mathrm{GHZ}}\rangle\langle\overline{\mathrm{GHZ}}|)$ results in the state $\rho_{\mathrm{GHZ},\overline{\mathrm{GHZ}}} = \frac{1}{2}(|000\rangle\langle 000| + |111\rangle\langle 111|)$, which is far from being entangled. Consequently, not all states deliver genuinely entangled states when mixed with their antistates. This discussion will be deepened in the next section for the generalised $|\mathrm{W}^g\rangle$ states with $n = 3$ qubits.

3.3.2 Genuine tripartite entanglement for $|\mathrm{W}^g\rangle$

Here, the entanglement of the generalised $|\mathrm{W}^g_3\rangle$ state, with an appropriate parameterisation defined as

$$|\psi\rangle = |\mathrm{W}^g_3\rangle = \cos(\alpha)\sin(\beta)|001\rangle + \sin(\alpha)\sin(\beta)|010\rangle + \cos(\beta)|100\rangle, \tag{3.8}$$

mixed with its antistate

$$|\overline{\psi}\rangle = |\overline{\mathrm{W}}^g_3\rangle = \cos(\alpha)\sin(\beta)|110\rangle + \sin(\alpha)\sin(\beta)|101\rangle + \cos(\beta)|011\rangle \tag{3.9}$$

will be investigated. For $\alpha, \beta \in [0, \frac{\pi}{2}]$ some special cases occur where no genuine tripartite entanglement can be detected. Firstly, for the cases of

a) $\alpha \in [0, \frac{\pi}{2}]$, $\beta = 0$

b) $\alpha = 0$, $\beta = \frac{\pi}{2}$ and

c) $\alpha = \frac{\pi}{2}$, $\beta = \frac{\pi}{2}$

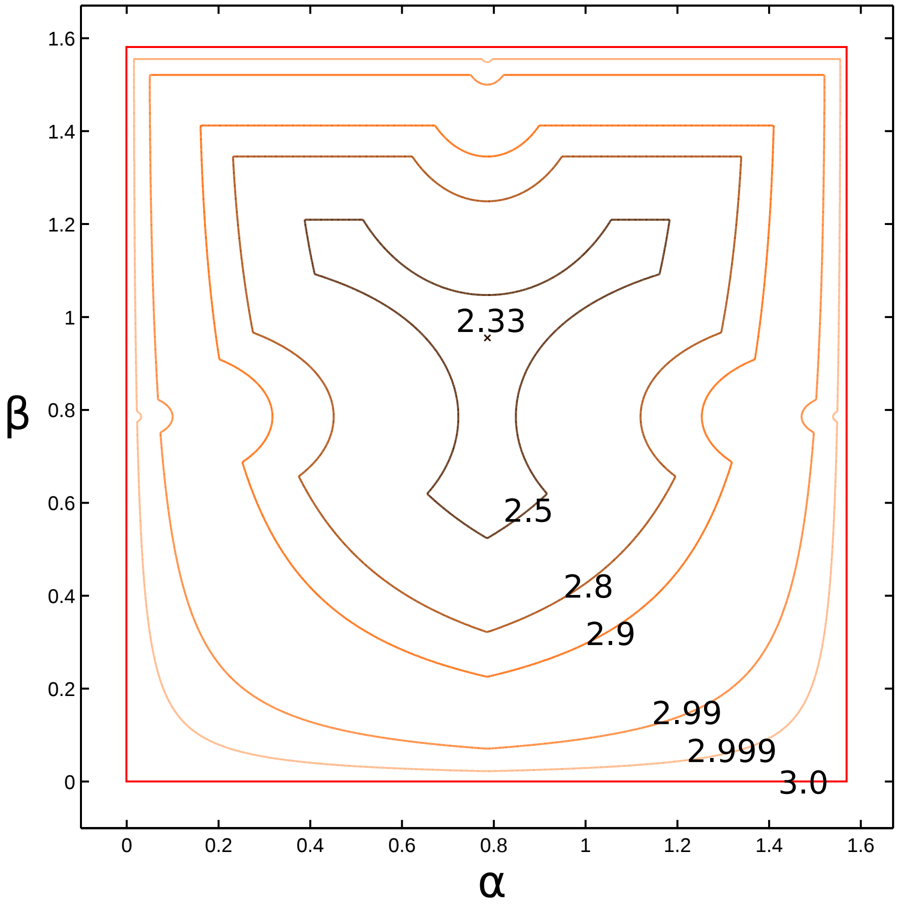

FIGURE 3.1: The upper bound for the maximally achieved value for $L = \max_{T^{\text{bi-prod}}}(T, T^{\text{bi-prod}})_G$ for the state of equal amounts of a generalised $|\psi\rangle = |\mathrm{W}_3^g\rangle$ state and its corresponding antistate $|\overline{\psi}\rangle = |\overline{\mathrm{W}_3^g}\rangle$ is shown as a contour plot. The outer frame, where $\alpha \in \{0, \frac{\pi}{2}\}$ or $\beta \in \{0, \frac{\pi}{2}\}$ holds, are the only states with no genuine tripartite entanglement. For all states within this frame, the maximally achieved value of L is clearly below $R = 3$, which indicates tripartite entanglement [13]. The spot at $\alpha = \pi/4 \approx 0.79$ and $\beta = \cos^{-1}\left(\sqrt{1/3}\right) \approx 0.96$ corresponds to the *no correlation* state $\rho_{\text{NC},3}$. For this state, the value of L can bounded from above by 7/3, clearly revealing entanglement.

the mixed state is fully separable. For example, in case a) the state becomes

$$\begin{aligned}\rho &= \frac{1}{2}(|\psi_{\beta=0}\rangle\langle\psi_{\beta=0}| + |\overline{\psi}_{\beta=0}\rangle\langle\overline{\psi}_{\beta=0}|) \\ &= \frac{1}{2}(|100\rangle\langle100| + |011\rangle\langle011|) \\ &= \frac{1}{2}\left(|1\rangle\langle1| \otimes |0\rangle\langle0| \otimes |0\rangle\langle0| + |0\rangle\langle0| \otimes |1\rangle\langle1| \otimes |1\rangle\langle1|\right),\end{aligned} \tag{3.10}$$

which is clearly not entangled at all. Case b) results in $\frac{1}{2}(|001\rangle\langle001| + |110\rangle\langle110|)$ while c) corresponds to $\frac{1}{2}(|010\rangle\langle010| + |101\rangle\langle101|)$. In all cases, the state is the mixture of two fully separable states.

Secondly, for specific choices of α and β the mixture of $|\psi\rangle$ and $|\overline{\psi}\rangle$ becomes bi-separable:

d) $\alpha \in (0, \frac{\pi}{2})$, $\beta = \frac{\pi}{2}$,

e) $\alpha = 0$, $\beta \in (0, \frac{\pi}{2})$ and

f) $\alpha = \frac{\pi}{2}$, $\beta \in (0, \frac{\pi}{2})$.

Case d) leads to a state

$$\begin{aligned}\rho &= \frac{1}{2}(|\psi_{\beta=\pi/2}\rangle\langle\psi_{\beta=\pi/2}| + |\overline{\psi}_{\beta=\pi/2}\rangle\langle\overline{\psi}_{\beta=\pi/2}|) \\ &= \frac{1}{2}\big[|0\rangle\langle 0| \otimes \big(\cos(\alpha)^2|01\rangle\langle 01| + \cos(\alpha)\sin(\alpha)|01\rangle\langle 10| \\ &\qquad\qquad + \cos(\alpha)\sin(\alpha)|10\rangle\langle 01| + \sin(\alpha)^2|10\rangle\langle 10|\big) \\ &\quad + |1\rangle\langle 1| \otimes \big(\cos(\alpha)^2|10\rangle\langle 10| + \cos(\alpha)\sin(\alpha)|10\rangle\langle 01| \\ &\qquad\qquad + \cos(\alpha)\sin(\alpha)|01\rangle\langle 10| + \sin(\alpha)^2|01\rangle\langle 01|\big)\big],\end{aligned} \tag{3.11}$$

being bi-separable according to $A|BC$. Analogously, the states of case e) are bi-separable with respect to the cut $B|AC$ while f) corresponds to $AB|C$.

The entanglement can be detected via the criterion using the correlation tensor as given in Sec. 2.4.1. For that purpose, one has to maximise over all bi-separable states. If then $L = \max_{\hat{T}^{\mathrm{bi-prod}}}(\hat{T}, \hat{T}^{\mathrm{bi-prod}})_G < (\hat{T}, \hat{T})_G = R$ holds, the state ρ with its correlation tensor $\hat{T}$ is genuinely entangled. As shown in Sec. A.5 and in [13], $R = 3$ holds for all α and β for the right-hand side. The optimisation over bi-separable states is performed to obtain values for the left-hand side L, the calculation for R and the result obtained for L is outlined in Sec. A.5. In full agreement with the already discussed cases of a) - f), only the states where $\alpha \in \{0, \frac{\pi}{2}\}$ or $\beta \in \{0, \frac{\pi}{2}\}$ are not genuinely tripartite entangled. This situation is illustrated in Fig. 3.1, where the respective value for the left-hand side L is shown for the possible values of α and β. All values with $L < 3$ indicate genuine tripartite entanglement. The limit value of 3 can only be reached if at least one of the expressions $\sin(\alpha)$, $\cos(\alpha)$, $\sin(\beta)$ or $\cos(\beta)$ vanishes.

3.4 Experimental preparation

This section explains the experimental setup for state preparation and describes the obtained measurement results. The experimental preparation of the state of interest $\rho_{\mathrm{NC},3} = 0.5(\rho_{\mathrm{W}_3} + \rho_{\overline{\mathrm{W}}_3})$ can be based on preparing the four qubit Dicke state $|D_4^{(2)}\rangle$ with

$$|D_4^{(2)}\rangle = \sqrt{\frac{1}{6}}\,(|HHVV\rangle + |HVHV\rangle + |HVVH\rangle + |VHHV\rangle + |VHVH\rangle + |VVHH\rangle)\,. \tag{3.12}$$

Afterwards, tracing out one qubit (without loss of generality the last qubit) [5] of the measured state results in the desired quantum state $\rho_{\mathrm{NC},3}$, i.e.

$$\mathrm{tr}_4(|D_4^{(2)}\rangle\langle D_4^{(2)}|) = \frac{1}{2}(|\mathrm{W}_3\rangle\langle\mathrm{W}_3| + |\overline{\mathrm{W}}_3\rangle\langle\overline{\mathrm{W}}_3|) = \rho_{\mathrm{NC},3}. \tag{3.13}$$

The full derivation of Eq. 3.13 can be found in Sec. A.4.

3.4.1 State preparation

To create the desired photonic four qubit Dicke state $|D_4^{(2)}\rangle$, a spontaneous parametric down conversion source with type II phase matching is used with pumping power high enough to enable fourfold coincidences. Initially, a $\mathrm{Nd}:\mathrm{YVO}_4$ laser with 10 W output power is used to pump a mode-locked Titan-Sapphire laser with pulse lengths of 130 fs and a repetition rate of 80 MHz [51]. The emitted light pulses with a centre wavelength of 780 nm are frequency-doubled by a lithium borate (LBO) crystal to achieve a wavelength of 390 nm. The pumping of the BBO (beta barium borate) crystal is then accomplished by using a bow-tie shaped cavity as depicted in Fig. 3.2. Essential requirements for adjustments of the cavity, like matching the round trip time of the laser pulse with the frequency of the additionally incoming pulses, i.e. locking the phases of the cavity to resonance [82], matching carrier envelope frequencies [83] and minimising the dispersion inside the cavity [84], are discussed in detail in [51, 85]. The BBO crystal inside the cavity is aligned for collinear type II phase-matching, see Sec. 2.6.1 and especially Eq. 2.67. In first order emission a pair of one horizontally and one vertically polarised photons are emitted while in second order emission two photons for each of those two polarisations are emitted. Thus, this second order emission events are of interest for the four qubit Dicke state presented here. Please note that also events of even higher orders may occur (see Sec. 2.6.1). The used pump power was $P = 0.66 \pm 0.06\,\mathrm{W}$. While a higher pump power results in a higher rate of events, higher orders of emission contribute more. The used pump power is a good compromise of state quality and measurement time.

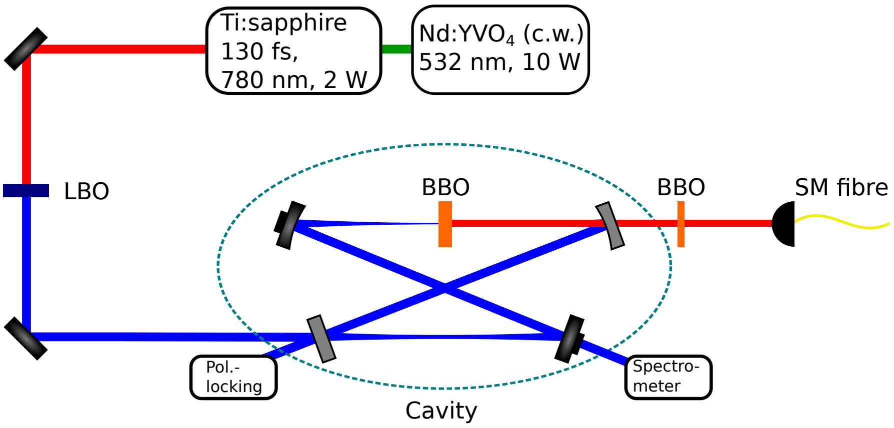

FIGURE 3.2: The SPDC source with type II phase matching (see Sec. 2.6.1) is used for creating entangled photons. Corresponding to the down conversion described in [85], a cavity with piezo-mounted mirrors is used which is pumped with a pulsed laser with 130 ns pulses @390 nm. A beta barium borate crystal of 1 mm thickness is used as the non-linear SPDC crystal inside the cavity. For compensation of walk-off effects between horizontal and vertical polarised photons, another BBO crystal with half thickness (0.5 mm) is used outside the cavity [85].

3.4.2 Analysis and measurement

After preparing the desired state as explained in the previous section and shown in Fig. 3.2, the state has to be measured. For that purpose a linear setup as shown in

Fig. 3.3 is used. The single mode fibre is directly connected to the aforementioned state preparation. Since a collinear configuration is applied, all created signal and idler photons are emitted in this spatial mode. An interference filter from Semrock with spectral bandwidth of 3 nm is used for spectral selection [51]. Afterwards, probabilistic (50 : 50, see Sec. 2.6.4) beamsplitters split up the incoming light to measure in the four spatial modes a, b, c and d. This leads to a permutationally invariant state of 4 qubits. Since the source is producing two photons with horizontal and two photons with vertical polarisation, the resulting state is the Dicke state $|D_4^{(2)}\rangle$[2]. In each spatial mode, a pair of Yttrium-vanadate crystals (YVO_4, see Sec. 2.6.3) is used to compensate birefringent phase shifts of the beamsplitters [51]. The YVO_4 crystals of 200 μm thickness are used pairwise with perpendicular orientation of optical axes [51]. This ensures a low order behaviour of the configuration and decreases wavelength dependencies. Finally, the qubits are analysed by using a polarisation analysis stage as described in Sec. 2.6.5

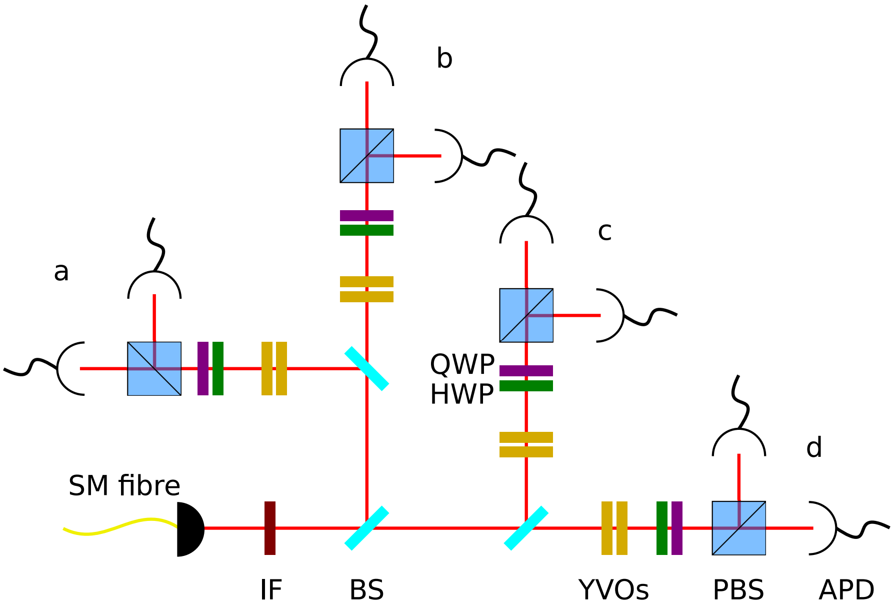

FIGURE 3.3: Since the prepared Dicke state $|D_4^{(2)}\rangle$ is permutationally invariant, the distribution of the photons into the four spatial modes a, b, c and d are performed with probabilistic beam splitters (BS). Spectral indistinguishability of the photons is improved with an interference filter. YVO_4 crystals are used for compensating birefringence of the beamsplitters [86] and setting the phase of the output modes [28]. They are used pairwise to achieve a zeroth order behaviour. Finally, the polarisation state in each mode is determined by using a polarisation analysis setup as explained in Sec. 2.6.5.

---

[2] While one wants to prepare the Dicke state with four qubits and two excitations, i.e. the symmetric superposition of all permutations of $|HHVV\rangle$, also higher orders may contribute. Consider the next higher order state with $\sqrt{\frac{1}{20}}(|HHHVVV\rangle + |HHVVVH\rangle + \cdots + |VVVHHH\rangle)$. Since only fourfold events are detected, this six qubit state can also contribute by losing two of its photons [86]. The measured state thus contains also admixtures of $|HHHV\rangle + |HHVV\rangle + \cdots + |VVVH\rangle$, which is directly obtained by losing two qubits, e.g. the last two qubits, of $|D_6^{(3)}\rangle$. While only the four qubit Dicke state with two excitations is aimed to be prepared, now admixtures of four qubit Dicke states with one and three excitations can be detected. By even smaller chance, the fourth order of the SPDC can be pumped, where eight photons are emitted. This leads to even lower additional admixtures of $|D_4^{(0)}\rangle = |HHHH\rangle$ and $|D_4^{(4)}\rangle = |VVVV\rangle$.

with a set of a half- and and a quarter-wave plate, a polarising beamsplitter and two single-photon avalanche photo diodes (APDs).

3.4.3 Tomography of $|D_4^{(2)}\rangle$

Here, the tomographic results of the state preparation are presented. First, the state obtained by a full quantum state tomography will be discussed. Afterwards, one qubit of the obtained quantum state will be projected on $|V\rangle$ and $|H\rangle$ to receive a $|\mathrm{W}_3\rangle$ and a $|\overline{\mathrm{W}}_3\rangle$ state, respectively. Furthermore, one qubit will be traced out to finally obtain the desired mixed state $\rho_{\mathrm{NC},3}$. This state will be used to show that its *full correlations* vanish while still being genuinely tripartite entangled.

Full tomography

The prepared state is tomographically analysed by means of the Pauli tomography scheme [15] with overcomplete measurement data (see Sec. 2.7), i.e. measurements are performed for all different combinations of 3^4 Pauli matrices ($\sigma_x^{\otimes 4}$, $\sigma_x \otimes \sigma_x \otimes \sigma_x \otimes \sigma_y$, $\ldots$, $\sigma_z^{\otimes 4}$). While already $4^4 - 1$ outcomes are sufficient for a complete tomography [15], here an overcomplete tomography is carried out. For each of the 3^4 measurement settings all 2^4 possible outcomes are monitored at the same time since for each of the four qubits two possible outcomes exist. Thus, $3^4 \cdot 2^4 = 6^4$ measurement outcomes are registered.

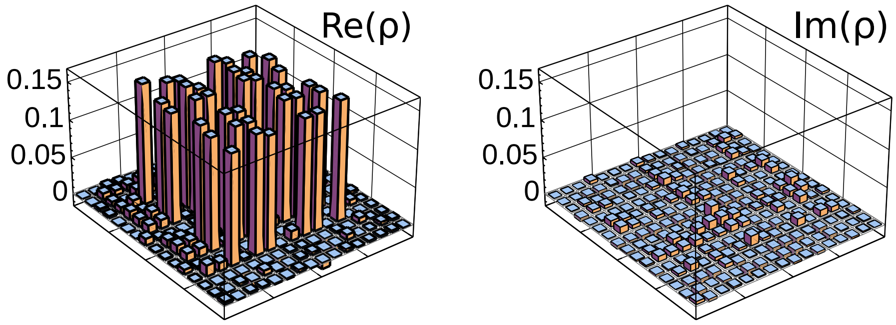

FIGURE 3.4: Real and imaginary part of the reconstructed density matrix obtained by full quantum state tomography of the experimental $|D_4^{(2)}\rangle$ state. The fidelity with respect to the target state $|D_4^{(2)}\rangle$ is $F(\rho, |D_4^{(2)}\rangle) = 0.92$.

The total measurement time of the executed tomography is more than 132 hours, where about $N_S \approx 1846$ events are detected per basis setting on average. The resulting state, whose density matrix is shown in Fig. 3.4, gives a fidelity with respect to the desired Dicke state $|D_4^{(2)}\rangle$ of $F\left(\rho, |D_4^{(2)}\rangle\right) = 0.920 \pm 0.005$. The overlap with states contributing due to higher order emissions is $F\left(\rho, |D_4^{(1)}\rangle\right) = 0.033 \pm 0.004$ and $F\left(\rho, |D_4^{(3)}\rangle\right) = 0.010 \pm 0.004$, respectively.

$|\mathrm{W}_3\rangle$ and $|\overline{\mathrm{W}}_3\rangle$ states and their mixture

For projecting one of the qubits of the measured Dicke state $|D_4^{(2)}\rangle$ onto $|V\rangle$ or onto $|H\rangle$ one obtains a $|\mathrm{W}_3\rangle$ and a $|\overline{\mathrm{W}}_3\rangle$ state, respectively. To obtain the mixture of those states according to Eq. 3.13, one qubit of the measured Dicke state has to be traced out. Thus, obtaining the data for the "no correlation state" $\rho_{\mathrm{NC},3}$ can be easily done.

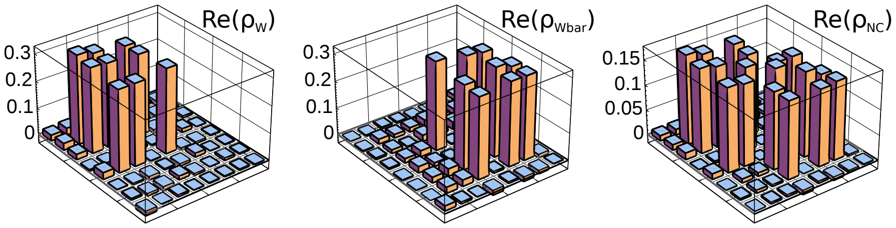

FIGURE 3.5: The real parts of the density matrices of $|\mathrm{W}_3\rangle$, $|\overline{\mathrm{W}}_3\rangle$ and $\rho_{\mathrm{NC},3}$ are shown. The imaginary parts of these states are close to zero (theoretically vanishing) and are therefore not depicted. The states are obtained from the measured four qubit state by projecting one qubit onto $|V\rangle$, onto $|H\rangle$ or by tracing it out, respectively.

In Fig. 3.5 the real parts of the density matrices for the measured state with one qubit projected onto $|V\rangle$ and onto $|H\rangle$ are shown as well as for tracing out one qubit.

3.4.4 Vanishing tripartite correlations

$\sigma_z \otimes \sigma_z \otimes \sigma_z$ correlation

Correlations are derived from the relative number of outcomes for a $\sigma_z^{\otimes 3}$ measurement for the $|\mathrm{W}_3\rangle$ and the $|\overline{\mathrm{W}}_3\rangle$ state. In Fig. 3.6 their relative number of events for each outcome in the $\sigma_z \otimes \sigma_z \otimes \sigma_z$ basis is shown for projecting one qubit onto $|V\rangle$ ($|\mathrm{W}_3\rangle$) and $|H\rangle$ ($|\overline{\mathrm{W}}_3\rangle$) as well as for tracing out one qubit ($\rho_{\mathrm{NC},3}$). The $|\mathrm{W}_3\rangle$ state's events VHH ("udd"), HVH ("dud") and HHV ("ddu") are expected each with probability $\frac{1}{3}$, which is in good agreement with the shown measured relative number of outcomes. The $|\overline{\mathrm{W}}_3\rangle$ state in contrast theoretically exhibits the events VVH ("uud"), VHV ("udu") and HVV ("duu") with probability of $\frac{1}{3}$, while all other outcomes are impossible. The experimental results are shown in the middle plot of Fig. 3.6. Evaluating the correlation according to Eq. 2.80, the high correlation for the $|\mathrm{W}_3\rangle$ state of $T_{ZZZ}^{\rho_{\mathrm{W}_3},\mathrm{exp}} = 0.904 \pm 0.033$ (theoretically $T_{ZZZ}^{\rho_{\mathrm{W}_3},\mathrm{theo}} = 1$) and the anticorrelation of $T_{ZZZ}^{\rho_{\overline{\mathrm{W}}_3},\mathrm{exp}} = -0.914 \pm 0.035$ (theoretically $T_{ZZZ}^{\rho_{\overline{\mathrm{W}}_3},\mathrm{theo}} = -1$) for the $|\overline{\mathrm{W}}_3\rangle$ state, can be calculated[3]. The prepared state

---

[3] $T_{ZZZ} \equiv T_{333}$. This section will use the alphanumerical indexing of the correlation tensor most of the time. For shortness, the measurement of $\sigma_z \otimes \sigma_z \otimes \sigma_z$ will also be denoted as a "measurement of ZZZ". See the definition of Pauli matrices in Sec. 2.1 for details.

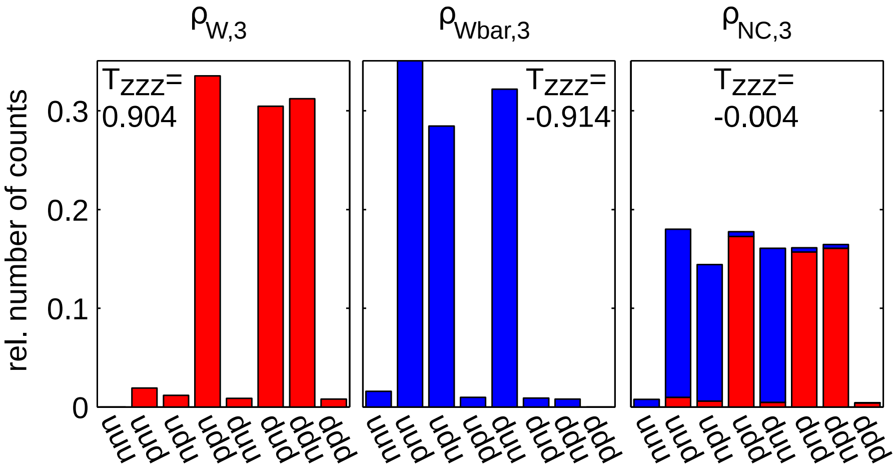

FIGURE 3.6: A measurement in the $\sigma_z^{\otimes 3}$ basis is performed. The relative number of counts for each outcome is shown here, where "u" and "d" denote up and down and are therefore corresponding for Z measurements to the outcomes V and H. The first two subfigures show the distribution of outcomes to all possible outcomes of a $\sigma_z \otimes \sigma_z \otimes \sigma_z$ measurement for the states $\rho_{\mathrm{W},3}$ and $\rho_{\overline{\mathrm{W}},3} \equiv \rho_{\mathrm{Wbar},3}$ obtained by projecting one qubit onto $|V\rangle$ and $|H\rangle$, respectively. In the last subfigure, their mixture is shown. Therefore, the distributions that are not symmetric in respect of exchanging "u" and "d" are now mixed to a distribution where $f_{uuu} \approx f_{ddd}$, $f_{uud} \approx f_{ddu}$, etc. Computing the correlations according to Eq. 2.80 leads to the given values of the high correlation for the $\rho_{\mathrm{W},3}$, the high anticorrelation for the $\rho_{\overline{\mathrm{W}},3}$ state as well as the almost uncorrelated result for the mixed state.

$\rho_{\mathrm{NC},3}$ has only a very low correlation of $T_{ZZZ}^{\rho_{\mathrm{NC},3},\mathrm{exp}} = -0.004 \pm 0.014$[4]. Therefore, the experimentally measured value for $T_{ZZZ}^{\rho_{\mathrm{NC},3},\mathrm{exp}}$ vanishes within its error bars.

Full correlations

In the previous section the measurement basis $\sigma_z \otimes \sigma_z \otimes \sigma_z$ was chosen to show the vanishing full correlation T_{ZZZ}. This feature is by no means restricted to this basis setting, but in contrast holds for all appearing full correlations. Because the considered state is a three qubit state, the correlation tensor has $4^3 = 64$ entries, of which $3^3 = 27$ are full correlations. In Fig. 3.7 and Fig. 3.8 the full correlations of the $|\mathrm{W}_3\rangle$ and $|\overline{\mathrm{W}}_3\rangle$ states are shown, where the red bars denote the experimental values with the respective error bars. The grey shaded extensions of the bars correspond to the theoretical values that one expects for the respective state and particular full correlation. For the $|\mathrm{W}_3\rangle$ state (Fig. 3.7) six full correlations are expected to be $\frac{2}{3}$ as can also be seen in Tab. 3.1. These correlations are those with all permutations of XXZ and YYZ, i.e. $T_{XXZ} =$

---

[4] Please note that the correlation of the mixed state is not necessarily exactly given by the sum of the correlations of the states obtained by projecting one qubit. In general, $T_{ZZZ}^{\rho_{\mathrm{NC},3},\mathrm{exp}} = \frac{1}{2}\left(T_{ZZZ}^{\rho_{\mathrm{W}_3},\mathrm{exp}} + T_{ZZZ}^{\rho_{\overline{\mathrm{W}}_3},\mathrm{exp}}\right)$ does not exactly hold for experimentally prepared states. Consider, for instance, that the last qubit of $|D_4^{(2)}\rangle$ shall be traced out to obtain the desired state $\rho_{\mathrm{NC},3}$. If now more events are measured where this qubit is found to be, e.g., $|V\rangle$ than $|H\rangle$, the projection onto $|V\rangle$ comes with better statistics and, thus, contributes more to the mixed state. Consequently, the mixed state is not exactly the equal sum of the projections. $\rho_{\mathrm{NC},3}^{\mathrm{exp}} \approx \frac{1}{2}\left(\rho_{\mathrm{W},3}^{\mathrm{exp}} + \rho_{\overline{\mathrm{W}},3}^{\mathrm{exp}}\right)$ is only an approximation if $\rho_{\mathrm{W},3}^{\mathrm{exp}}$ and $\rho_{\overline{\mathrm{W}},3}^{\mathrm{exp}}$ are obtained by projections and $\rho_{\mathrm{NC},3}^{\mathrm{exp}}$ by tracing out one qubit.

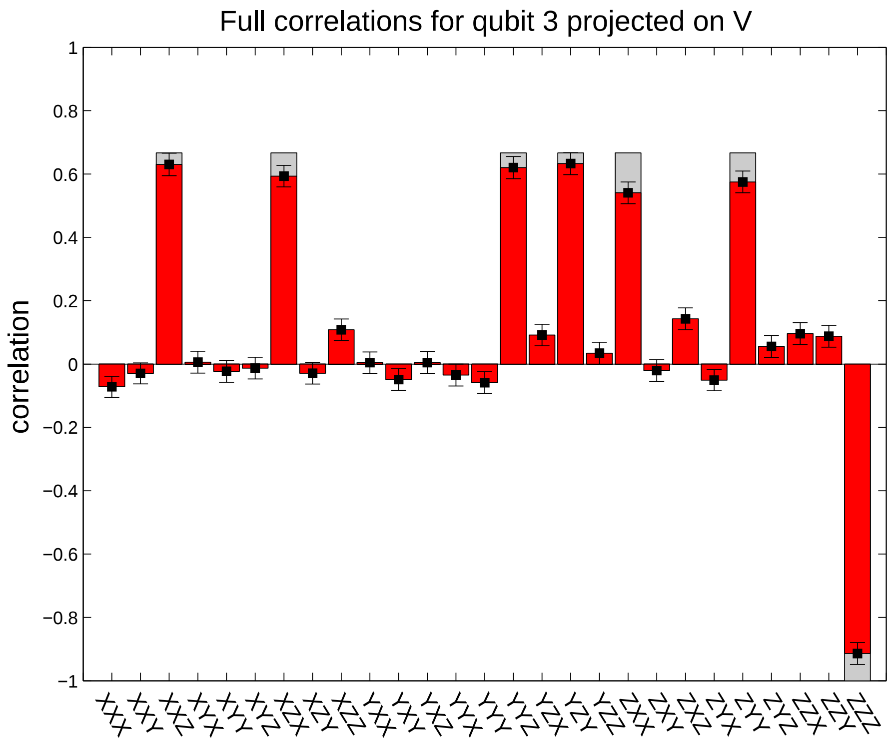

FIGURE 3.7: The full correlations of $|\mathrm{W}_3\rangle$ are depicted for measuring $XXX \equiv \sigma_x \otimes \sigma_x \otimes \sigma_x$, $XXY \equiv \sigma_x \otimes \sigma_x \otimes \sigma_y$, $XXZ \equiv \sigma_x \otimes \sigma_x \otimes \sigma_z$, $\ldots$, $ZZY \equiv \sigma_z \otimes \sigma_z \otimes \sigma_y$, $ZZZ \equiv \sigma_z \otimes \sigma_z \otimes \sigma_z$. Six full correlations are close to their theoretical prediction of $\frac{2}{3}$ while the remaining 20 full correlations are estimated to vanish, which is in good agreement with the experimental results.

$T_{XZX} = T_{ZXX} = T_{YYZ} = T_{YZY} = T_{ZYY} = \frac{2}{3}$. The correlation of $\sigma_z^{\otimes 3}$ reveals maximal anticorrelation of $T_{ZZZ} = -1$. The outcomes in all remaining 20 measurement bases are predicted to be not correlated at all, i.e. $T_{ijk} = 0$ for all remaining ijk. Fig. 3.8 depicts the full correlations of the $|\overline{\mathrm{W}}_3\rangle$ state. Again, 20 full correlations are vanishing. The remaining seven full correlations are just opposite to those of the $|\mathrm{W}_3\rangle$ state such that $T_{ZZZ} = 1$ and the six correlations which were said to be $\frac{2}{3}$ are expected to be $-\frac{2}{3}$ now. Although the experimental values may slightly deviate from the theoretical predictions in many cases, the general behaviour of the full correlations can also be read from the experimental measured data.

Finally, in Fig. 3.9 the full correlations of the mixed state $\rho_{\mathrm{NC},3}$ are shown. The error bars are denoting the interval of one standard deviation of the correlations. Therefore, 68.3% of the correlations are expected to vanish within their error bars, which is in good agreement with the measured correlations, where 21 of 27 correlations vanish within one standard deviation, while six correlations do not. By using error bars of 2σ one would expect 95.4% of the correlations to vanish, i.e. 26 correlations, which again is in good agreement with the results. Merely the correlation T_{XXX} vanishes only for error bars of approximately 3σ, but this is still in good accordance with the predictions. Please note the different scaling of the figures for the $|\mathrm{W}_3\rangle$ (Fig. 3.7) and $|\overline{\mathrm{W}}_3\rangle$ (Fig. 3.8) states and the one for the full correlations of Fig. 3.9.

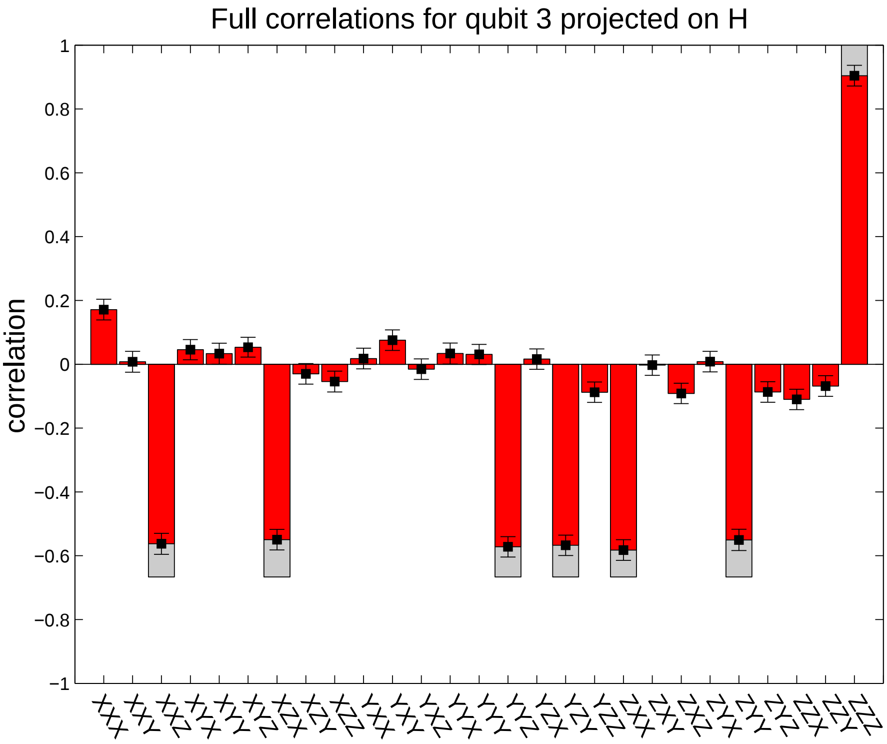

FIGURE 3.8: According to Fig. 3.7, the full correlations of $|\overline{\mathrm{W}}_3\rangle$ are shown. While the same 20 full correlations are expected to vanish, the 7 correlations that deviate from zero for the $|\mathrm{W}_3\rangle$ state have correlations with opposite sign for the $|\overline{\mathrm{W}}_3\rangle$ state.

Independence of measurement settings

It could be shown previously that all full correlations of the $\rho_{\mathrm{NC},3}$ state are vanishing. For this case, all 3^3 *Pauli directions* were considered. One might claim that this feature does only hold due to the Pauli measurement scheme. If one would use different measurement bases, so one could assert, full correlations unequal to zero may occur. Here, this proposition will be refuted by using a rotated measurement basis setting. Exemplarily, the measurement of $(\sin(\theta)\sigma_y + \cos(\theta)\sigma_z) \otimes \sigma_z \otimes \sigma_z \equiv rZZ$ is taken into account with $\theta = \pi/4$. For this case of a rotated basis setting, the relative number of counts per possible outcome is depicted in Fig. 3.10. As in the corresponding figure for measuring $\sigma_z^{\otimes 3}$ (Fig. 3.6) the states $|\mathrm{W}_3\rangle$ and $|\overline{\mathrm{W}}_3\rangle$ still exhibit a correlation. The correlation for the projection onto $|V\rangle$, such that the $\rho_{\mathrm{W},3}^{\mathrm{exp}}$ state is obtained, delivers for the rotated measurement setting of $(\sin(\theta)\sigma_y + \cos(\theta)\sigma_z) \otimes \sigma_z \otimes \sigma_z$ the value $T_{rZZ}^{\rho_{\mathrm{W},3,\mathrm{exp}}} = -0.610 \pm 0.035$, where the theoretical value is found to be $T_{rZZ}^{\rho_{\mathrm{W},3,\mathrm{theo}}} = -\sqrt{\frac{1}{2}} \approx -0.707$. Correspondingly, the projection onto $|H\rangle$ gives $T_{rZZ}^{\rho_{\overline{\mathrm{W}},3,\mathrm{exp}}} = 0.606 \pm 0.033$, $T_{rZZ}^{\rho_{\overline{\mathrm{W}},3,\mathrm{theo}}} = \sqrt{\frac{1}{2}} \approx 0.707$, while measurements for the mixed state, obtained by tracing out one qubit of the measured four qubit state, are almost completely uncorrelated ($T_{rZZ}^{\rho_{\mathrm{NC},3,\mathrm{exp}}} = 0.020 \pm 0.024$). Small deviations between the experimental and the theoretical values can be explained by contributions of higher order emissions as explained in Sec. 3.4.2.

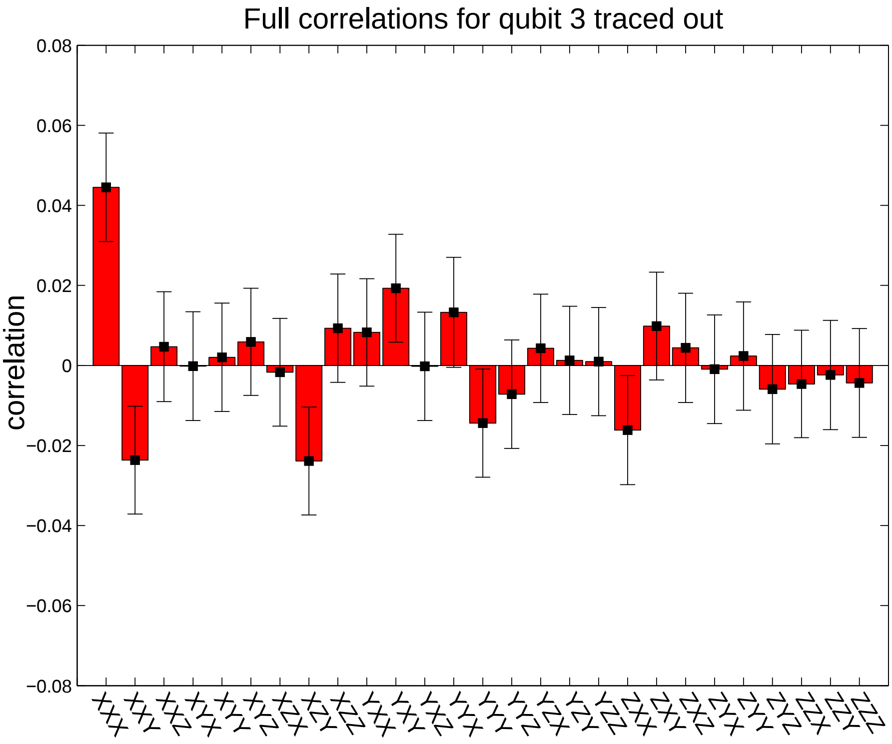

FIGURE 3.9: Mixing the states $|\mathrm{W}_3\rangle$ and $|\overline{\mathrm{W}}_3\rangle$, the correlations are given by their respective averages. Therefore, theoretically, all full correlations should vanish. Here, the full correlations of the experimentally prepared state are shown. Most of the correlations vanish within one standard deviation. Please note the different scaling compared to the plots of Fig. 3.7 and Fig. 3.8

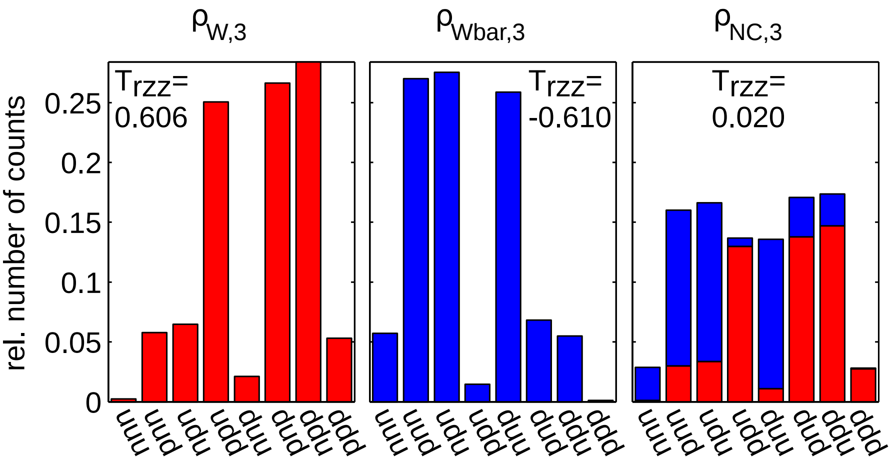

FIGURE 3.10: Measurement of $\frac{\sigma_y+\sigma_z}{2}\otimes\sigma_z\otimes\sigma_z$: while the states obtained by projections of one qubit show correlations for this choice of measurement, the state given by tracing out one qubit exhibits uncorrelated results. The relative number of counts for each outcome is almost symmetric under exchange of "u" and "d" for the mixed state $\rho_{\mathrm{NC},3}$ which leads to a vanishing correlation according to Eq. 2.80. The "r" as the index of the correlation tensor elements abbreviates the rotated direction, i.e. $\sigma_r = \frac{\sigma_y+\sigma_z}{2}$ in this case.

Non-full correlations

According to the previously shown relative number of counts for a rotated measurement and the thereby resulting vanishing correlation, the correlations for measurements with arbitrary angles are determined. For that purpose, a measurement is performed in $\sigma_r \otimes \sigma_z \otimes \sigma_z$ basis, where $\sigma_r = \sin(\theta)\sigma_y + \cos(\theta)\sigma_z$ denotes a rotated basis. In Fig. 3.11 the correlations for different angles are shown. For the considered $\rho_{\text{NC},3}$ state the full correlation along the corresponding direction is shown (green curve). Furthermore, non-full correlations are depicted (red and black curves). As already explained in this chapter's introduction (Sec. 3.2), the fact that full correlations are vanishing does not allow to infer anything about non-full correlations. Fig. 3.11 shows strikingly that, although the full correlations vanish, non-full correlations may occur for some basis settings. These non-full correlations oscillate with the chosen measurement setting, such that a measurement of for instance $\pm\sigma_z \otimes \sigma_z \otimes \sigma_0$ $(\theta = 0,\ \theta = \pi)$ of the used state reveals correlation up to $\mp\frac{1}{3}$, while the measurement results of the setting $\sigma_y \otimes \sigma_z \otimes \sigma_0$ $(\theta = \pi/2)$ are uncorrelated.

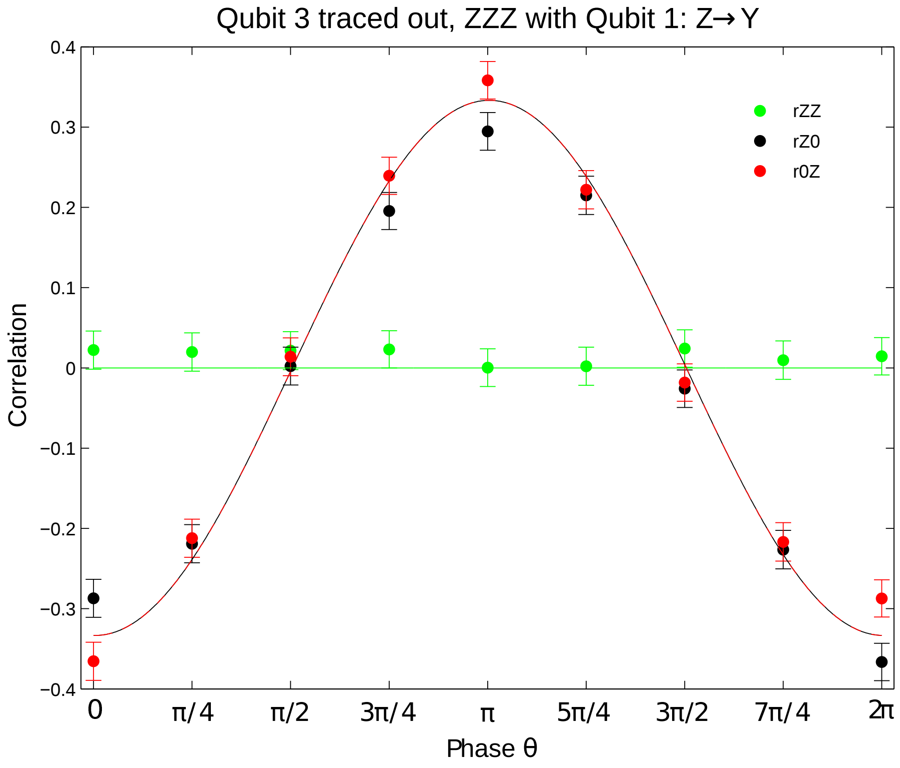

Figure 3.11: The measurement setting is rotated such that measurements are performed in $\sigma_r \otimes \sigma_z \otimes \sigma_z$ basis with $\sigma_r = \sin(\theta)\sigma_y + \cos(\theta)\sigma_z$. The green curve shows the theoretical prediction for the full correlations, which are all vanishing irrespectively of the used angle θ. For the black curve the non-full correlations between the first two qubits are considered, i.e. they correspond to T_{rZ0}, while the red curve shows the theoretical values for the correlations T_{r0Z}, where r denotes the rotated basis setting. Furthermore, the measured values are shown together with their standard deviations. Although some deviations occur, the measured values for the correlations are in good agreement with the theoretical values and reflect tellingly the fact that non-full correlations may still persist even if all full correlations vanish.

3.4.5 Genuine tripartite entanglement

The second demand of the mixed state $\rho_{\mathrm{NC},3}$ beside having vanishing full correlations is to be genuinely tripartite entangled. Constructing a state with vanishing full correlations, but without this second demand would also lead to such trivial results like the maximally mixed state, where the only contributing correlation is $T_{111} = 1$. Therefore, it will be shown in this section that the experimental state is also genuinely tripartite entangled, that is, the state is not only not fully separable, but also not bi-separable. The ansatz to reveal the genuine tripartite entanglement is the one already introduced in Sec. 2.4.1 [48]. Thus, one has to find the bi-separable state $\rho^{\mathrm{bi-prod}}$ with the correlation tensor $\hat{T}^{\mathrm{bi-prod}}$ which maximises the inner product $(\hat{T}, \hat{T}^{\mathrm{bi-prod}})_G$ for a given metric G and the correlation tensor $\hat{T}$ of the state under study. If one can find a metric G such that the maximum over all states $\rho^{\mathrm{bi-prod}}$ is less than the value of the inner product $(\hat{T}, \hat{T})_G$, the state is not bi-separable and consequently genuinely entangled. This was more explicitly discussed in Sec. 3.3.2 for the case of generalised $|\mathrm{W}^g\rangle$ states as defined in Eq. 2.51, where genuinely tripartite entangled states of a more general form are studied with vanishing full correlations. Here, the entanglement of only the state $\rho_{\mathrm{NC},3}$ is studied.

For this purpose, the metric G is chosen such that only correlations of the state $\rho_{\mathrm{NC},3}$ are involved, i.e. $G_{\Pi(xx0)} = G_{\Pi(yy0)} = G_{\Pi(zz0)} = 1$ where Π denotes all permutations of the indices. All other entries of G are chosen to be zero. Consider $\hat{T}$ to be the correlation tensor of the theoretical state $\rho_{\mathrm{NC},3}$, then the maximally achieved value for the separability criterion is [13]

$$\max_{\hat{T}^{\mathrm{bi-prod}}} \left(\hat{T}, \hat{T}^{\mathrm{bi-prod}}\right)_G = \frac{7}{3}, \tag{3.14}$$

which is obtained for the state $\rho^{\mathrm{bi-prod}} = (\cos(\theta)\,|PP\rangle - \sin(\theta)\,|MM\rangle) \otimes |P\rangle$ with $|P\rangle = \sqrt{\frac{1}{2}}\,(|0\rangle + |1\rangle)$ (see Sec. 2.1) and $\theta = \frac{1}{2}\tan^{-1}\left(\frac{3}{4}\right)$. If the value of $(\hat{T}, \hat{T}^{\mathrm{exp}})_G$ is greater than the limit value of $\frac{7}{3}$, genuine tripartite entanglement is detected. Using the correlation tensor $\hat{T}^{\mathrm{exp}}$ of the experimentally obtained state ρ^{exp}, one finds with the given metric G that $(\hat{T}, \hat{T}^{\mathrm{exp}})_G = 2.86 \pm 0.014$, clearly beating the threshold of $7/3 \approx 2.33$. Consequently, the experimentally prepared state is detected to be genuinely tripartite entangled.

3.5 Conclusion

Naively, one would assume that a genuinely entangled state has non-vanishing full correlations, representing the correlations between the outcomes of all involved parties. Furthermore, a state where all full correlations cannot be distinguished from zero could intuitively be assumed to be at least bi-separable. Instead, in this chapter the considered incoherent summation of the (symmetric) $|\mathrm{W}_3\rangle$ and the $|\overline{\mathrm{W}}_3\rangle$ state was experimentally proven to be genuinely tripartite entangled. Moreover, its full correlations were shown to be that small that they are not significantly distinguishable from zero. To indicate that not only the choice of Pauli measurement directions causes the full correlations to vanish, also rotated measurement directions were exemplarily considered. The discrimination between knowing the exact measurement outcomes of each party and only knowing the correlations between all the other measurement outcomes is important for understanding this feature, as it was discussed and illustrated with the help of an example. Finally, the insights from studying the symmetric $|\mathrm{W}_3\rangle$ and its antistate were used to enquire the generalised three qubit $|\mathrm{W}_3^g\rangle$ state. It was shown that only very special choices of the coefficients α and β lead to a separable state when mixing with its antistate. As soon as all three states, $|001\rangle$, $|010\rangle$ and $|100\rangle$, contribute to the superposition of $|\mathrm{W}_3^g\rangle$, the resulting state is genuinely tripartite entangled.

The question whether this phenomenon is restricted to cases of odd numbers of qubits is still open. It might be possible to find genuinely entangled states for even number of qubits n with vanishing n-partite correlations if more than two states are incoherently summed. Correspondingly, if it is sufficient to incoherently sum two pure states with an even number of qubits, is not ascertained, yet.

Chapter 4

Bias of Estimators

4.1 Introduction

Quantum state tomography [14] is the process to determine an unknown quantum state by measuring the outcomes by means of a certain measurement protocol like the Pauli scheme given in Sec. 2.7. The quantum state tomography is a widely-used method [87–91]. For that purpose, it is important to control the experimental parameters such that the measured quantum state is in good agreement with the desired state. Besides this natural experimental challenge, the finite statistics of measurements may cause further inconvenience. The measurement process delivers count values c_i^S for the measurement outcome i in the used basis setting S. Normalising these counts to the total number of counts of the respective basis, one obtains relative numbers. These frequencies can only be approximations with finite precision of probabilities of the underlying quantum state. Directly using the obtained frequencies as probabilities delivers a matrix $\hat{\rho}_{\mathrm{LIN}}$ which does not necessarily represent a (physical) density matrix, i.e. $\hat{\rho}_{\mathrm{LIN}} \ngeq 0$. Since physical density matrices are needed to evaluate non-linear functions, methods to ensure the physicality are applied.

Although strictly speaking not being correct, matrices ρ with $\rho \ngeq 0$ will be called "states" throughout this chapter to avoid lengthy phrases. Also, the quotation marks will often be skipped for those unphysical states to improve the readability while the clarification about the physicality is made where relevant. In this chapter, different ways are discussed how to estimate a quantum state or at least particular quantities of interest of the underlying quantum state. This analysis is based on a large amount of states such that statistical distributions can be considered. Thus, unless otherwise noted, the used quantum states are obtained by numerical simulations.

One requirement of all quantum state reconstruction methods is that the reconstructed state is close to the underlying quantum state. Assuming that the state prepared in the laboratory may give a fidelity of $F(\rho_{\mathrm{exp}}, \rho_{\mathrm{theo}}) = 0.8$ with respect to the theoretical state ρ_{theo}, it is a further, natural and reasonable demand that the reconstructed state gives $F(\rho_{\mathrm{reconstructed}}, \rho_{\mathrm{theo}}) = 0.8$ on average as well. Due to the finite statistics of the measurement, some deviations may occur, but for successful state reconstruction the theoretical value should be at least within the error bars of the reconstructed quantum state. This analysis will show that in general this is not the case.

4.2 State reconstruction

Here, an overview of different strategies to obtain the density matrices of quantum states from measured data is given. In the first part the most naive and straightforward ansatz is presented, where the obtained frequencies are interpreted to be probabilities. Its drawback, i.e. the possibility to give unphysical density matrices, will be discussed. This directly motivates the search for various methods that ensure physical results. To conclude this section, methods of practical use to compute error bars will be shown.

4.2.1 Using frequencies as probabilities

According to Sec. 2.7, one measurement strategy to determine an unknown n qubit state is to measure all 3^n tensor products of the three Pauli matrices σ_x, σ_y and σ_z. In each measurement setting, 2^n different outcomes are possible. The positive-operator-valued-measure (POVM) [92] elements for a given setting S are denoted by M_i^S, the absolute number of counts of outcome i in setting S is labelled by c_i^S and their respective sum in a given basis by N_S. Therefore, the frequency of outcome i when measuring setting S is given by $f_i^S = c_i^S/N_S$. According to Born's rule [93] the probability of outcome i in basis S for a quantum state ρ_0 is given by $P_{\rho_0}(i|S) = \mathrm{tr}(\rho_0 M_i^S)$ where the frequency f_i^S is an approximation to the probability $P_{\rho_0}(i|S)$.
The most straight-forward way to obtain a density matrix by the count data is to find the state $\hat{\rho}_{\mathrm{LIN}}$ that fulfils the condition $f_i^S = P_{\hat{\rho}_{\mathrm{LIN}}}(i|S)$ for all outcomes i in all settings S. This can be done directly for a complete set of $4^n - 1$ measurement results [15]. For an overcomplete measurement protocol as described in Sec. 2.7, the 6^n obtained measurement results over-determine the set of equations. The state in best agreement with the overcomplete set of measurement results is determined either by an explicit fit, i.e. $\hat{\rho}_{\mathrm{LIN}} = \arg\min_\rho (f_i^S - P_\rho(i|S))^2$ [94], or one uses the reconstruction scheme as defined in Eq. 2.77 [28]. Unfortunately, this does not ensure $\hat{\rho}_{\mathrm{LIN}} \geq 0$. Although it is possible to use $\hat{\rho}_{\mathrm{LIN}}$ to estimate for example the fidelity of the reconstructed state with respect to a pure target state (see Sec. 2.4.3), some non-linear measures like the entropy (see Sec. 2.4.6) or the negativity of a partial transpose (see Sec. 2.4.1) of a state cannot be evaluated meaningfully. The possibly occurring negative eigenvalues λ_i of $\hat{\rho}_{\mathrm{LIN}}$ could for instance cause the entropy $S(\rho) = -\sum_{i=1}^{2^n} \lambda_i \ln(\lambda_i)$ to be an imaginary number, which is away from any physical interpretation.
Therefore, different schemes to ensure the physicality of the reconstructed states are applied.

4.2.2 State reconstruction with ensured physicality

The general approach of quantum state reconstruction methods that guarantee to deliver a physical density operator is to use an optimisation with the constraint $\rho \geq 0$. The task is to find the density operator that is in *best* agreement with the measured data. Various definitions of the meaning of "best" finally result in different methods. Mathematically speaking one defines this property by

$$\hat{\rho} = \arg\max_{\rho \geq 0} T(\rho|f) \tag{4.1}$$

with the target function T which quantifies the agreement of a density operator ρ with the observed set of given frequency data f.

Least squares

An intuitive way to describe the agreement of the observed frequencies with the probabilities of a quantum state is given by a *least-squares* (LS) ansatz. For general weights $w(i|S)$ the target function can be expressed as

$$T_{\mathrm{LS}} = -\sum_{i,S} w(i|S)\left(f_i^S - P_{\rho_0}(i|S)\right)^2, \tag{4.2}$$

i.e. for the case of $f_i^S = P_{\rho_0}(i|S)$ the target function vanishes while for increasing deviations T_{LS} decreases.
Different choices for the used weights $w(i|S)$ are possible. For the most prominent case of *free least-squares* (FLS) [15] the weights are the inverse probabilities[1] such that

$$T_{\mathrm{FLS}} = -\sum_{i,S} \frac{1}{P_{\rho_0}(i|S)}\left(f_i^S - P_{\rho_0}(i|S)\right)^2. \tag{4.4}$$

Consequently, deviations for low probabilities have greater weights than those for larger $P_{\rho_0}(i|S)$.

Maximum likelihood

Another commonly used [87, 88] target function to obtain a quantum state with $\rho \geq 0$ is given by the *maximum likelihood* (ML) estimation [16]. With this approach, one is looking for the state ρ whose probability to deliver the given frequency data f is largest. As a result, the likelihood of the desired state ρ is largest. Because the probability for a measurement outcome i when measuring the operator M_i^S of the candidate state ρ is given by $P_{\rho_0}(i|S) = \mathrm{tr}(\rho_0 M_i^S)$, the likelihood function of observing a series of N events $i_1, \ldots, i_N$ in bases $S_1, \ldots, S_N$ is given by [95]

$$\tilde{\mathcal{L}}(\rho) = P_\rho(i_1|S_1)P_\rho(i_2|S_2)\cdots P_\rho(i_N|S_N) = \prod_{j=1}^{N} P_\rho(i_j|S_j). \tag{4.5}$$

Ordering and regrouping the events according to their outcome, the function now reads as [96]

$$\tilde{\mathcal{L}}(\rho) = \prod_{i,S} P_\rho(i|S)^{c_i^S}. \tag{4.6}$$

---

[1] If the number of counts N_S per basis setting S are not equal for all settings, this term should be slightly modified. According to [15]

$$T_{\mathrm{FLS}} = -\sum_{i,S} \frac{1}{N_S P_{\rho_0}(i|S)}\left(c_i^S - N_S P_{\rho_0}(i|S)\right)^2 = -\sum_{i,S} \frac{N_S}{P_{\rho_0}(i|S)}\left(f_i^S - P_{\rho_0}(i|S)\right)^2. \tag{4.3}$$

For equal number of measurement repetitions for all settings, the factor N_S does not change the weighting.

Because of the monotony of the logarithm it is also possible - and easier - to maximise instead of Eq. 4.6 its logarithm, the so called *log-likelihood*:

$$\mathcal{L}_c(\rho) = \log(\tilde{\mathcal{L}}(\rho)) = \sum_{i,S} c_i^S \log\left(P_\rho(i|S)\right). \tag{4.7}$$

Renormalising this function by the total number of events N_S per respective basis setting, the (log-)likelihood becomes now $\mathcal{L}_f(\rho) = \sum_{i,S} f_i^S \log(P_\rho(i|S))$. Therefore, this leads to the target function

$$T_{\mathrm{ml}} = \sum_{i,S} f_i^S \log(P_{\rho_0}(i|S)), \tag{4.8}$$

which again is optimised such that according to Eq. 4.1 $\rho \geq 0$ holds. Please note that Eq. 4.8 holds only if all the numbers of counts per basis setting N_S are equal for each setting S. Because this does not apply for measured states in most cases, those states may be reconstructed by using Eq. 4.7 instead.

There are also modifications of the maximum likelihood approach used based on statistical considerations conceivable[2]. One exemplary variation of ML is the *hedged maximum likelihood* (HML) estimation [18]. Additionally to the used likelihood function $\mathcal{L}_{\mathrm{ML}}(\rho) = P(f|\rho)$, a hedging function will be used such that [18]

$$\mathcal{L}_{\mathrm{HML}} = \mathcal{L}_{\mathrm{ML}} h(\rho) = \mathcal{L}_{\mathrm{ML}} \det(\rho)^\beta. \tag{4.9}$$

Therefore, the target function one has to maximise over reads for the HML case with $\beta > 0$ [21] and with the same statistics N_S for all settings as

$$T_{\mathrm{HML}} = \sum_{i,S} f_i^S \log(P_{\rho_0}(i|S)) + \beta \log(\det(\rho_0)). \tag{4.10}$$

As a direct consequence, the hedged maximum likelihood approach favours states with larger determinant. Since the density matrix expressed in its eigenvalues $\{\lambda_i\}_i$ and its eigenvectors $\{|\psi_i\rangle\}_i$ reads as $\rho = \sum_{i=1}^{2^n} \lambda_i |\psi_i\rangle\langle\psi_i|$, the determinant becomes $\det(\rho) = \prod_{i=1}^{2^n} \lambda_i$. Thus, rank-deficit states, where $\lambda_i = 0$ holds for at least one i, have a vanishing determinant, lead to an infinite value of the target function. Therefore, these rank-deficit states are forbidden and will not be reconstructed as such. The hedged maximum likelihood approach hence prevents states with low rank, but in return ensures better error estimation for parametric bootstrapping [18, 21], see also Sec. 4.2.4.

The state estimates, i.e. for example the reconstructed state by means of the maximum likelihood target function $\hat{\rho}_{\mathrm{ML}} = \arg\max_{\rho\geq 0} T_{\mathrm{ml}}(\rho|f)$ and $\hat{\rho}_{\mathrm{FL}} = \arg\max_{\rho\geq 0} T_{\mathrm{fls}}(\rho|f)$ for the free least-squares ansatz, respectively, can be used to evaluate functions of the density operators directly. While the entropy $S(\hat{\rho}_{\mathrm{LIN}})$ of the linearly reconstructed state is a quantity without physical interpretation for $\hat{\rho}_{\mathrm{LIN}} \not\geq 0$, the entropy $S(\hat{\rho}_{\mathrm{ML}})$ or $S(\hat{\rho}_{\mathrm{FL}})$ of the reconstructed state corresponds to the desired quantity. All kinds of measures can be obtained with $\hat{\rho}_{\mathrm{ML}}$ and $\hat{\rho}_{\mathrm{FLS}}$ while the quality of the least-squares and the maximum likelihood estimates will be analysed later on.

---

[2] For a motivation, consider tossing a coin. If all outcomes are, e.g., *head*, the maximum likelihood ansatz gives a probability of unity that the outcome is *head* and zero for *tail*. In contrast, by a finite number of tosses the probability for *tail* cannot be said to be zero since a small, but finite probability is still in accordance with the measurement outcomes. Variations of maximum likelihood take this into account and prefer more mixed states even if their likelihood according to Eq. 4.7 would be slightly less.

4.2.3 Convex optimisation

To reconstruct the quantum state via fitting, Eq. 4.1 has to be solved. Irrespective of the used target function, i.e. whether one prefers the free least-squares (Eq. 4.4), another variant of the least-squares (Eq. 4.2), the maximum likelihood approach (Eq. 4.8) or the hedged maximum likelihood (Eq. 4.10), the task to solve is by the additional condition of $\rho \geq 0$ a constraint problem [97]. Thus, the expression to optimise can be changed into [21]

$$\hat{\rho} = \arg\max_{\rho}(T(\rho|f) + t\log(\det(\rho))), \tag{4.11}$$

such that the *barrier term* expression $T_{\text{barrier}} = t\log(\det(\rho))$ ensures the physicality of ρ explicitly, but therefore the optimisation itself runs over all ρ. At boundary when leaving the space of physical density operators $\rho \geq 0$, at least one eigenvalue vanishes [21]. Therefore, the determinant vanishes when approaching the border of the space of physical states and thus, the logarithm diverges according to $\lim_{\rho\to\det(\rho)=0}\log(\det(\rho)) = -\infty$. Consequently, the step of quitting the physical subspace is punished. Note that in general unphysical states are not forbidden with this barrier term themselves, but only the step of reaching them. Thus, it is crucial to start with a physical state. In practice, a good starting state for the optimisation is the maximally mixed state ρ_{WN}.

A helpful parametrisation of ρ is given by

$$\rho(x) = \frac{1}{2^n}\sum_{i=1}^{4^n} x_i S_i \tag{4.12}$$

with $x_1 = 1$. x_1 is kept fix to ensure $\text{tr}(\rho(x)) = 1$. Since this parametrisation is fully equivalent to Eq. 2.18, the x_i correspond to the elements of the correlation tensor, while S_i are the basis elements with $S_i = \sigma_{i_1} \otimes \sigma_{i_2} \otimes \cdots \otimes \sigma_{i_n}$, where $i_j \in \{0,1,2,3\}$ indexes the Pauli matrix σ_{i_j} of the jth qubit.
Instead of maximising Eq. 4.11, the problem is turned into a minimisation with [21]

$$\hat{\rho} = \arg\min_{\rho}(\underbrace{-T(\rho|f)}_{\tilde{T}(\rho|f)} - t\log(\det(\rho))) \tag{4.13}$$

where $\tilde{T}(\rho|f) = -T(\rho|f)$ and $\tilde{T}_{\text{barrier}} = -T_{\text{barrier}}$, i.e. the tilde denotes the inverted signature. Showing the convexity of the problem can be done by inquiring the second derivative of the optimisation given in Eq. 4.13. The barrier term's first and second derivatives read as [97]

$$\frac{\partial \tilde{T}_{\text{barrier}}}{\partial x_i} = -\,\text{tr}\left(\rho(x)^{-1}S_i\right), \tag{4.14}$$

$$\frac{\partial^2 \tilde{T}_{\text{barrier}}}{\partial x_i \partial x_j} = \text{tr}\left(\rho(x)^{-1}S_i\rho(x)^{-1}S_j\right). \tag{4.15}$$

Then for the maximum likelihood target function $\tilde{T}_{\text{ML}}$, one obtains [21]

$$\frac{\partial \tilde{T}_{\text{ML}}}{\partial x_i} = -\sum_k \frac{f_k}{\text{tr}(\rho(x)M_k)}\,\text{tr}(S_i M_k), \tag{4.16}$$

$$\frac{\partial^2 \tilde{T}_{\text{ML}}}{\partial x_i \partial x_j} = \sum_k \frac{f_k}{\text{tr}(\rho(x)M_k)^2}\,\text{tr}(S_i M_k)\,\text{tr}(S_j M_k), \tag{4.17}$$

while the derivatives for the free least-squares approach read as

$$\frac{\partial \tilde{T}_{\mathrm{FLS}}}{\partial x_i} = \sum_k \frac{1-f_k^2}{\mathrm{tr}(\rho(x)M_k)^2}\,\mathrm{tr}(S_i M_k), \tag{4.18}$$

$$\frac{\partial^2 \tilde{T}_{\mathrm{FLS}}}{\partial x_i \partial x_j} = \sum_k 2\frac{f_k^2}{\mathrm{tr}(\rho(x)M_k)^3}\,\mathrm{tr}(S_i M_k)\,\mathrm{tr}(S_j M_k). \tag{4.19}$$

Since all second derivatives are non-negative, the Hessian matrix fulfils $\nabla^2(\tilde{T}(\rho|f) - t\log(\det(\rho))) > 0$ for $\tilde{T} = \tilde{T}_{\mathrm{ML}}$ as well as for $\tilde{T} = \tilde{T}_{\mathrm{FLS}}$. Hence, the maximum likelihood and the free least-squares approaches are convex problems, i.e. there is only a single local minimum, which therefore coincides with the global minimum. Thus, a valid criterion for the algorithm to converge is the $p = 2$ norm of the gradient of $\tilde{T}(\rho|f) - t\log(\det(\rho))$. The smaller this absolute value is, the closer ρ is to the desired minimum [95].

A valid way to monitor the convergence behaviour of the algorithm's implementation is by applying it on simulated states. Since in this case the desired ρ is known and chosen to be a positive semi-definite matrix, the output state $\hat{\rho}_{\mathrm{ML}}$ and $\hat{\rho}_{\mathrm{FLS}}$, respectively, have to be in good agreement with the input state ρ_{theo}. A verification of this behaviour is given in Fig. 4.1, where the trace distance of the obtained state $\hat{\rho}_{\mathrm{ML}}$ with respect to the theoretical state is shown.

The barrier term $\tilde{T}_{\mathrm{barrier}} = -t\log(\det(\rho))$ equates to the additional expression of the hedged maximum likelihood approach (Eq. 4.10) for appropriate t and β. Similar as the hedged maximum likelihood target function, the barrier term punishes states with high purity. The determinant of the density matrix can be expressed as the product of its eigenvalues, i.e. $\det(\rho) = \prod_i \lambda_i$. Therefore, if $\lambda_i = 0$ holds for at least one λ_i, the determinant vanishes and the logarithm diverges. Thus, it is crucial to gradually decrease the value of t and hence the influence of the barrier term. While in the first iteration $t = 1$ the range of the influence of the barrier term is wide, its punishment becomes more peaked for decreasing t. After at most 10 iterations, t is decreased according to[3] $t \rightarrow t/10$. If the algorithm converged earlier, i.e. the gradient undershoots some threshold, t is decreased before 10 iterations.

For more details concerning the implementation of the convex optimisation and especially ideas for modifications, please consult the appendix B.

4.2.4 Bootstrapping

All described methods to reconstruct a physical density operator by means of measured frequency data are point estimates only. Therefore, each application of a reconstruction scheme results in a single density matrix without any statistical information. To overcome this, a method called *bootstrapping* is established [22, 98]. By means of Monte-Carlo simulations [99] the observed data sets are used to generate new sets of frequency data. The fluctuations of the new samples are then used to derive the error bars of the estimate. Depending on the concrete way of Monte-Carlo sampling, two kinds of bootstrapping are distinguished. After presenting the two commonly used methods of bootstrapping, later the quality of this technique will be discussed.

---

[3] Consider for example a state with $\lambda_1 = \cdots = \lambda_{15} = 10^{-5}$ and $\lambda_{16} \approx 1$. The determinant is approximately $\det(\rho) \approx (10^{-5})^{15} = 10^{-75}$. Therefore, $\log(\det(\rho)) \approx -75/\log_{10}(e)$. For $t = 1$, this term is of relevance since the likelihood is of the same order of magnitude. Later on, $t \rightarrow 10^{-10}$ and thus, the punishment by the barrier term is negligible.

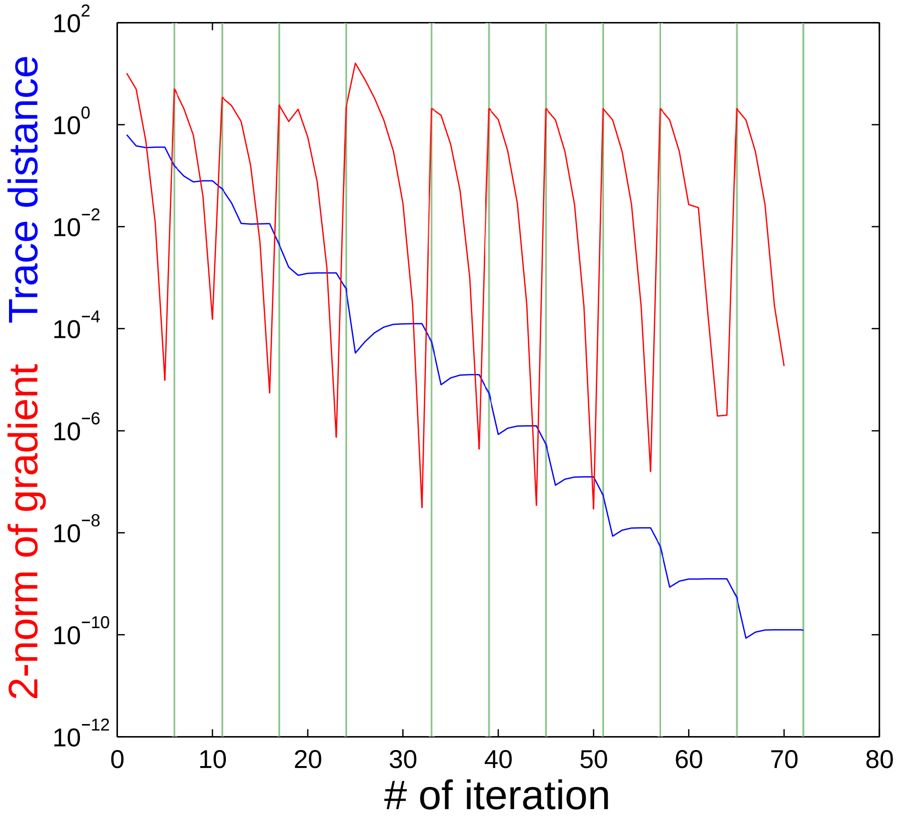

FIGURE 4.1: The logarithm of the trace distance (blue) between the physical input state $\rho_{\text{theo}} = |D_4^{(2)}\rangle\langle D_4^{(2)}|$ and the optimised state $\hat{\rho}_{\text{ML}}^{(i)}$ is shown for using the maximum likelihood estimator during the iterations of the algorithm. Each plateau of the trace distance plot corresponds to a certain value of the barrier term's coefficient t. The 2-norm of the gradient of the optimisation problem is also shown (red). At each step when the gradient undershoots a given threshold the decrease of the barrier term's coefficient t is invoked. This causes a change in the optimisation function, whereby the gradient suddenly increases again. Thus, this has to be done until the influence of the barrier term is steep, but only of short range. Here, a maximum likelihood fit was applied. Similar results were obtained for FLS.

Parametric bootstrapping

One way to obtain a kind of error bars is by applying the so called *parametric bootstrapping* [98]. First of all, the experimental data f_{obs} is used to estimate the density operator $\hat{\rho}_{\text{fit}}$ with the help of a fitting method like FLS or ML. Afterwards, the probabilities for each measurement outcome of the reconstructed state are calculated according to $P_{\hat{\rho}_{\text{fit}}}(i|S) = \text{tr}(\hat{\rho}_{\text{fit}} M_i^S)$. With these probabilities, new samples of frequency data are simulated according to a multinomial probability distribution. Afterwards, the sets of frequency data are used to reconstruct a set of density matrices. Finally, the quantity of interest can be equipped with error bars by determining the standard deviation of the respective quantity on all reconstructed density operators of all sets of simulated frequency data.
The method of simulating a quantum state based on a theoretical quantum state will be explained further in Sec. 4.4.1.

Non-Parametric bootstrapping

Besides the parametric bootstrapping, it is also possible to determine error bars of respective quantities by using the *non-parametric bootstrapping* method [98]. While the general approach equals the parametric bootstrapping, the way of generating the new samples of frequencies differs in this case. In contrast to using the probabilities of the reconstructed state like in the parametric case, the observed frequencies f_i^S are interpreted as probabilities and therefore directly used to generate new samples. Thus, no intermediate fitting takes place.

4.3 A definition of bias

An estimator is called *unbiased* if it delivers the true answer on average, i.e. its estimations are not biased towards any direction. Applying this general rule to quantum state estimation, the mean of the resulting $\hat{\rho}$ coincides with the underlying quantum state ρ_0, i.e. [100]

$$\mathbb{E}_{\rho_0}(\hat{\rho}) = \sum_f P_{\rho_0}(f)\hat{\rho}(f) = \rho_0, \tag{4.20}$$

where f denotes in this case a whole set of frequency data, i.e. $f \equiv \{f_i^S\}_{i,S}$. $P_{\rho_0}(f)$ denotes the probability that the underlying state ρ_0 provides the frequencies f. With this definition, the difference of the mean of an estimator of ρ_0 and ρ_0 itself is called *bias* [100]. If a samples of estimates is analysed , not only the bias of the estimator accounts, but also statistical fluctuations of the sample. The *mean squared error* (MSE) can be used to further investigate the different contributions to the scatter. For example, in case of the fidelity estimation, the mean squared error reads [100]

$$\mathrm{MSE(F)} = \mathbb{E}_{\rho_0}[(\hat{F} - F_0)^2] = \mathbb{V}_{\rho_0}(\hat{F}) + [\mathbb{E}_{\rho_0}(\hat{F}) - F_0]^2. \tag{4.21}$$

Thus, the mean squared error decomposes into two fundamentally different contributions, i.e. into the statistical fluctuations of the sample $\mathbb{V}_{\rho_0}(\hat{F})$ and the bias, that is the deviation of the sample's mean $\mathbb{E}_{\rho_0}(\hat{F})$ and the estimator's true mean F_0. Please note that the observation of a bias is not a very special phenomenon of quantum state estimation, rather a well known effect in statistics [100]. The probably most famous example of a bias is given by a Gaussian distribution whose variance has to be determined [101]. Assume to have obtained N samples $(X_i)_i$ from the Gaussian distribution $\mathcal{N}(\mu, \sigma^2)$ with mean μ and variance σ^2. The sample mean may be determined as $\overline{X} = \frac{1}{N}\sum_i X_i$ which directly delivers the mean μ. An intuitive estimator to obtain the sample's variance could be $\hat{s}^2 = \frac{1}{N}\sum_i (X_i - \overline{X})^2$, whose expectation value does not coincide with the underlying variance σ^2 since [102]

$$\begin{aligned}\mathbb{E}\left(\hat{s}^2\right) &= \frac{1}{N}\mathbb{E}\left(\sum_{i=1}^{N}(X_i-\mu+\mu-\overline{X})^2\right)\\ &= \frac{1}{N}\mathbb{E}\left(\sum_{i=1}^{N}(X_i-\mu)^2-2\sum_{i=1}^{N}(X_i-\mu)(\overline{X}-\mu)+N(\overline{X}-\mu)^2\right)\\ &= \frac{1}{N}\mathbb{E}\left(\sum_{i=1}^{N}(X_i-\mu)^2-N(\overline{X}-\mu)^2\right)\\ &= \mathbb{V}\left(X\right)-\mathbb{V}\left(\overline{X}\right)=\frac{N-1}{N}\sigma^2<\sigma^2\end{aligned} \tag{4.22}$$

Instead, the usually preferred estimator for the variance is obtained by correcting it according to $\hat{\sigma}^2=\hat{s}^2\frac{N}{N-1}$, which gives an unbiased estimator for the variance. The bias vanishes for infinite sample size N since $\lim_{N\to\infty}\frac{N}{N-1}=1$.

Biased estimators can lead to wrong results when not used with caution. But since the bias can in some cases be small compared to the statistical fluctuations and unbiased estimators may be hard to find, biased estimators can be employed when being aware of the problem. Here, the order of magnitude of the bias and statistical scatter will be studied, whereby the validity of the used quantum state estimators ML and FLS will be questioned.

4.3.1 Impossibility of general and unbiased estimator

By a heuristic proof, it can be shown that no quantum state reconstruction method can be designed such that on one hand all delivered quantum states are positive and on the other hand the reconstruction method is unbiased [23]. An illustrative analogue for this proposition is shown in Fig. 4.2. Due to the constraint of $\rho\geq 0$, the tail of a probability

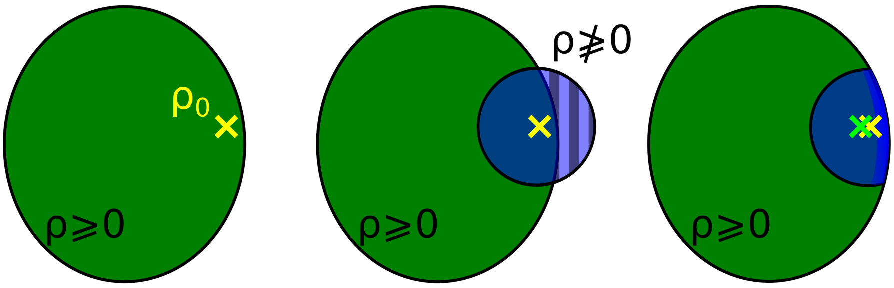

FIGURE 4.2: A physical state ρ_0 is prepared in an experimental setup (or simulated numerically). ρ_0 is shown in the left figure sketching the state space. The yellow marked state lies within the space of physical states $\rho\geq 0$ (green). When this state is used to sample with finite statistics from it, the linearly reconstructed states lie around ρ_0 as the blue shaded area suggests in the middle figure. The right part of the blue circle (blue and white stripes) does not lie within the physical space of states, consequently $\rho\geq 0$ does not hold for these states. A reconstruction method that ensures the physicality of the obtained state causes these states with $\rho\ngeq 0$ to be shifted into the area of $\rho\geq 0$. Their mean is consequently shifted. Therefore, the mean of the reconstruction using a fit $\overline{\rho_{\text{fit}}}$ (light green mark) is shifted compared to ρ_0 (yellow mark).

density distribution is cut such that the mean of the remaining sample of states is shifted relative to the initial distribution. Therefore, the restriction to $\rho \geq 0$ then necessarily leads to a bias.

Fully equivalent to this illustrative, but very heuristic argumentation, the non-existence of a general and unbiased quantum state estimator can be shown according to [23]. Consider a quantum state estimator $\hat{\rho}$ that ensures the physicality constraint of $\hat{\rho} \geq 0$. It will be shown by a proof by contradiction that no *general* quantum state estimator can be found that gives *physical* (i.e. $\hat{\rho} \geq 0$) and *unbiased* results. The unbiasedness can be express by the condition

$$\mathbb{E}_{\rho_0}(\hat{\rho}) = \sum_f P_{\rho_0, N_S}(f)\hat{\rho}(f) = \rho_0. \tag{4.23}$$

$P_{\rho_0,N_S}(f)$ denotes the probability that a theoretical quantum state ρ_0 delivers the set of frequency data f when measuring ρ_0 with N_S measurement repetitions per setting. For a finite number of measurement repetitions, $P_{\rho_0,N_S}(f) > 0$ holds for more than one[4] set of frequency data f.

Now assume that an unbiased estimator $\hat{\rho}$ with $\hat{\rho} \geq 0$ according to Eq. 4.23 would exist. The incoherent sum of reconstructed states $\hat{\rho}(f)$ based on the frequency data f has to coincide with the underlying quantum state ρ_0. For a pure state $\rho_0 = |\psi_0\rangle\langle\psi_0|$, this is only possible if the quantum state estimator delivers the underlying state $|\psi_0\rangle$ for all sets of frequency data with $P_{\rho_0,N_S}(f) > 0$, i.e. $\hat{\rho} = |\psi_0\rangle\langle\psi_0|$. Obviously, this estimator is not general. Particularly, note that for each set of frequency data f, a nondenumerable amount of states correspond to the same data set[5]. Thus, even if the state ρ_0 is reported unbiasedly by the estimator $\hat{\rho}$, another state ρ_1 cannot be determined unbiasedly if both, ρ_0 and ρ_1, correspond to the same set of frequency data when measured with N_S measurement repetitions.

Consequently, a quantum state estimator cannot be unbiased, general and always delivering a physical state at the same time when based on data sets with finite statistics.

4.3.2 Problems arising for non-linear functions

Furthermore, please be aware that even if the state estimator is unbiased this does not necessarily lead to an unbiased estimation of non-linear functions of the state. Assume that the condition of unbiasedness would hold, i.e. $\mathbb{E}_{\rho_0}(\hat{\rho}) = \sum_f P_{\rho_0}(f)\hat{\rho}(f) = \rho_0$. Now consider a function $Q \equiv Q(\rho)$ defined on the space of states. If the expectation value of this quantity shall be compared with its theoretical prediction, $Q(\rho_0) = Q\left(\sum_f P_{\rho_0}(f)\hat{\rho}(f)\right)$ directly follows for the theoretical value. If Q is a linear function,

---

[4] Only for infinite statistics ($N_S \to \infty$), $P_{\rho_0,N_S}(f) = \delta(f - f_0)$, where f_0 denotes the set of frequency data corresponding to ρ_0.

[5] Consider the simple one qubit state $|\psi_0\rangle = \sqrt{\frac{1}{\sqrt{2}}}|H\rangle + \sqrt{1 - \frac{1}{\sqrt{2}}}|V\rangle$. For example, the real probability $P_{|\psi_0\rangle}(0|Z) = \mathrm{tr}\left(M_0^Z|\psi_0\rangle\langle\psi_0|\right) = \mathrm{tr}\left(\frac{1}{2}(\sigma_0 + \sigma_z)|\psi_0\rangle\langle\psi_0|\right) = \sqrt{\frac{1}{2}}$ can be approximated by the rational frequency number f_0^Z, only. For infinite statistics the frequency f_0^Z converges to $P_{|\psi_0\rangle}(0|Z)$, i.e. $f_0^Z \to \sqrt{\frac{1}{2}}$.

this can be simplified such that

$$Q(\rho_0) = Q\left(\sum_f P_{\rho_0}(f)\hat{\rho}(f)\right) = \sum_f P_{\rho_0}(f)Q(\hat{\rho}(f)) = \mathbb{E}_{\rho_0}(Q(\hat{\rho})) \tag{4.24}$$

holds. Therefore, a linear function together with an unbiased quantum state estimator leads to unbiased results of this function. Instead, if the considered function Q is non-linear, this simplification cannot be executed. Consequently, evaluating a non-linear function on a set of unbiased quantum states does not necessarily lead to an unbiased distribution of the function values. Additionally, also if the bias of the quantum state estimator is small, non-linear functions may give strongly biased results.

4.4 Examination of state reconstruction techniques

The biasedness of quantum state estimators will be analysed by using simulated states. Thus, a theoretical state, for which the probabilities for each measurement outcome are known, is used to simulate sets of frequency data according to their distributions for a finite number of measurement repetitions. As an exemplary state, the Greenberger-Horne-Zeilinger $|\text{GHZ}\rangle$ state (see Sec. 2.5.2) encoded with 4 qubits will be considered: $|\text{GHZ}\rangle = \frac{1}{\sqrt{2}}(|0000\rangle + |1111\rangle)$. The state will be admixed with white noise according to $\rho_0 = p|\text{GHZ}\rangle\langle\text{GHZ}| + (1-p)\rho_{\text{WN}}$. This state is chosen as it first avoids problems associated with pure states and as it This simple model allows to obtain values for the fidelity that are comparable to those of experimental setups. Since a fidelity value of $F_0 = F(\rho_0, |\text{GHZ}\rangle) = 0.8$ is in the range of typically experimentally achievable fidelities, p is chosen to accomplish this, i.e. setting $p = (0.8 - F(|\text{GHZ}\rangle, \rho_{\text{WN}}))/(1 - F(|\text{GHZ}\rangle, \rho_{\text{WN}})) = 0.78\overline{6}$ results in a theoretical fidelity value of 80%. Now, the probabilities of ρ_0 are used to sample with a multinomial distribution from as will be explained in Sec. 4.4.1. One would intuitively expect that the fidelity of the quantum state after reconstruction with respect to the $|\text{GHZ}\rangle$ state is around 0.8. Even if a slight deviation might occur, at least for repeatedly doing this, the expected distribution of the fidelities of states reconstructed with FLS and ML, respectively, would fluctuate around the theoretical value of 0.8.

4.4.1 Simulation method

The fundament of the analysis are numerical simulations. For that purpose, a theoretical state ρ_0 is used. For a given measurement protocol (see Sec. 2.7) the probabilities for all projective measurements are calculated, i.e. one calculates the probabilities $P(i|S) = \text{tr}(\rho_0 M_i^S)$ for all outcomes i of all settings S with M_i^S as, for example, given in Eq. 2.78. For each setting S, these probabilities are used to sample N_S times with a given probability distribution. Here, a multinomial distribution for the number of counts c_i^S in each basis setting S is employed according to

$$P\left(c_0^S = x_0, c_1^S = x_1, \ldots, c_{2^n}^S = x_{2^n}\right) = \frac{N_S!}{x_0! \cdot x_1! \cdots x_{2^n}!} P(0|S)^{x_0} \cdot P(1|S)^{x_1} \cdots P(2^n|S)^{x_{2^n}} \tag{4.25}$$

with the number of measurement repetitions N_S. The obtained data is a set of 6^n numbers of counts c_i^S for the presented overcomplete protocol. Finally, the density matrices are reconstructed by means of the presented methods in Sec. 4.2 for each set of counts.

4.4.2 Fidelity of reconstructed Greenberger-Horne-Zeilinger states

With the presented method of simulation the maximum likelihood and the free least-squares reconstructions are compared with the direct linear inversion. For the above mentioned state ρ_0, i.e. a Greenberger-Horne-Zeilinger state with admixed white noise, one expects a fidelity with respect to the pure $|\mathrm{GHZ}\rangle$ state of $F_0 = 0.8$. The simulation is based on $N_S = 100$ measurement repetitions per basis setting S. 500 sets of frequency data are simulated that are reconstructed with the aid of the linear inversion, the maximum likelihood reconstruction and the free least-squares method.
For evaluating the fidelity from the typically unphysical states $\hat{\rho}_{\mathrm{LIN}}$ determined by means of the linear reconstruction one obtains indeed $\overline{F}_{\mathrm{LIN}} = 0.800 \pm 0.012$. However, if the states are reconstructed with maximum likelihood and free least-squares, the resulting fidelities are not distributed around 0.8 at all. Maximum likelihood reconstruction gives a significant bias and $\overline{F}_{\mathrm{ML}} = 0.788 \pm 0.010$ while reconstructing the state by means of the free least-squares method leads to even lower fidelities ($\overline{F}_{\mathrm{FLS}} = 0.749 \pm 0.010$). Although the distributions overlap, both fitting methods reduce in almost all cases the fidelity, as can be seen from the lower part of Fig. 4.3. Interestingly, while the width of ML is similar to the width of the fidelity distribution of FLS, the distribution of the difference between both fits is narrower, which is depicted in the *FLS-ML* distribution. The narrow distribution of *FLS-ML* indicates a stronger correlation between the reported states via maximum likelihood and free least-squares than between the outcomes of *LIN* and *ML* on one hand and *LIN* and *FLS* on the other hand.

As one would intuitively suppose, the bias of the reconstruction methods depends on the statistics, i.e. the number of measurement repetitions N_S per basis setting S. This bias is neither special for the GHZ state nor for estimating the fidelity. Instead, in Sec. 4.4.3 an overview of the deviations for different states and statistics N_S will be given. Firstly, for $N_S = 100$ state parameters of the Greenberger-Horne-Zeilinger state will be varied. Afterwards, the dependency of the bias of different states on the statistics N_S will be investigated.

4.4.3 Different state parameters

Although the deviations between the reconstructed and the theoretical state were shown for the 4 qubit Greenberger-Horne-Zeilinger (GHZ) state with $21.\overline{3}\%$ admixed white noise, this effect is not restricted to this case. In this section different parameters of the initial state are varied to further study the bias of the state estimator.

Fidelity bias for different GHZ states

Here, different types of Greenberger-Horne-Zeilinger states are considered by either varying the number of qubits or the amount of admixed white noise. The theoretical state belongs to the class of states $\rho_0 = p|\mathrm{GHZ}_n\rangle\langle\mathrm{GHZ}_n| + (1-p)\rho_{\mathrm{WN}_n}$, where p is chosen such that the fidelity between the theoretical state ρ_0 and the target state $|\mathrm{GHZ}_n\rangle$ fulfils

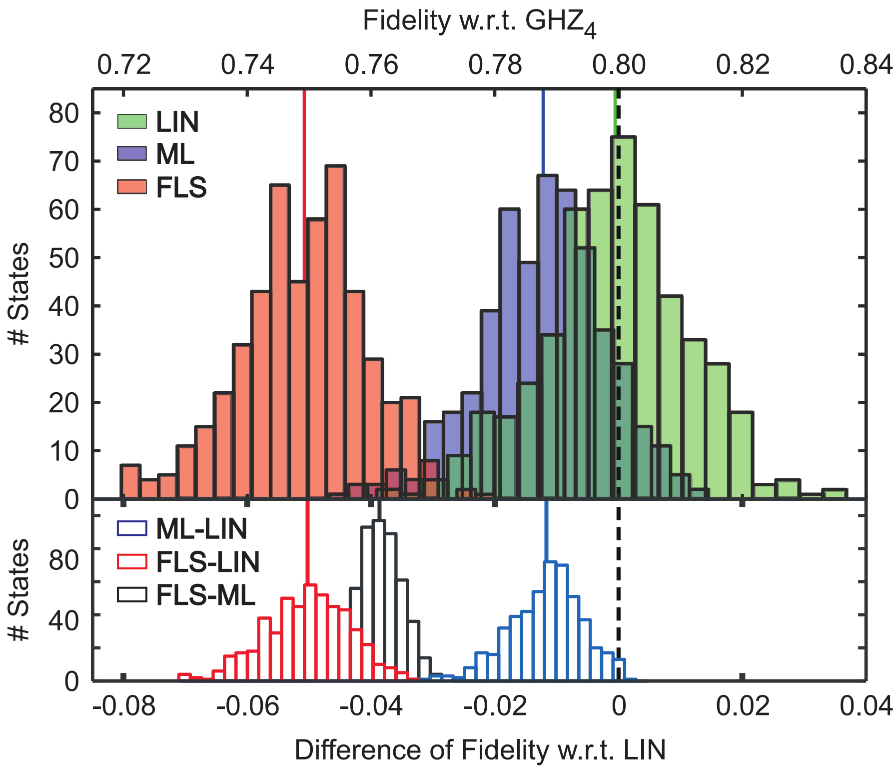

FIGURE 4.3: 500 states are sampled with a multinomially distributed number of counts per outcome based on the simulated frequencies (see Sec. 4.4.1) of the theoretical state ρ_0 which is composed of a four qubit GHZ state with admixed white noise. $N_S = 100$ measurements are performed per setting S. While the average fidelity using linear inversion (LIN) is close to the true value of 0.8, which is depicted by the dashed line, the mean values for free least-squares (FLS) and maximum likelihood (ML) state reconstruction deviate from it. The lower part of the figure shows the distribution of the deviations for each single state for the different reconstruction methods such as *ML-LIN* which corresponds to the difference of the fidelity of the state reconstruction using maximum likelihood and the linear estimated fidelity.

$F_0 = F(\rho_0, |\mathrm{GHZ}_n\rangle) = 0.8$. While for increasing number of qubits n the fidelity decreases only slightly for the maximum likelihood reconstructed states, a stronger decrease for FLS can be seen, cf. Fig. 4.4. The standard deviation of the fidelity of the studied 500 states decreases for increasing numbers of qubits n, although the total amount of repetitions per basis setting is always set to $N_S = 100$. This effect can be explained by the fact that the fidelity is composed of a larger quantity of different numbers of counts c_i^S for increasing number of qubits and basis settings. This reduces the statistical error.

Additionally, one is also interested in the dependency of the reported fidelity on the purity of the inital state. Therefore, the amount of admixed white noise will be varied such that target fidelities F_0 between 1/16 (only white noise without any amount of a GHZ state) and 0.9 are used to generate again 500 states for each ρ_0. Recall from the introduction in Sec. 4.4 that the parameter p is given by

$$p = \frac{F_0 - F(|\mathrm{GHZ}\rangle, \rho_{\mathrm{WN}})}{1 - F(|\mathrm{GHZ}\rangle, \rho_{\mathrm{WN}})}. \tag{4.26}$$

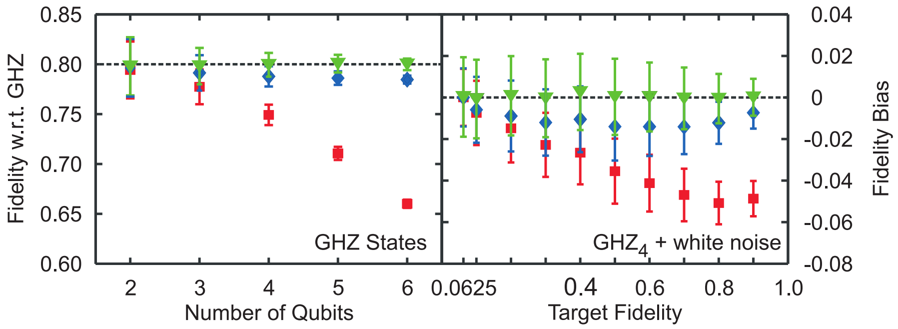

Figure 4.4: The fidelity dependency on the number of qubits (left subfigure) and on the target fidelity (right subfigure) is shown. The right subfigure depicts the fidelity bias $F(\rho_{\text{estimation}}, |\text{GHZ}\rangle) - F(\rho_0, |\text{GHZ}\rangle)$ for different initial fidelities F_0. The analysis is based on 500 simulated states (see Sec. 4.4.1) and $N_S = 100$ measurement repetitions per basis setting S.

While for low target fidelities the bias almost vanishes, its influence is largest for the most interesting regime of $F_0 \in [0.5; 0.9]$. The largest deviation between F_0 and $\overline{F}_{\text{FLS}}$ is found for higher target fidelities than in the case of using the maximum likelihood method. For all chosen F_0, FLS gives more biased fidelities than ML, whose results are shifted from F_0, anyway.

Fidelity bias for different classes of states

In Fig. 4.5 different classes of 4 qubit states are investigated with respect to the fidelity between the theoretical state and the state obtained with either one of the fit methods ML and FLS or with direct evaluation of the linearly obtained state. Therefore, the fidelity which again is expected to be $F_0 = 0.8$ is shown for a GHZ state (Sec. 2.5.2), the Smolin state (Sec. 2.5.7), a Dicke state (Sec. 2.5.3), a W state (Sec. 2.5.4), a Cluster state (Sec. 2.5.6) and a separable state ($\propto (|H\rangle + |P\rangle)^{\otimes 4}$) for a varying number of simulated measurement repetitions N_S per setting S. The mentioned effect is not negligible for any of the states in the regime of a typical number of measurement repetitions while for high values of N_S this effect becomes almost indiscernible. The state reconstruction with FLS gives in all cases worse results than using ML which still suffers from a significant bias. In contrast, the linearly evaluated fidelity is a good estimate since it fluctuates around the true value of $F_0 = 0.8$ for all values of N_S. Please note that the fidelity can only be evaluated linearly if it is evaluated with respect to a pure state. Because the measured (in this case simulated) state is in general a mixed state, a linear evaluation can only be accomplished if the theoretical state is pure. Therefore, a linear evaluation of the fidelity is not possible, e.g., for the Smolin state as it is a mixed state [73].

4.4.4 Error calculation

Because quantum state estimators like the maximum likelihood or free least-squares methods are point-estimators, no information about the statistical behaviour can be inferred from a single estimation. Therefore, an often used method to learn about the standard deviation of quantities of interest is to use bootstrapping as explained in Sec.

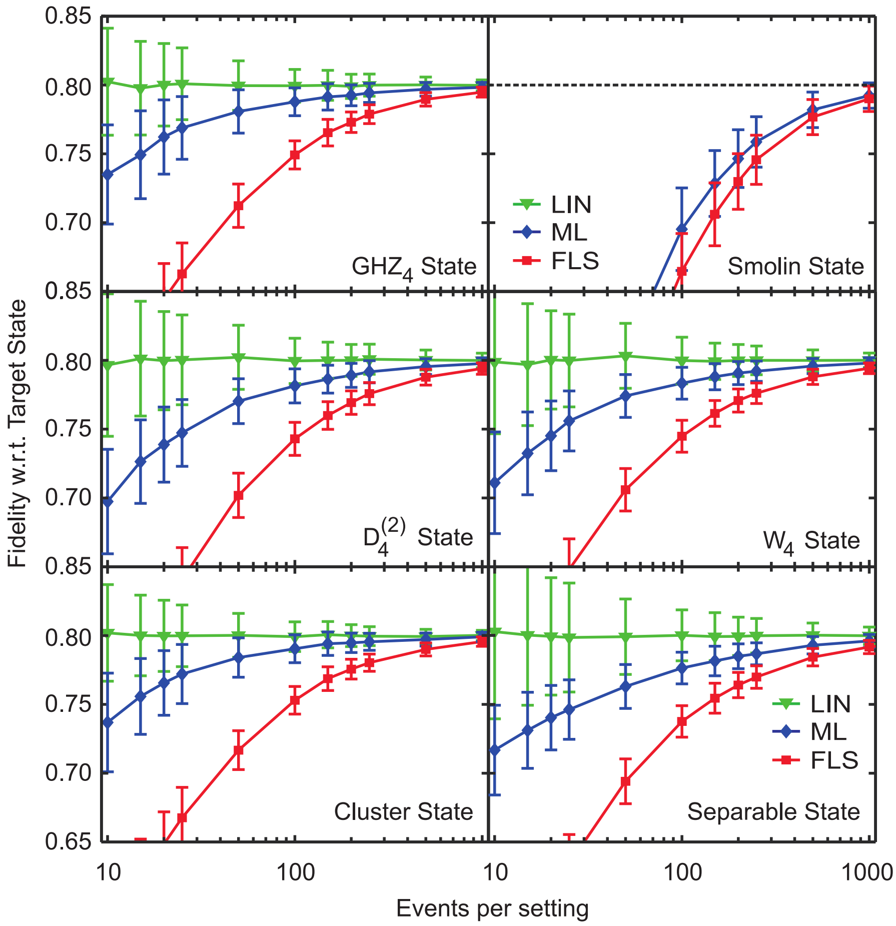

FIGURE 4.5: The dependency of the fidelity on the number of measurement repetitions N_S per basis setting for various states is shown. The decrease of the fidelity with respect to the theoretical state cannot only be seen for GHZ states, but is also observable for all studied states like the Smolin state, the Dicke state $|D_4^{(2)}\rangle$, the $|\mathrm{W}\rangle$ state, the Cluster state and a separable state, each considered with $n = 4$ qubits. Estimating the fidelity of reconstructed states via a fitting method underestimates the true value of $F_0 = 0.8$. In the case of the Smolin state no linearly estimated fidelity could be obtained since if both, the studied state and the reference state, are mixed, the fidelity is not a linear function, cf. Eq. 2.33.

4.2.4. For the case of a Greenberger-Horne-Zeilinger state with admixed white noise ($F_0 = F(\rho_0, |\mathrm{GHZ}\rangle) = 0.8$) bootstrapping will be used to evaluate the quality of the obtained error bars. In Fig. 4.6 bootstrapping is applied for all four combinations of maximum likelihood and free least-squares estimation as well as for parametric and non-parametric bootstrapping. The distributions are based on the same 500 simulated states as in Fig. 4.3. For each obtained estimate 100 new frequency data are generated from the reconstructed state (parametric bootstrapping) and from the original frequency data directly (non-parametric bootstrapping). These new generated 100 sets of frequency data for each state are used to compute their respective mean and standard deviation. In Fig. 4.6 the histogram shows the distribution of the mean values of the bootstrapping results as well as their standard deviation (grey shaded area).

In the most drastic case, as shown in the lower left part of Fig. 4.3, the determined value deviates significantly from the theoretical value. Since the fidelity is reduced

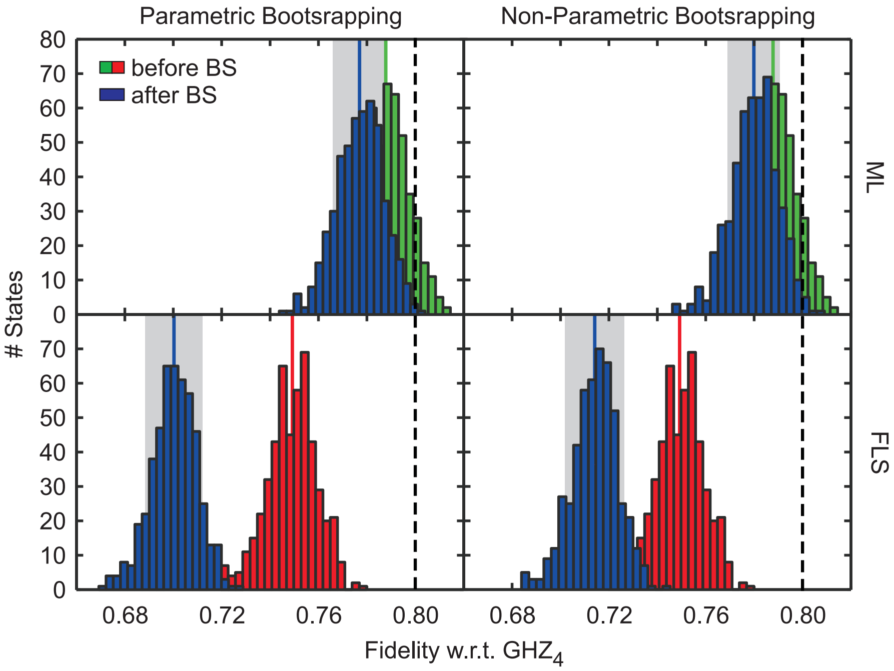

FIGURE 4.6: Applying bootstrapping is a standard method for determining the standard deviation of quantities of interest. Here, bootstrapping is used to find the standard deviation of the fidelity for the 4 qubit GHZ state with admixed white noise. The histograms denoted by "before BS" represent the fidelity estimates for the states obtained by ML and FLS fitting methods, respectively. In the case of parametric bootstrapping, each of these 500 states are used to generate 100 new sets of frequency data, which are again used for the fitting method. The 500 means and standard deviations, each taken over the set of these 100 newly sampled data, are shown in the histogram labelled with "after BS" as the continuous blue lines and the grey shaded areas, respectively. In the non-parametric case, the seed for the bootstrapping corresponds to the linear inverted states, cf. the distribution denoted by "LIN" in Fig. 4.3.

by using a fitting result for estimating it, this effect occurs twice for the parametric bootstrapping. Neither the theoretical value of $F_0 = 0.8$ is within one standard deviation of the reported fidelity ($F_{\mathrm{FLS}}^{\mathrm{para}} = 0.700 \pm 0.012$) nor the estimated fidelity for the initial FLS fit. None of the bootstrapping methods reports a result which is in agreement with the theoretical fidelity of 80% since $F_{\mathrm{FLS}}^{\mathrm{para}} = 0.700 \pm 0.012$, $F_{\mathrm{FLS}}^{\mathrm{non-para}} = 0.714 \pm 0.012$, $F_{\mathrm{ML}}^{\mathrm{para}} = 0.777 \pm 0.011$ and $F_{\mathrm{ML}}^{\mathrm{non-para}} = 0.780 \pm 0.011$.
Obviously, the reported standard deviation is of the right order of magnitude for the fidelity as can be seen in Fig. 4.3. The standard deviation of the linear inversion (see Sec. 4.4.2) is about 0.012. A very similar value can be obtained by means of the bootstrapping method (0.011 for maximum likelihood and 0.012 for free least-squares). Nevertheless, the estimated error bars are not evaluated for the state one is interested in, but instead for a biased state. Furthermore, the true value lies not within the interval given by the mean of the estimate and the error bars obtained by the standard deviation of the bootstrapping. Consequently, bootstrapping is a highly questionable method to obtain error bars when using an estimator with a significant bias as shown for ML and FLS state estimators. Although the error bars can be roughly estimated in this example for the linear fidelity, the error bars of non-linear function may not be determined correctly. In [23], the bootstrapping method of the Fisher information [103] is analysed. There, one learns that the standard deviation obtained via bootstrapping can be far too small.

In this thesis, another method to find statistical information will be presented, that is also applicable for non-linear functions like the Fisher information.

4.5 Linear evaluation

Since quantities like the von Neumann entropy or the purity cannot be evaluated on states with $\rho \ngeq 0$, fitting methods like ML and FLS were applied, as they ensure the physicality of the obtained state. In the previous section it could be shown that these estimators suffer from a significant bias. Therefore, one has to decide whether the statistics is large enough such that the bias is negligible or whether fitting cannot be applied at all.

In the previously shown figures the fidelity was employed as a measure of the reconstructed state. While only in the case of the Smolin state comparing the in general unphysical state $\hat{\rho}_{\mathrm{LIN}}$ with its theoretical state was not possible, cf. Fig. 4.5, in all other cases the fidelity could have been evaluated linearly, as the fulfilment of the condition $\rho \geq 0$ was not required. In this section, a method will be presented how to evaluate non-linear, but convex or concave quantities of interest by using linear approximations to the functions, i.e. directly from the data without employing intermediate fitting methods like maximum likelihood or free least-squares. Furthermore, a way how to estimate (conservative) confidence regions for the quantities will be explained.

4.5.1 Linear approximation of convex & concave quantities

Evaluating quantities of interest for unphysical "states" gives results without any information value. For example, consider the entropy $S(\rho) = -\operatorname{tr}(\rho \log(\rho))$ which might give imaginary values for $\rho \ngeq 0$. A possible approach to overcome this problem is to linearly approximate non-linear functions. For a convex function a lower bound can be determined by an appropriately chosen linear function, as well as all concave functions can be upper bounded by straight lines (or in general by hyperplanes for multivariate functions). Luckily, most of the interesting functions are either convex or concave functions. Here, an overview of some of these functions will be given together with their derivatives and a recipe, how to compute confidence regions for these quantities.

Fig. 4.7 illustrates the linear approximation for an univariate case. Consider a convex function $Q(x)$, i.e. $Q(\alpha x_1 + \beta x_2) \leq \alpha Q(x_1) + \beta Q(y_2)$ for $\alpha \in [0, 1]$ and $\beta = 1 - \alpha$. A lower bound for $Q(x)$ can be determined by finding the tangent in any expansion point x_0. Therefore, $Q(x) \geq Q(x_1) + Q'(x_1)(x_0 - x_1)$ holds $\forall x_1$ [101]. Fig. 4.7 already clarifies that the inequality holds for all choices of x_1 while the inequality is more tight for some expansions than for others. The linear approximation by evaluating in x_1 gives a valid expression (blue line), but building up the inequality with x_2 may be the better choice (red line). The value of the linearisation around x_2, evaluated at x, is closer to the function value $Q(x)$ than if one used the linearisation around x_1. The previous explanation for $x \in \mathbb{R}$ can be generalised to be used with density matrices. Therefore, with an adequate parametrisation of $\rho \equiv \rho(x)$ linearly determined bounds for $Q[\rho(x)]$

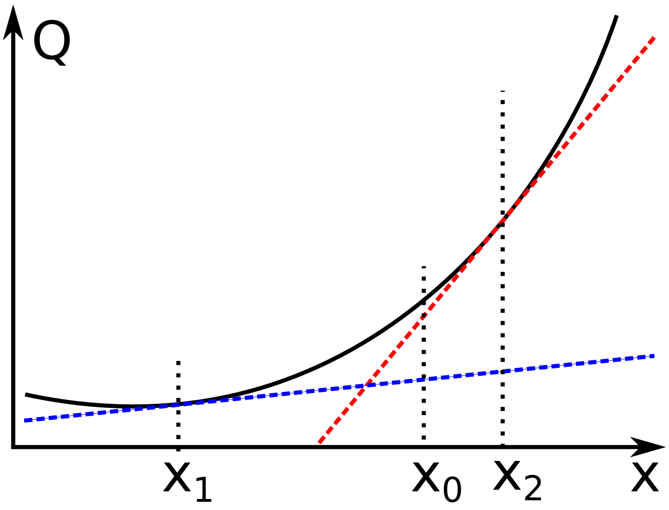

FIGURE 4.7: Convex functions can be bounded from below by linear functions. Irrespective of the evaluation point, a valid inequality $Q(x_0) \geq Q(x_1) + Q'(x_1)(x_0 - x_1)$ can be found, while some choices are better than others since their informative value is larger. Linear approximation in x_1 gives a valid information about $Q(x_0)$ while one retrieves more information by using a more appropriate evaluation point as x_2, as depicted by the red line. The value of the red approximation evaluated in x_0 is closer to $Q(x_0)$ than when using the blue approximation.

can be found, i.e.

$$Q_1[\rho(x)] \geq Q_1[\rho(x_0)] + \nabla Q_1[\rho(x_0)]^T(x - x_0) = \mathrm{tr}(L_1\rho(x)) \tag{4.27}$$

$$Q_2[\rho(x)] \leq Q_2[\rho(x_0)] + \nabla Q_2[\rho(x_0)]^T(x - x_0) = \mathrm{tr}(L_2\rho(x)) \tag{4.28}$$

for a convex quantity Q_1 and a concave quantity Q_2, respectively. Corresponding to the discussed univariate case, the quality, i.e. the information value, of the approximation of a function of density operators depends on the state that the function was expanded around. For practical usage a linear operator L has to be determined based on linear approximation. While L can be found by expanding around different states, i.e. a L_a by expanding around $\rho(y_a)$ and L_b by using $\rho(y_b)$, it is an essential step to find a proper operator L. For example getting to know that $\mathrm{tr}\left[\rho(x)^2\right] \geq \frac{1}{2^n}$ for a n qubit state, is unquestionably a veritable statement but without any information value since $\mathrm{tr}[\rho^2] \in [1/2^n; 1]$ holds for all n qubit states ρ, anyway. When another linear approximation would deliver a larger value for the lower bound of the purity, finding this more informative L is the current task. There are different approaches to find the best suited state to expand around. One might choose the theoretical or the expected state as $\rho(y)$. But since the experimentally measured state could strongly differ from this - or the theoretical state is even completely unknown - one might also use fitting approaches based on the measured data. Thus, the linearisation can for instance be built with $\rho(y) = \hat{\rho}_{\mathrm{ML}}$, where $\hat{\rho}_{\mathrm{ML}}$ may be obtained via the maximum likelihood fitting from the linearly reconstructed state. Note that $\rho(y)$ must fulfil $\rho(y) \geq 0$ as will be explained later. To find the best, i.e. tightest bound, one could also optimise over all physical states $\rho(y)$ and determine the extremal value (maximal value for a convex function, minimal value for a concave function) for the linear approximation. It is still an open question what the best strategy to find the linear operator L is.

Function	$Q(x)$
Concave	
Entropy (see 2.4.6)	$-\operatorname{tr}(\rho\log\rho)$
Uhlmann fidelity (see 2.4.3)	$(\operatorname{tr}[\sqrt{\sqrt{\sigma}\rho\sqrt{\sigma}}])^2$
Convex	
Purity (see 2.4.5)	$\operatorname{tr}(\rho^2)$
Fisher information [103]	$2\sum_{ij}\frac{(\lambda_i-\lambda_j)^2}{\lambda_i+\lambda_j}H_{ij}H_{ji}$
Negativity (see 2.4.1)	$(\sum_i \lvert\lambda(\rho_{AB}^{T_A})_i\rvert - 1)/2 = \sum_{\lambda(\cdot)<0}\lvert\lambda(\rho_{AB}^{T_A})\rvert$

Table 4.1: Convex and concave functions can be linearly evaluated to estimate the quantity also for non-physical density matrices. Note that lower bounds for the purity, the Fisher information or the negativity are of avail as well as an upper bound for the entropy. In contrast, in most cases one would be more interested in a lower bound for the fidelity instead of bounding it from above. Due to its concavity, this is unfortunately not achievable. The logarithm for evaluating the entropy is used to base two [5]. The λ_i in the definition of the Fisher information denotes the ith eigenvalue of ρ as well as $|\psi_i\rangle$ labels the corresponding eigenstate while H_{ij} identifies $\langle\psi_i|H|\psi_j\rangle$ with the operator H to compute the Fisher information.

In Tab. 4.1 an overview of important quantities of interest is given. For concave functions like the entropy or the fidelity, upper bounds can be found by using also states with $\rho \not\geq 0$. Convex functions like the purity, the Fisher information or the negativity can be bounded by below. In the following, the method how to compute the linear approximation of these quantities will be explained.

Composing linear operators

An appropriate parametrisation of the states is again given by $\rho = \mathbb{I}/2^n + \sum_i x_i S_i$ which is equivalent to the parametrisation in Eq. 4.12 and to the decomposition of ρ into a correlation tensor (Eq. 2.19), except for the factor of $1/2^n$ to be absorbed now into the x_i. In principle, all choices for a basis $\{S_i\}_i$ with $\operatorname{tr}(S_iS_j) = \delta_{i,j}$ are valid. Here, using the Pauli matrices as a basis seems the most appropriate option and will be used in the following. For a quantity $Q(x) = Q[\rho(x)]$ being convex in x and with $\rho(x) \geq 0$ a lower bound will be determined. For that purpose, one has to find a state $\rho_1(x_0)$ with $\rho_1(x_0) \geq 0$ where a Taylor expansion [101] up to first order of the quantity has to be constructed. Then, for the state ρ_0 obtained by linear inversion, for which $\rho_0 \geq 0$ may not necessarily hold, an inequality can be constructed, i.e.

$$Q[\rho_0] \geq \operatorname{tr}[\rho_0 L(\rho_1(x_0))]. \tag{4.29}$$

Here, the linear operator $L(\rho_1(x_0))$ decomposes into

$$L(\rho_1(x_0)) = l_0\mathbb{I} + \sum_i l_i S_i, \tag{4.30}$$

Function	$\partial_i Q(x)$
Upper bounds	
Entropy	$-\operatorname{tr}[S_i(\log\rho-\mathbb{I})]$
Uhlmann fidelity	$2\sqrt{Q(x)}\sum_j \frac{1}{2\sqrt{\mu_j}}\langle w_j\vert\sqrt{\sigma}S_i\sqrt{\sigma}\vert w_j\rangle$
Lower bounds	
Purity	$2\operatorname{tr}(S_i\rho)$
Fisher information	$4\sum_{ijk}\frac{\lambda_i\lambda_j+\lambda_i\lambda_k+\lambda_j\lambda_k-3\lambda_i^2}{(\lambda_i+\lambda_j)(\lambda_i+\lambda_k)}H_{ij}S_{jk}H_{ki}$
Negativity	$-\operatorname{tr}(\rho_{AB}\Pi^{T_A})$

Table 4.2: The partial derivatives of the functions listed in Tab. 4.1. The concave quantities (entropy and fidelity) can be bounded from above, while for the convex measures (purity, Fisher information and negativity) lower bounds can be estimated. In the formula of the fidelity, μ_i and $|w_i\rangle$ label the eigenvalues and the associated eigenvectors of $\sqrt{\sigma}\rho\sqrt{\sigma}$; σ belongs to the target state that ρ is compared with. For the negativity the projector Π^{T_A} onto the eigenspace belonging to the negative eigenvalues of $\rho_{AB}^{T_A}$ has to be computed.

where the coefficients l_0 and l_i are determined by

$$l_0 = Q[\rho_1(x_0)] - \sum_i y_i \frac{\partial}{\partial x_i} Q(x_0), \tag{4.31}$$

$$l_i = \frac{\partial}{\partial x_i} Q(x_0). \tag{4.32}$$

Since the partial derivatives of the quantity are needed to be computed, a list of those derivatives[6] for the considered quantities is given in Tab. 4.2. With this scheme it is possible to estimate or at least bound quantities also for non-physical states. This procedure will be further illustrated by an example.

Consider now the convex function of the purity $P(\rho)=\operatorname{tr}\left(\rho^2\right)$. Using the parametrisation $\rho=\mathbb{I}/2^n+\sum_j x_j S_j$, the purity becomes $P(\rho)=\operatorname{tr}\left(\rho^2\right)=\operatorname{tr}\left(\left(\mathbb{I}/2^n+\sum_j x_j S_j\right)^2\right)$. The partial derivative of the function with respect to x_i is needed for building up the linear operator. Thus,

$$\frac{\partial P(\rho(x))}{\partial x_i} = \operatorname{tr}\left(\frac{\partial\left(\mathbb{I}/2^n+\sum_j x_j S_j\right)^2}{\partial x_i}\right) = \operatorname{tr}\left(2\left(\mathbb{I}/2^n+\sum_j x_j S_j\right)S_i\right) = 2\operatorname{tr}(\rho S_i). \tag{4.33}$$

Then, one has to find a state $\rho_1(y)$ to construct the linear approximation. This state has to fulfil the physicality constraint of $\rho_1(y)\geq 0$ since the purity $P(\rho_1(y))$ as a non-linear

---

[6] Tobias Moroder. Unpublished.

function will be evaluated for this state. Thus, the linear operator becomes

$$\begin{aligned} L(\rho_1(y)) =& l_0\mathbb{I} + \sum_i l_i S_i \\ =& \mathbb{I}\left(\operatorname{tr}\left(\rho_1(y)^2\right) - 2\sum_i y_i \operatorname{tr}(\rho_1(y)S_i)\right) + 2\sum_i \operatorname{tr}(\rho_1(y)S_i)S_i \\ =& \mathbb{I}\operatorname{tr}\left(\rho_1(y)^2\right) + 2\sum_i \operatorname{tr}(\rho_1(y)S_i)(S_i - y_i\mathbb{I}). \end{aligned} \tag{4.34}$$

Although Eq. 4.34 looks rather lengthy, it fully corresponds to a Taylor expansion up to the first order of $\operatorname{tr}(\rho^2)$ when evaluating around $\rho = \rho_1(y)$. Now, this operator can be used to find a lower bound for the purity of another state $\rho(x)$ by calculating $\operatorname{tr}(\rho(x)L(\rho_1(y))$. Note that $\rho(x) \geq 0$ does not necessarily need to hold because this expression corresponds to a linear evaluation where possible unphysicalities do not cause any problems.

Before examples of this method will be presented, the calculation of error bars will be discussed.

4.5.2 Error calculation

Although it is possible to determine the error of each matrix element of the linearly estimated density matrix by using Gaussian error propagation, these error bars are not very useful when one is interested in single quantities like the fidelity. In this case, directly reporting the error of the quantity, that one is interested in, is of more avail. Firstly, this section argues why error bars in the state space are not most appropriate. Instead, using the errors of the obtained measurement data by means of Gaussian error propagation to compute the error of a quantity is more appropriate. Unfortunately, this cannot be done for non-linear quantities. To overcome this, the error calculation based on the Hoeffding inequality [24] will be presented. With this calculus one finds an expression for estimating the confidence intervals of the functions directly.

Error bars in state space

Assume that a confidence region C_ρ in the state space can be used to easily find the confidence region C_Q for the quantity of interest Q such that the confidence interval of Q is for example given by

$$\mathrm{C}_Q = [\min_{\rho\in C_\rho} Q(\rho); \max_{\rho\in C_\rho} Q(\rho)], \tag{4.35}$$

i.e. one tries to find the minimal and maximal values for Q for all states ρ that are within the confidence region C_ρ. The illustrative example of [23] will be recapitulated here to emphasise why this way of constructing confidence intervals is not the best choice. Suppose a probability density function $f(\vec{x})$ with $\vec{x} \in \mathbb{R}^2$ is given where

$$f(\vec{x}) = \frac{1}{2\pi\sqrt{\det(\Sigma)}} \exp\left(-\frac{1}{2}(\vec{x}-\vec{\mu})^T\Sigma^{-1}(\vec{x}-\vec{\mu})\right) \tag{4.36}$$

corresponds to a two dimensional normal distribution with the positive-definite covariance matrix $\Sigma \in \mathbb{R}^{2\times 2}$ and the mean $\vec{\mu} \in \mathbb{R}^2$. For simplicity, assume $\Sigma = \sigma \mathbb{I}_{2\times 2}$ and without loss of generality $\vec{\mu} = \vec{0}$. The confidence region for a sample according to this distribution can be build up in different ways. The probability density function of Eq. 4.36 is shown in Fig. 4.8 in the left plot. One might want to construct a symmetric confidence region for this function as shown by the red cylindrical shape. The probability value obtained by integrating over the enclosed area of the red shape is the same as if one chooses a confidence region like the green confidence region. While both choices are valid for one and the same confidence level, the regions look rather different. If one asks for the confidence region of the function of the projection onto the x direction, the green choice is smaller than the symmetric, red one. Therefore, recall the Gaussian probability density function of Eq. 4.36. If one wants to find a confidence region corresponding to one standard deviation, the confidence region shall contain 68.3% of the possible results. Since the probability density is normalised, it must hold $\int_C f(\vec{x})d\vec{x} = 0.683$. Choosing a symmetric region, one obtains

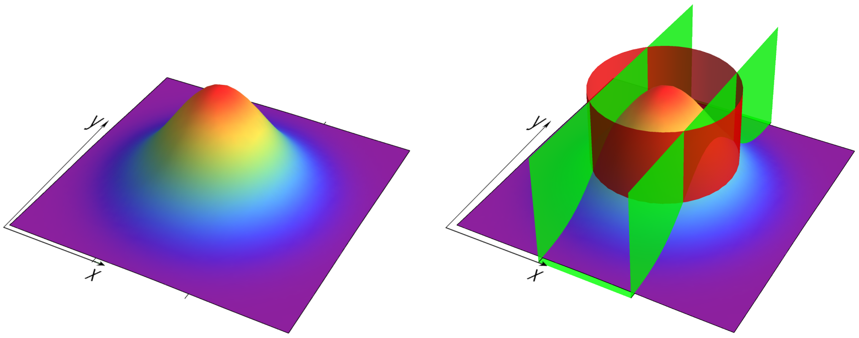

FIGURE 4.8: The probability density function of Eq. 4.36 as a three dimensional plot (left). Two exemplary choices for confidence regions are shown and finally compared for this function (right). The green confidence region $C^{(1)} = [-\sigma, \sigma] \times [-\infty, \infty]$ gives the smallest possible spread in x direction, while being maximally extended along y. The red choice in contrast is chosen symmetrically with $C^{(2)} = \{\sqrt{x^2+y^2} \leq 1.52\sigma\}$ [23]. This schematically shown behaviour becomes even more dramatical for more dimensions.

$$0.683 = \int_0^{2\pi} \mathrm{d}\phi \int_0^R \mathrm{d}r r \frac{1}{2\pi\sigma} \exp\left(-\frac{1}{2}\frac{r^2}{\sigma^2}\right) = 1 - \exp\left(-\frac{1}{2}\frac{R^2}{\sigma^2}\right) \Rightarrow R \approx 1.52\sigma. \quad (4.37)$$

Consequently, the symmetric choice of the confidence region is not the best option if one is interested only in a single value. Thus, the confidence region becomes smallest if especially defined for the quantity one is interested in. Consider that one is interested in the projection in x direction, only. Then,

$$0.683 = \int_{-x_0}^{x_0} \mathrm{d}x \int_{-\infty}^{\infty} \mathrm{d}y \frac{1}{2\pi\sigma} \exp\left(-\frac{1}{2}\frac{x^2+y^2}{\sigma^2}\right) \Rightarrow x_0 = \sigma. \quad (4.38)$$

If one wants to give confidence regions for the density matrix itself, these regions become huge while the confidence region of for example the fidelity might be much smaller. Naturally, the error bars of a linear measure like the fidelity can be determined directly by means of Gaussian error propagation. How error bars for linearly approximated non-linear functions can be determined, will be explained in the following section.

Direct calculation of confidence intervals

For example, consider the bipartite negativity (Sec. 2.4.1), that depends on the eigenvalues of the partial transpose of the state. It is difficult to express the bipartite negativity in terms of the measured count data and, therefore, to estimate the error bars by means of Gaussian error propagation. In contrast, the error information is obtained by means of the linear approximation of the non-linear function. A direct way to find a confidence region for the quantities of interest is given by applying Hoeffding's inequality [24]. Consider n independent real random variables X_i bounded within a certain interval, i.e. $X_i \in [a, b]\ \forall X_i$. Hoeffding derived an inequality to bound the probability that the sum of the random variables deviates by a certain amount from its respective expectation value [24]. For that purpose, denote the sum of the random variables with $S = \sum_i X_i$. Then [104]

$$P(|S-\overline{S}| \geq \alpha) \leq 2\exp\left(-\frac{2\alpha^2}{\sum_{i=1}^{n}(b_i-a_i)^2}\right) = \delta, \tag{4.39}$$

where $P(|S-\overline{S}| \geq \alpha)$ denotes the probability that S deviates more than α from its mean $\overline{S}$. While with Eq. 4.39 the probability that a certain deviation occurs can be computed, it is more interesting to find the confidence region belonging to a certain probability. Therefore, one is interested in an expression for $\alpha(\delta)$ describing the maximal deviation α for a given confidence level δ:

$$\alpha(\delta) = \sqrt{-\frac{1}{2}\sum_{i=1}^{n}(b_i-a_i)^2 \ln\left(\frac{\delta}{2}\right)}. \tag{4.40}$$

If one is interested in for example a 90% confidence region, the probability that the outcome deviates more than this region should be less than 10%, thus $\delta = 0.10$. Hence, δ may be replaced by $1-\gamma$, where γ labels the confidence level.

Regard now a linear function[7] $Q_L(\rho_0) = \mathrm{tr}(L\rho_0)$. A confidence region is given by Eq. 4.39 and Eq. 4.40 for $P\left(Q_L(\rho_L) \in [Q_L(\rho_0) - C_L^{2s}; Q_L(\rho_0) + C_L^{2s}]\right) \geq \gamma$ with

$$C_L^{2s} = \frac{\alpha}{\sqrt{N_S}} = \sqrt{-\frac{h^2\ln((1-\gamma)/2)}{2N_S}}. \tag{4.41}$$

Here, $h^2 = \sum_{i=1}^{n}(b_i-a_i)^2 = \sum_S(\max_i l_i^S - \min_i l_i^S)^2$. Since $\max_i l_i^S$ $(\min_i l_i^S)$ denotes the value of the linear operator that can be achieved maximally (minimally) in basis S. Therefore, the linear operator has to be decomposed into its measurement projectors[8]. The number of repetitions N_S is assumed to be equal for all settings S. Since the former approach is based on linearising convex or concave function, one does only get a bound by

---

[7] $Q_L(\rho_0) = \mathrm{tr}(L\rho_0) = \mathrm{tr}\left(\sum_{i,S} P(i|S)M_i^S L\right) = \sum_{i,S} P(i|S)\,\mathrm{tr}(M_i^S L) = \sum_{i,S} P(i|S) l_i^S$.

[8] The linear operator L can be parametrised in different ways:

$$L = \underbrace{l_0\mathbb{I} + \sum_i l_i S_i}_{\text{Decomposition into e.g. Pauli basis}} = \underbrace{\sum_{i,S} l_i^S M_i^S}_{\text{Decomposition into measurement projectors}}, \tag{4.42}$$

where the first decomposition $L = l_0\mathbb{I} + \sum_i l_i S_i$ corresponds to the construction of the linear operator as given in Eq. 4.30. The second identity, i.e. $L = \sum_{i,S} l_i^S M_i^S$ is the parametrisation of L in terms of the measurement projectors M_i^S. Note that the coefficients l_i in the first decomposition are distinct from l_i^S in the second one.

an inequality. Therefore, two-sided confidence intervals are not needed. Instead, a more meaningful ansatz leads to a one-sided confidence region weakening the inequality. With this, one finally obtains an expression for the probability that the *true* value $Q_L(\rho_0)$ of a convex (concave) function exceeds (undershoots) a threshold determined by the data and afterwards relaxed by an error bound. Thus, $P_{\rho_0}[Q_L(\rho_0) \geq Q(\hat{\rho}_{\mathrm{LIN}}(f)) - C_L] \geq \gamma$ holds for convex functions, while concave functions fulfil $P_{\rho_0}[Q_L(\rho_0) \leq Q(\hat{\rho}_{\mathrm{LIN}}(f)) + C_L] \geq \gamma$. Here, C_L is given by

$$C_L \equiv C_L^{1s} = \sqrt{-\frac{h^2 \ln(1-\gamma)}{2N_S}}. \tag{4.43}$$

Note the difference of the factor 2 in the equations for two-sided C_L^{2s} (Eq. 4.41) and one-sided C_L^{1s} (Eq. 4.43) confidence regions. An exemplary evaluation of this Hoeffding confidence bound can be found in Sec. A.6 as well as in [104, 105]. While these error bars may not be optimal, they are only based on the assumption of independent random variables, are easy to evaluate.

Consider for example the bipartite negativity (Sec. 2.4.1), that depends on the eigenvalues of the partial transpose of the state. Since it is difficult to express the bipartite negativity in terms of the measured count data, the error information is obtained by means of the linear approximation of the non-linear function. Exemplarily using the Hoeffding calculus demonstrates that error bars will be overestimated. Nevertheless, they help to rate the estimated value.

4.5.3 Examples

After establishing the fundamental methods for linear estimating different quantities of interest, some examples will further clarify this technique. Therefore, the negativity of a partially transposed state (see Sec. 2.4.1) will be considered. States that are separable with respect to the chosen partial transposition deliver a zero-value for the negativity, while entangled states can be detected by giving a positive value for this bipartite negativity. Here, the negativity of a product state and a Greenberger-Horne-Zeilinger state will be investigated. The considered product state is again $|\psi\rangle \propto (|H\rangle + |P\rangle)^{\otimes 4}$. Consequently, this product state is fully separable, irrespectively of the used partial transposition. Fig. 4.9 shows the negativity of the partial transpose of the reconstructed state when the entanglement with respect to the cut $AB|CD$ is evaluated. The linearly estimated lower bound for the negativity is compared with the values obtained for the reconstructed states using maximum likelihood and free least-squares fitting methods. The state is theoretically expected to show no entanglement, which is indicated by the dashed line. Indeed, the linear estimation delivers a lower bound for the negativity, which on average does not detect any entanglement. ML (blue) and FLS (red) in contrast detect entanglement. Please note that the linear estimation does not make any statement about the state being separable, but at least does not detect unexpected entanglement. While in Fig. 4.3 the fidelity was underestimated by the reconstruction methods that are based on fitting, now the situation is more delicate. Even fully separable states are detected to be entangled when using fitting methods for data with finite statistics. The L operator, that is used according to Eq. 4.28, is in this case obtained by expanding the negativity around a state given by a maximum likelihood fit on the data [23]. One uses a set of frequency data to reconstruct a state $\hat{\rho}_{\mathrm{ML}}$ which then is employed to build up the operator L. Afterwards, this operator is applied to another set of data to finally obtain $\mathrm{tr}(\rho L)$. If instead the theoretical state is considered to build up

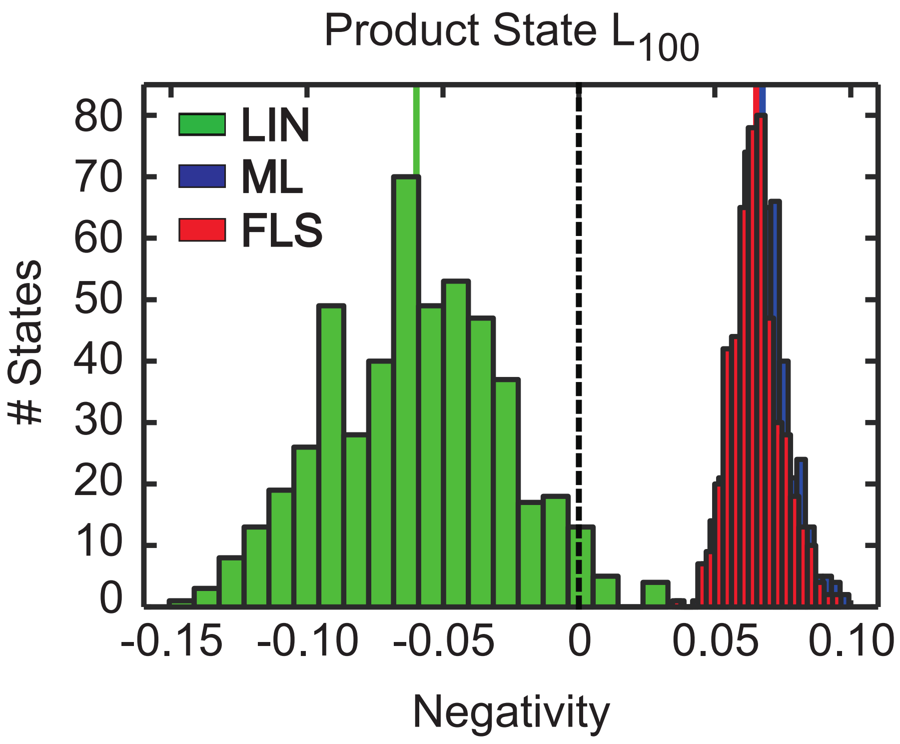

Figure 4.9: The bipartite negativity of a product state for simulated states with $N_S = 100$ counts per basis setting and admixed white noise such that $F_0 = 80\%$. The linear estimation (green histogram) gives a lower bound for the true value (dashed line), while maximum likelihood (blue histogram) and free least-squares (red) are used to reconstruct the 500 simulated states for which the negativity is calculated. The mean of the lower bounds (green line) give valid results, while both fitting methods detect more entanglement than present in the underlying fully separable quantum state.

the linear operator, this operator equals the null operator, always delivering the value 0 for the lower bound of the bipartite negativity, irrespective of the evaluation state $\rho(x)$. This again gives rise to the question how to find the optimal operator L.

The negativity is not only overestimated for a fully separable state but also for the already considered GHZ state, which can be seen in Fig. 4.10 (right subfigure). Both, the blue and the red histograms, show that the negativity obtained for maximum likelihood and free least-squares is above the theoretical expectation value for the GHZ state with admixed white noise ($F_0 = 80\%$ as used above) of 0.380, which is again shown as the dashed line. This emphasises the dependence of the estimated bound on the linearisation operator L. Indeed, in both cases the lower bound (green histogram) is on average below the theoretical value, while for better choices - as for example the operator L_{theo} - the mean of the obtained bounds is closer to the theoretical value than for worse estimated operators. The 68.3% confidence region for the operator L_{theo} is obtained with the Hoeffding calculus to be $C_L = 0.031$. This is in full agreement with the considerations of the linear approximation of convex quantities as illustrated in Fig. 4.7.

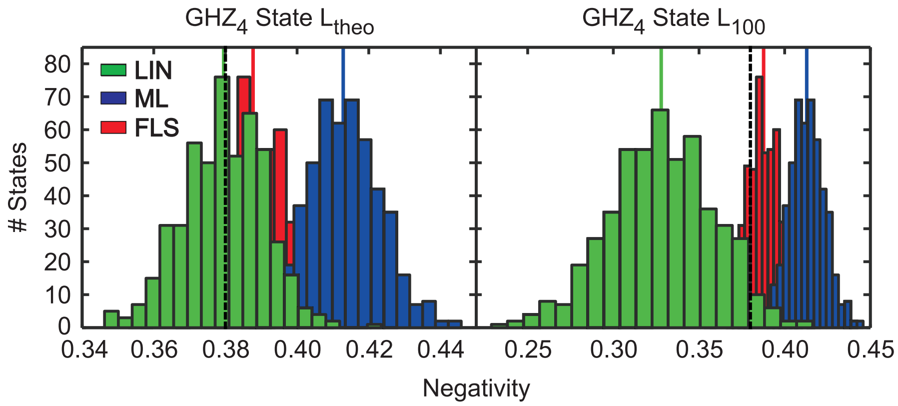

FIGURE 4.10: The bipartite negativity of a Greenberger-Horne-Zeilinger state with four qubits and admixed white noise ($F_0 = 80\%$). The linear approximation delivers lower bounds for the bipartite negativity. The quality of the approximation depends on the chosen linear operator. The expectation value of 0.380 is obtained for the lower bound if the operator is constructed by using the theoretical state (left figure). Using the Hoeffding calculus, one finds a 68.3% error bound at $0.380 - 0.031$. If instead the operator is based on fitting procedures (right figure), the obtained bound is less tight to the true value.

4.6 Eigenvalue modification

The reconstruction of a density operator based on experimental data often leads to unphysical states, as explained in detail previously. The frequently mentioned constraint of $\rho \geq 0$ is equivalent to the property that all eigenvalues of ρ have to be non-negative[9]. It may be possible to obtain an unbiased estimate of the state if making some appropriate assumptions on the underlying state and therefore dropping the generality of the state estimator.

States prepared in the laboratory are never pure due to experimental imperfections. This motivates the noise model of

$$\rho = p\rho_0 + (1-p)\rho_{\mathrm{WN}}, \tag{4.44}$$

where white noise is admixed to an underlying quantum state ρ_0. Hence, for the limit cases of $p = 1$ no additional white noise is admixed to ρ_0 and for $p = 0$ the state ρ only consists of white noise. For practical purposes the case of $p \in (0,1)$ will mainly be of interest. Throughout this section, a specialised quantum state reconstruction method for states of this noise model will be presented. Here, the noise model is exemplarily chosen to show the gain in quality of a quantum state estimation with a priori knowledge. If not only the noise model is predetermined, but also the amount of admixed white noise is assumed to be approximately known, the quantum state estimator can be further improved.

---

[9] $\rho \geq 0 \iff \langle\phi|\rho|\phi\rangle \geq 0\ \forall|\phi\rangle$. Choosing $\rho = \sum_i \lambda_i|\psi_i\rangle\langle\psi_i|$ and $|\phi\rangle = |\psi_j\rangle$, one obtains $\langle\psi_j|\sum_i \lambda_i|\psi_i\rangle\underbrace{\langle\psi_i|\psi_j\rangle}_{\delta_{i,j}} = \lambda_j \geq 0$.

Firstly, the behaviour of the state's eigenvalues will be examined when a set of finite measurement repetitions is simulated. Afterwards, motivated by insights from this study, a specialised quantum state estimator will be proposed and analysed.

4.6.1 Behaviour of eigenvalues due to finite statistics

Because of the admixture of white noise in the noise model given in Eq. 4.44 it is helpful to investigate the behaviour of the eigenvalues of the maximally mixed state (Sec. 2.5.8) first. The eigenvalues of the n qubit maximally mixed state ρ_{WN} are 2^n-fold degenerate with $\lambda_1 = \cdots = \lambda_{2^n} = \frac{1}{2^n}$. The state can be parametrised in its eigenvalues λ_i and its eigenvectors $|\psi_i\rangle$ according to

$$\rho = \sum_{i=1}^{2^n} \lambda_i |\psi_i\rangle\langle\psi_i|, \tag{4.45}$$

where the $\{|\psi_i\rangle\}_i$ is a set of any 2^n mutually orthonormal states, each contributing with the same probability $\lambda_i = \frac{1}{2^n}$. While for infinite count statistics the linearly reconstructed state equals ρ_{WN}, for finite number of counts per basis setting the eigenvalues' degeneracy is in general lifted. By investigating the eigenvalues of the $n = 4$ qubit maximally mixed state with $N_S = 100$ counts per basis setting, one finds that the distribution of the eigenvalues corresponds to a distribution with a semi-circular shape. In Fig. 4.11 the histogram of the eigenvalues of the mentioned state is shown. The distribution of the eigenvalues is centred around the value $\frac{1}{2^n}$, which corresponds to the degenerate

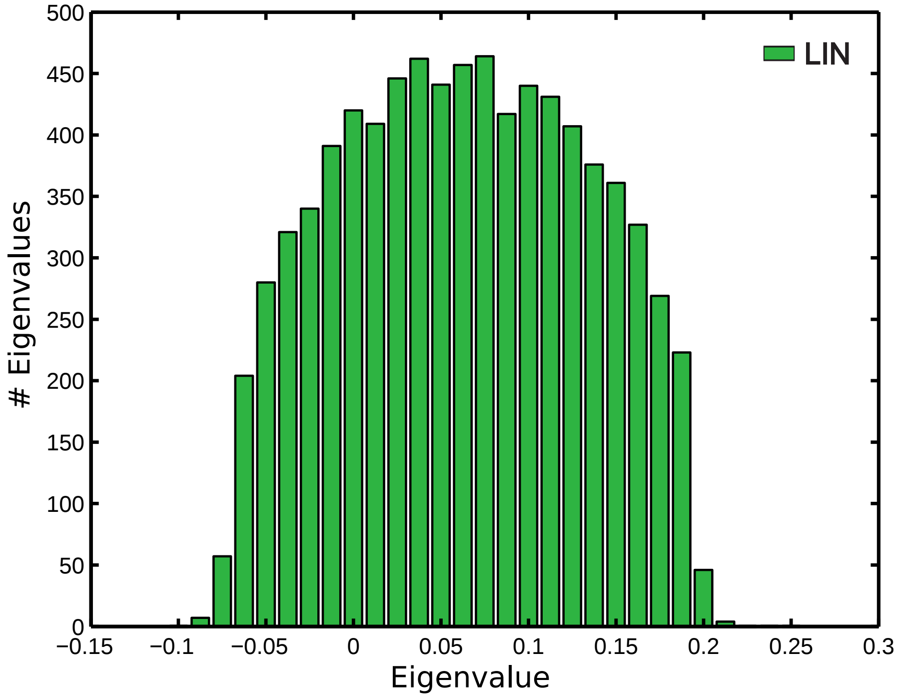

FIGURE 4.11: The distribution of the 2^4 eigenvalues of 500 sampled maximally mixed states ρ_{WN} with $n = 4$ qubits is shown. The simulated states are obtained when applying a Poissonian probability distribution to obtain a certain outcome according to Eq. 2.78. The expectation value for the total number of counts per basis setting was given by $N_S = 100$.

eigenvalue for the theoretical case. The radius of the semi-circular distribution can be empirically estimated to be around $R \approx 0.15$ for this case of 4 qubits and 100 counts per basis setting. As one would also naively assume, the spectral radius decreases with the statistics, which will be further investigated in Sec. 4.6.2.
The spectral probability density function of random matrices is its subject of research, cf. [106–109]. A grand goal of this field of study is to find the statistics of eigenvalues of random matrices with $N \times N$ elements for $\lim_N \to \infty$. The results obtained from random matrix theory suggest a semi-circular distribution of the spectral probability density function for matrices with elements of finite variances - at least for the case of infinitely large matrices. This serves as a motivation for empirical investigations for a low number of qubits.

While all eigenvalues with $\lambda_i \in [0, 1]$ can be interpreted as probabilities, this interpretation fails for $\lambda_i < 0$ (and for $\lambda_i > 1$, which as well may occur by reconstructing pure states with finite statistics). To overcome the problem of negative eigenvalues, one approach is to estimate how much the finite statistics could cause a scatter of the eigenvalues and finally try to revert this scatter. In this case the degeneracy is lifted towards a semi-circular distribution, i.e. the eigenvalues are found with high probability in a range of $[M - R; M + R]$, where R and M label the radius and the midpoint of the interval, respectively. A state ρ_0 with admixed white noise has $2^n - \text{rank}(\rho_0)$ degenerate eigenvalues at $\lambda_{\text{rank}(\rho_0)+1} = \cdots = \lambda_{2^n} = \frac{1-p}{2^n}$, where p denotes the amount of admixed white noise as given in the noise model of Eq. 4.44. These degenerate eigenvalues can cause the non-physicality of the state by becoming negative. The method presented here is based on the idea that eigenvalues within this interval are modified such that they are set to their respective average. The main concept of this quantum state estimation is easy and straight-forward, but the concrete implementation can be done in different ways. One of those will be briefly presented.

Error of matrix elements

According to Sec. 2.7 and to the given linear state reconstruction procedure presented there, the non-full correlation tensor elements are subject to lower statistical deviations than the full correlations. Thus, with the parametrisation given in Eq. 2.19 some matrix elements scatter more than others. Investigating the simple $n = 1$ qubit example for the density matrix, i.e.[10]

$$\rho = \frac{1}{2} \begin{pmatrix} T_0 + T_Z & T_X - iT_Y \\ T_X + iT_Y & T_0 - T_Z \end{pmatrix}, \tag{4.46}$$

the larger variance of the anti-diagonal elements compared to the diagonal elements is obvious as T_1 does not fluctuate. While in the single qubit case only two groups of entries with same variances exist (diagonal and anti-diagonal elements), in general the matrix elements of a n qubit density matrix can be put into $n + 1$ groups of same variances. Again, the lowest variance occurs for the $2^n \binom{n}{0}$ diagonal entries, which depend not only on the full correlation $T_{ZZ\ldots Z}$, but also on the correlations $T_{00\ldots 0}$, $T_{00\ldots Z}$, ..., $T_{ZZ\ldots 0}$, i.e. all possible tensor products of σ_z and σ_0. The second lowest variance can be found for the $2^n \binom{n}{1}$ elements, to which amongst others the correlations $T_{00\ldots 0X}$, $T_{00\ldots 0Y}$, $T_{00\ldots X0}$, ..., $T_{Y0\ldots 00}$ contribute. Finally, the largest statistical scatter appears in the $2^n \binom{n}{n}$ anti-diagonal elements corresponding to $T_{XX\ldots X}$, $T_{XX\ldots Y}$, $T_{XY\ldots X}$, ..., $T_{YY\ldots Y}$.

---

[10] To distinguish full and non-full correlations, here the correlation tensor will be indexed with 0, X, Y and Z.

Here, the largest statistical scatter occurring for the maximally mixed state will be deduced with the help of Gaussian error propagation to estimate the dependence of the variance on the statistics. Since the variance of the anti-diagonal terms is largest, the discussion will be restricted to these elements to obtain an upper bound of the variances. The standard deviation for all anti-diagonal elements $\{\rho_{i,2^q-i}\}_i$ is equal, therefore exemplarily the element $\rho_{1,2^n}$ is considered:

$$\Delta\rho_{1,2^n} = \sqrt{\sum_{i,S}\left(\frac{\partial\rho_{1,2^n}}{\partial c_i^S}\right)^2 (\Delta c_i^S)^2}. \tag{4.47}$$

As already argued before, the anti-diagonal terms depend on correlations of the measurements $\sigma_x\otimes\cdots\otimes\sigma_x\otimes\sigma_x$, $\sigma_x\otimes\cdots\otimes\sigma_x\otimes\sigma_y$, $\sigma_x\otimes\cdots\otimes\sigma_y\otimes\sigma_x$, $\ldots$, $\sigma_y\otimes\cdots\otimes\sigma_y\otimes\sigma_y$. For each basis of those 2^n settings 2^n outcomes are possible, therefore 4^n summands are left in the summation over outcomes i and settings S. After carrying out the partial derivatives (see Sec. A.2.2), one finally obtains

$$\Delta\rho_{1,2^n} = \frac{1}{N_S^2}\sqrt{\sum_S\left(\sum_{i;\text{odd}} c_i^S\right)^2 \sum_{i;\text{even}} (\Delta c_i^S)^2 + \left(\sum_{i;\text{even}} c_i^S\right)^2 \sum_{i;\text{odd}} (\Delta c_i^S)^2}, \tag{4.48}$$

where *even* and *odd* denote even or odd parity of the corresponding outcome. With the assumption of Poissonian distributed numbers of counts, Eq. 4.48 can be estimated for the n qubit maximally mixed state. According to Sec. A.7 one finds that the upper bound for the largest occurring variances - i.e. those on the antidiagonal of the matrix - is decreasing with the number of measurement repetitions N_S such that

$$\Delta\rho_{1,2^n} \leq \sqrt{\frac{2^n}{N_S}}. \tag{4.49}$$

With the aforementioned arguments, this expression is also an upper bound for all other variances of the density matrix.

4.6.2 Dependency of the semi-circular distribution on the number of counts

Fig. 4.11 showed the histogram of the eigenvalues of 500 maximally mixed states simulated with finite count statistics of $N_S = 100$. The distribution clearly can be assumed to be semi-circular - or strictly speaking: semi-elliptic. The distribution is centred around $1/2^4$ and extends in this case about 0.15 in both directions, whereby the approximation

$$f(x) = \begin{cases} \frac{1}{Z}\sqrt{R^2-(x-x_0)^2}, & \text{for } x\in[-R,R]\ , \\ 0, & \text{otherwise} \end{cases} \tag{4.50}$$

for the spectral probability density function with the radius $R = 0.15$, the centre $x_0 = 1/2^4$ and the normalisation $Z = \int_{-R}^{R}\sqrt{R^2-(x-x_0)^2}\mathrm{d}x = \pi R^2/2$ is in good agreement with the observed histogram. The radius of the semi-elliptic function $f(x)$, i.e. the spectral probability density function, is clearly dependent on the statistics which is given by the number of measurement repetitions N_S per basis setting S. This dependency can be read from Fig. 4.12, where for the case of the maximally mixed state with $n = 4$

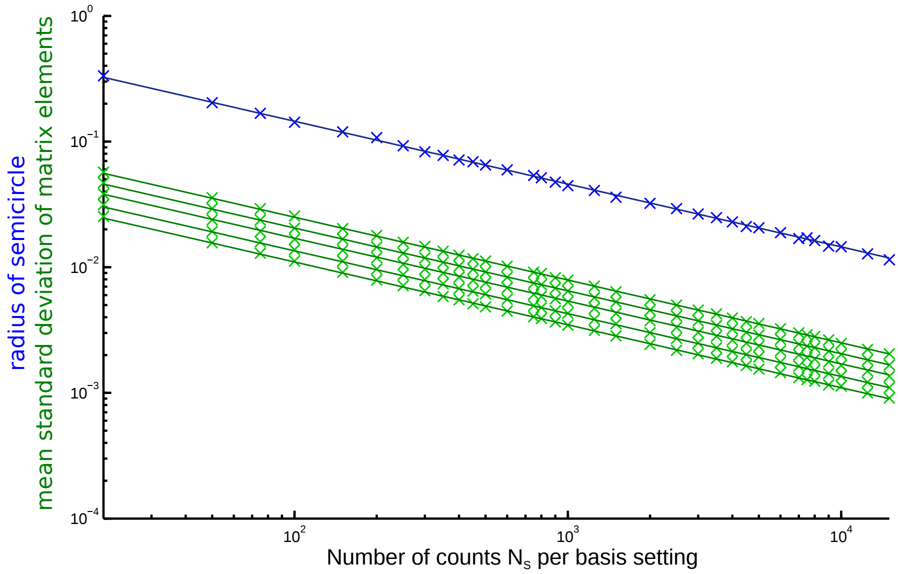

Figure 4.12: Radius of the maximally mixed state with 4 qubits for finite statistics in a log-log plot. The obtained radii for each histogram as shown in Fig. 4.11 decrease according to $\sqrt{\frac{1}{N_S}}$ with the statistics N_S. The behaviour corresponds to the standard deviations of the different matrix elements, which are grouped into five different classes as described in Sec. 4.6.1.

qubits the radius is determined for different N_S. Furthermore, the standard deviation of the different groups of matrix elements is calculated according to the discussion in Sec. 4.6.1. Eq. 4.49 leads to a relationship of $\frac{1}{\sqrt{N_S}}$ between the total number of counts N_S per basis setting and the estimation for the upper bound of the standard deviation of the matrix elements. This relationship is perfectly reflected in Fig. 4.12, where the standard deviations of the different groups of matrix elements are shown. This figure shows that also the radius decreases with an $\frac{1}{\sqrt{N_S}}$ behaviour which corresponds to a straight line in a log-log-type plot like Fig. 4.12. The semi-elliptic shape of the distribution's approximation $f(x)$ allows to estimate which eigenvalues originate from a set of degenerate eigenvalues and are distributed around some mean due to the finite statistics. This motivates the following state estimation method.

4.6.3 Averaging eigenvalues

The method based on the empirical insights from studying the distribution of eigenvalues is very straightforward. For the chosen noise model of $\rho = p\rho_0 + (1-p)\rho_{\mathrm{WN}}$ one observes a distribution of eigenvalues around their value of degeneracy of $\frac{1-p}{2^n}$ (see Sec. 4.6.1). The measured frequency data are used for a linear state reconstruction $\hat{\rho}_{\mathrm{LIN}}$. Afterwards, an eigendecomposition of this matrix is performed yielding the eigenvalues $\{\lambda_i\}_i$ and the respective eigenstates $\{|\psi_i\rangle\}_i$. If the lowest eigenvalue $\lambda_1 = \lambda_{\min}$ is found below zero, i.e. $\hat{\rho}_{\mathrm{LIN}} \ngeq 0$, the average $\overline{\lambda}$ over the eigenvalues $\{\lambda_1, \ldots, \lambda_k\}$ within the interval $I = [\lambda_1; \frac{p}{2^n} - \lambda_1]$ is calculated. If no further assumption for p is made, a conservative definition for I is used with $I = [\lambda_1; \frac{p}{2^n} - \lambda_1]$. Afterwards, the obtained value is used for all eigenvalues of the given interval I. Finally, the state estimate is given by the reconstructed state according to

$$\hat{\rho}_{\text{MOD}} = \sum_{i=1}^{2^n} \lambda_i |\psi_i\rangle\langle\psi_i| = \sum_{i=1}^{k} \overline{\lambda} |\psi_i\rangle\langle\psi_i| + \sum_{i=k+1}^{2^n} \lambda_i |\psi_i\rangle\langle\psi_i|. \tag{4.51}$$

If further assumptions are made, the procedure can be modified. Consider the case that one knows the amount of admixed white noise to be small, i.e. one guesses a value for p in the noise model. Since the degenerate eigenvalues are expected to be $\frac{1-p}{2^n}$, the interval for the eigenvalues to average can be adjusted. Suppose that one may use for example $p \leq 0.05$. Then, the eigenvalues of the admixed white noise are found at $\frac{0.05}{2^n}$ or below for the theoretical state. Assuming a symmetric distribution of the eigenvalues due to the finite statistics, they will be spread within $[\lambda_1; \frac{0.05}{2^n} - \lambda_1]$. This further assumption can preserve low eigenvalues as will be shown later with experimental data.
If not further stated, the quantum state estimator based on this procedure will be called *MOD* to emphasise the modification of the eigenvalue. The estimates will be denoted by $\hat{\rho}_{\text{MOD}}$. The variation of *MOD* based on further assumptions will be labelled with MOD_b.

Finally, another more adaptive variant of this procedure is conceivable. One might average over the eigenvalues $\{\lambda_1, \ldots, \lambda_k\}$, such that

$$\overline{\lambda} = \min_{k; \sum_{i=1}^{k} \lambda_i \geq 0} \frac{1}{k} \sum_{i=1}^{k} \lambda_i \tag{4.52}$$

is fulfilled. Consequently, one averages over the lowest k eigenvalues until their sum is non-negative anymore. This procedure ensures that the estimated state fulfils[11] $\rho \geq 0$.

4.6.4 Persistence of major contributions

One can express the state ρ in terms of its eigenvalues according to Eq. 4.45. Obviously, large eigenvalues contribute most to the state and carry the essential information while the small eigenvalues are often of no special interest. The presented method modifies the small eigenvalues, that can possibly be found below zero, to ensure $\rho \geq 0$, but does not affect the larger eigenvalues. Consider for example a state like the often mentioned GHZ state with admixed white noise such that $F_0 = 80\%$. The large eigenvalue $\lambda_{\max}$ of this state is on average $\lambda_{\max} = 0.8$. Theoretically, the corresponding eigenstate is the pure GHZ state. Since only the eigenvalues of the admixed white noise are modified, the contribution of the GHZ state stays the same. One might face the problem that the measured state contains contributions of coloured noise with small eigenvalues. Inherently, if these contributions are within the range of the eigenvalues that have to be modified, the eigenvalues of the coloured noise are finally set to the averaged value. These contributions, that are indistinguishable from the eigenvalues of the noise, are therefore not reconstructed correctly by this method. This problem will be discussed later when treating the non-generality of the method.

---

[11] Consider the case of a pure state with finite statistics where for example $\lambda_{\max} > 1$. Then, the average of all other eigenvalues is negative. In this case the adaptive variant averages over all eigenvalues, delivering the maximally mixed state. Obviously, this result is strongly biased. Nevertheless, all estimates fulfil $\hat{\rho}_{\text{MOD}} \geq 0$.

4.6.5 Obtaining the error of the state

The method to reconstruct a quantum state with the described modification on the eigenvalues of the state is statistically speaking a point estimation method exactly like the already discussed maximum likelihood and free least-squares estimators. Unlike a region estimator, the point estimator does not provide any information about the quality of the estimated value [110]. Consequently, one has to come back to bootstrapping methods to evaluate error bounds, as described in Sec. 4.2.4. Since the state estimation can be - when applied to states belonging to the chosen noise model - unbiased, non-parametric and parametric bootstrapping methods do not further shift the mean of the obtained quantity. Here, one does not report the standard deviation of a quantity evaluated of another state, as it happens when using bootstrapping with fitting methods for error estimation. In contrast, the standard deviation obtained by bootstrapping is evaluated by means of the desired state.

4.6.6 Numerical investigation

Finally, the performance of the proposed method for quantum state estimation will be investigated using numerical methods. In correspondence to the investigation of the bias of maximum likelihood and free least-squares state estimators, the mainly considered state is a four qubit $|\mathrm{GHZ}\rangle$ state with admixed white noise such that its fidelity with respect to the pure $|\mathrm{GHZ}\rangle$ is 80% ($F_0 = 80\%$, $p = 0.78\overline{6}$ in the noise model). At first, the quantum state estimation by modification of eigenvalues is tested with a linear measure. Here, the fidelity with respect to the theoretically expected state is regarded. Afterwards, also non-linear measures are treated. Since the direct evaluation of the state obtained by linear inversion is not meaningful, the values are compared to those obtained by linearisation of the respective convex or concave quantities as previously explained. Furthermore, the respective values for states reconstructed using maximum likelihood and free least-squares methods are shown for comparison.

Fidelity

The fidelity is again used as a linear measure[12] to compare the values determined by the states of linear inversion (Sec. 4.2.1) with those of the states obtained after modifying eigenvalues. Fig. 4.13 shows the great overlap of the distributions of the fidelities for the two compared estimators. The linearly inverted state gives a value of $\overline{F}_{\mathrm{LIN}} = 0.799 \pm 0.012$, while the quantum state reconstruction by modification of the eigenvalues delivers the same value of $\overline{F}_{\mathrm{MOD}} = 0.799 \pm 0.012$. Both results are in good agreement with the theoretical expectation of $F_0 = 0.8$. In contrast, Sec. 4.4 showed that the fidelity of maximum likelihood and free least-squares estimators are far off the expected value ($\overline{F}_{\mathrm{ML}} = 0.788 \pm 0.010$, $\overline{F}_{\mathrm{FLS}} = 0.749 \pm 0.010$). The good agreement of the reconstructed states $\hat{\rho}_{\mathrm{MOD}}$ with the expectation can also be explained by the argumentation of Sec. 4.6.4. The largest eigenvalue - the main contribution to the state - remains unmodified in this case. The eigenstate corresponding to this largest eigenvalue has a great overlap with the $|\mathrm{GHZ}\rangle$ state, whereby no major change of the fidelity is expected by this reconstruction method for this type of state.

---

[12] The fidelity is only linear if compared with a pure state like the $|\mathrm{GHZ}\rangle$ state in this case. Then, $\langle\psi|\rho|\psi\rangle$ is linear in x for the given parametrisation of $\rho \equiv \rho(x)$.

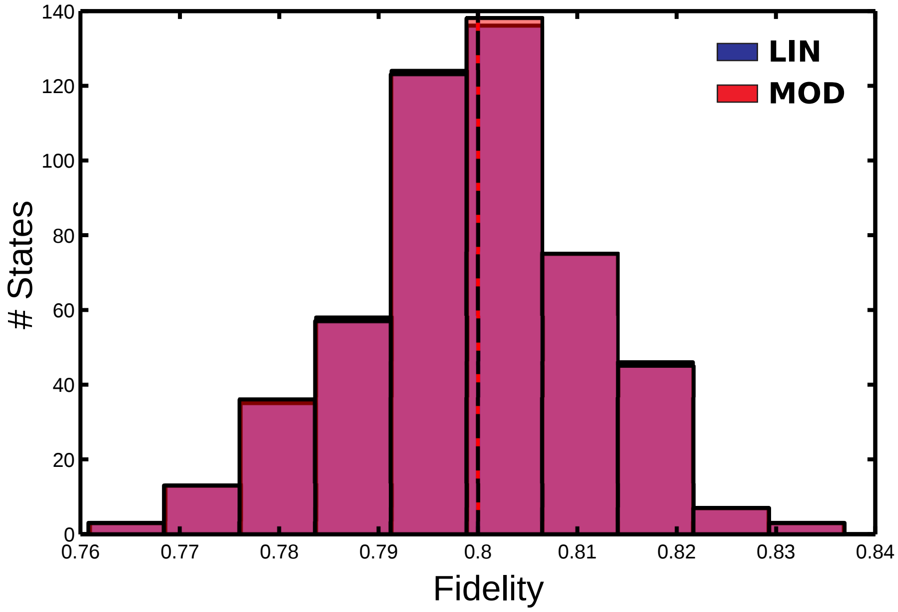

FIGURE 4.13: GHZ states simulated with 500 different diced sets of frequencies with admixed white noise and $N_S = 100$ counts per basis setting, are evaluated. Here, the proposed state estimation is used to determine the fidelity with respect to the pure 4 qubit GHZ state (red). The blue histogram shows the distribution of the directly linearly evaluated fidelity. This distribution is almost completely hidden by the state estimator based on modification of the eigenvalues (*MOD*), indicating a nearly perfect overlap. The theoretical value of $F_0 = 80\%$ is given by both estimators.

Non-linear measures

Unlike the fidelity, non-linear measures cannot be evaluated meaningfully on the linearly inverted state if $\hat{\rho}_{\mathrm{LIN}} \geq 0$ does not hold. Therefore, these non-linear functions are compared between the reconstructed state $\hat{\rho}_{\mathrm{MOD}}$ and the lower or upper bound, respectively, obtained by the linearisation as described in Sec. 4.5. The first quantity to be investigated is the purity. Since the purity $\mathrm{tr}(\rho^2) = \sum_i \lambda_i^2$ corresponds to the sum of the squared eigenvalues, one could imagine that this quantity is biased easily by applying a method modifying the eigenvalues. To discriminate the direct linear evaluation (*LIN*) from the estimate obtained by linear approximation, the latter will be called *BND*. This name is chosen to emphasise the fact that one obtains (one-sided) $\underline{\text{b}}$o$\underline{\text{u}}$$\underline{\text{nd}}$s, solely.

In contrast to one's expectation, the purity of the reconstructed states $\hat{\rho}_{\mathrm{MOD}}$ is similarly distributed as the lower bounds obtained by linearly approximating the purity, which is depicted in Fig. 4.14. Please note that the linear approximation does not necessarily give a value close to the real value, but only delivers a one-sided bound. Nevertheless, the fact that the theoretical expected value is within the error bounds of the red distribution of *MOD*, suggests a good quality of the estimation, at least for the considered state. For the eigenvalue modification, the purity is obtained to be $\overline{P}_{\mathrm{MOD}} = 0.653 \pm 0.019$, while the linear bound is determined to be 0.631 ± 0.019, when expanding around the states reconstructed with *MOD*. To linearise around the state found by modification of the eigenvalues, another 100 counts per basis setting are recorded. Afterwards, this additional set of data is used to find a state $\hat{\rho}_{\mathrm{MOD}}$. This state is applied to construct a linear operator L according to Sec. 4.5. These bounds will be denoted by BND_1. Using

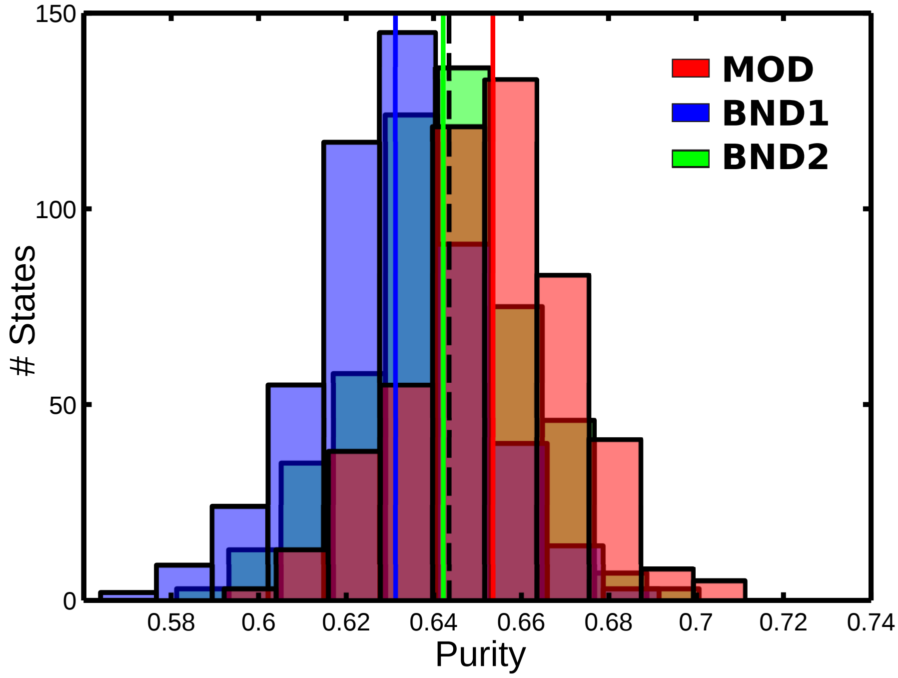

FIGURE 4.14: The purity of 500 Greenberger-Horne-Zeilinger states with admixed white noise ($F_0 = 80\%$) is determined via the reconstruction method by modifying the eigenvalues (*MOD*, red). The blue histogram shows the distribution of the lower bounds (BND_1) to the purity by a linear approximation. The linear approximation is done according to Sec. 4.5, where the linear operator is obtained by using $\hat{\rho}_{\mathrm{MOD}}$ with additional data. The green distribution (BND_2) shows the lower bounds for the purity when constructing the linear operator with the theoretical state.

instead the linear operator for the theoretical value, one obtains 0.642 ± 0.019 for the lower bound BND_2. The theoretically expected value of $P_0 = 0.643$ is therefore within the error bars of the purity of the reconstructed state and in agreement with the lower bounds, respectively.

Since this study is based on a huge amount of simulated states, the illustration in Fig. 4.14 shows the distribution of the directly obtained bounds without calculation of confidence regions, only. The effect of a 68.3% Hoeffding confidence level is now calculated, but not shown in the figure. The confidence region for the linear operator based on the theoretical state gives a value of $C_L \approx 0.048$. Therefore, using the linear bound BND_2, one obtains on average the statement that with at least 68.3% probability the purity is larger than $0.642 - 0.048 = 0.594$. Indeed, this is a veritable information, but not very rigorous. The value for C_L for a 68.3% confidence level is for BND_1 on average[13] $\overline{C_L} = 0.059$, leading on average to the statement that the true value is greater than $0.631 - 0.059 = 0.572$ with a probability of at least 68.3%.

Also, the negativity is inspected as a quantity of interest by using the proposed quantum state reconstruction method. In Sec. 4.5.3, the bipartite negativity for the cut $AB|CD$ was inspected for free least-squares and maximum likelihood estimators. This bipartite negativity is also used to study the state estimator based on modification of the eigenvalues. The theoretical value of $N_{0,AB|CD} = 0.380$ is in good agreement with

---

[13] In this case 500 slightly different linear operators L exist, each with a slightly different confidence interval. Consequently, their average is considered here.

the linear bound obtained by constructing the L operator based on the theoretical state, as shown as the green distribution in Fig. 4.15. The 68.3% Hoeffding confidence values are for BND_2 $C_L = 0.031$ and for BND_1 on average $\overline{C_L} = 0.086$. The average value for the bipartite negativity for states reconstructed with the here proposed method (red distribution, *MOD*, $\overline{N}_{AB|CD}(\hat{\rho}_{\text{MOD}}) = 0.405 \pm 0.016$) is not far away from the theoretical value, but though the theoretical value does not lie within one standard deviation. This corresponds to the feature mentioned in Sec. 4.3.2. Even if the state estimator gives fairly unbiased results for the considered subset of states, a non-linear function of the states is not necessarily unbiased.

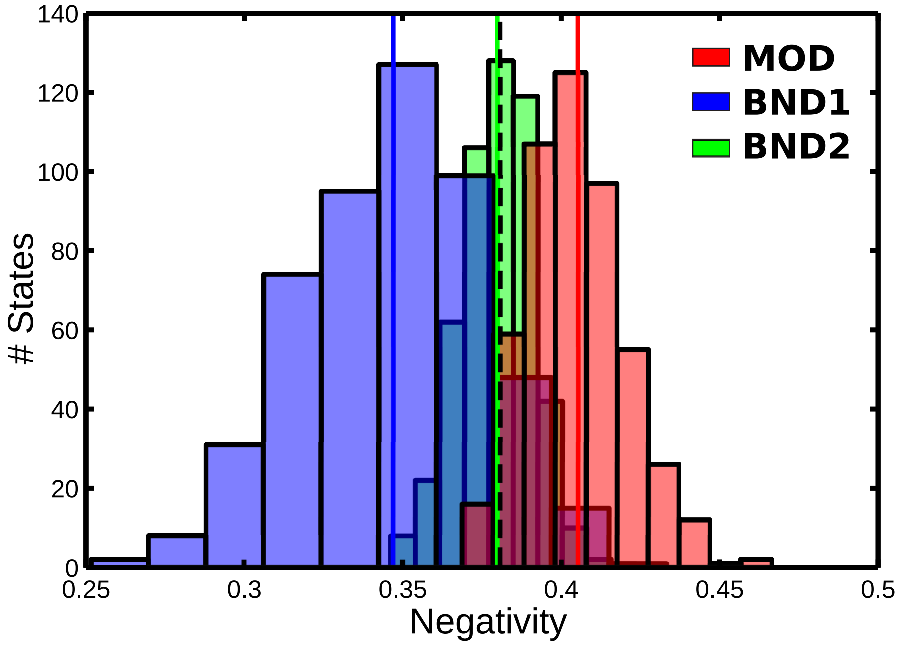

Figure 4.15: The distribution of the values obtained for the bipartite negativity (negativity of $\rho^{T_{AB}}$, see Sec. 2.4.1) is shown for the reconstruction based on modification of eigenvalues (*MOD*, red). This distribution is compared to the lower bounds obtained by using the linear approximation, when expanded around different states (expansion around theoretical state in green and around a state obtained by $\hat{\rho}_{\text{MOD}}$ in blue). The theoretically expected value of 0.380 is shown by the dashed line.

Finally, in Fig. 4.16 the distributions of the entropy are shown for the same three evaluation methods, i.e. for the modification of the eigenvalues *MOD*, for the linearisation around states obtained by modification of the eigenvalues BND_1 as well as for the linearisation around the theoretical state BND_2. Again, BND_1 uses another set of frequency data to reconstruct a state $\hat{\rho}_{\text{MOD}}$. This state is used to construct a linear operator, which is then applied to $\hat{\rho}_{\text{LIN}}$. Note that the last two methods only give upper bounds onto the entropy. The average value for the entropy of the states obtained by modification of eigenvalues is $\overline{S}_{\text{MOD}} = 1.465 \pm 0.071$ and is therefore in good agreement with the theoretical prediction of $S_0 = 1.503$. In contrast, the mean value of the entropy of the maximum likelihood estimates is $\overline{S}_{\text{ML}} = 1.236 \pm 0.049$ and therefore far off the expectation. By applying the Hoeffding calculus, one finds that the entropy is with probability of at least 68.3% larger than $1.502 - 0.071 = 1.431$. For BND_1, one obtains on average for the 500 used linear operators for the confidence interval according to Eq.

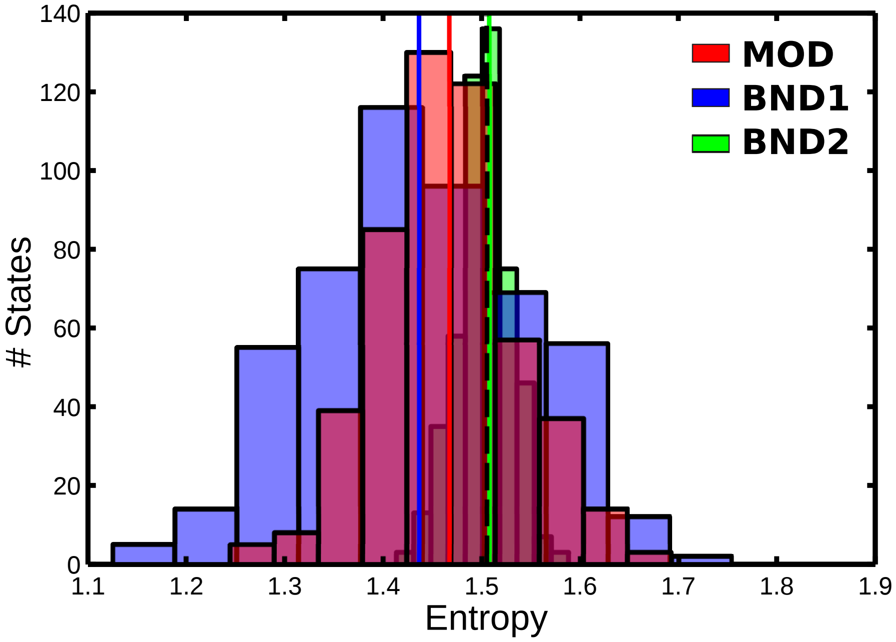

FIGURE 4.16: The entropy $\rho \log(\rho)$ is computed for the inspected $|\mathrm{GHZ}_4\rangle$ state with admixed white noise such that $F_0 = 80\%$. $N_S = 100$ measurement repetitions per basis setting are simulated. The theoretical value of 1.503 (dashed line) lies within the error bars of all three presented distributions. Again, the red distribution shows the entropy for the reconstructed state according to *MOD*, while the blue distribution, which mostly overlaps with the red distribution, delivers only upper bounds to the entropy. The operators BND_1 are expanded around the states $\hat{\rho}_{\mathrm{MOD}}$. In contrast, the green distribution BND_2 corresponds to the usage of the theoretical L operator.

4.43 a value of $\overline{C_L} = 0.080$. Therefore, one learns from BND_1 that the true value is on average with 68.3% probability larger than $1.434 - 0.080 = 1.354$.

Tab. 4.3 lists the values obtained for the considered four qubit Greenberger-Horne-Zeilinger with admixed white noise for different functions and different quantum state estimators or respectively approximations to the quantities. Note that the results for a direct evaluation with $\hat{\rho}_{\mathrm{LIN}}$ are far off for the non-linear measures and give meaningless values. The maximum value for the Fisher information of J_z achievable for a four qubit state is 16, while the linear inversion *LIN* gives 16.500 ± 48.347. Furthermore, the huge standard deviation of above 48 emphasises the absurdity of the result. Note that for the fidelity, the Fisher information, the entropy and the purity the values obtained by modifying the eigenvalues (*MOD*) as explained above leads to results where the theoretical predictions are well within the error bars. In the case of the bipartite negativity, which is used as an entanglement measure, the obtained value 0.405 ± 0.016 is above the expected value of 0.380.

Consequently, for the exemplarily studied state, the proposed method gives results generally in good agreement with the theoretical predictions or the linearly approximated value, respectively. Nevertheless, this estimator cannot give unbiased results for each and every quantum state and all non-linear functions, which was already claimed in the proposition in Sec. 4.3.2. On the other hand, if applicable, this quantum state estimator gives the chance to evaluate not only non-linear convex and concave functions. Indeed,

	Fidelity	Fisher info.	Entropy	Negativity	Purity
THEO	0.800	12.174	1.503	0.380	0.643
LIN	0.799 ± 0.012	16.500 ± 48.347	0.869 ± 0.135	0.615 ± 0.024	0.716 ± 0.020
MOD	0.799 ± 0.012	12.200 ± 0.226	1.465 ± 0.071	0.405 ± 0.016	0.653 ± 0.019
ML	0.788 ± 0.010	12.014 ± 0.444	1.236 ± 0.049	0.413 ± 0.011	0.636 ± 0.015
FLS	0.749 ± 0.010	11.287 ± 0.452	1.430 ± 0.048	0.388 ± 0.011	0.579 ± 0.014
BND_1	0.799 ± 0.012	12.106 ± 0.693	1.434 ± 0.110	0.346 ± 0.028	0.631 ± 0.019
BND_2	0.799 ± 0.012	12.149 ± 0.695	1.502 ± 0.027	0.379 ± 0.012	0.642 ± 0.019

TABLE 4.3: Overview of different methods to evaluate functions of the state. The theoretical values *THEO* are compared to the values obtained by reconstructing the state by means of the modification of eigenvalues (*MOD*), the maximum likelihood approach (*ML*) and the free least-squares fitting (*FLS*). The direct evaluation with the possibly unphysical matrix is shown in the row of *LIN*, while the lower and upper bounds are given by BND_1 and BND_2, respectively. BND_1 is obtained by constructing the linear operator with the help of the reconstructed state $\hat{\rho}_{\mathrm{MOD}}$, while BND_2 uses ρ_{theo} for finding L. The values given for the bounds are the mean and the standard deviation for the 500 states. The fidelity is calculated with respect to the pure $|\mathrm{GHZ}_4\rangle$ state. The considered Fisher information is based on the $J_z = \frac{1}{2}\left(\sigma_z \otimes \sigma_0^{\otimes 3} + \sigma_0 \otimes \sigma_z \otimes \sigma_0^{\otimes 2} + \sigma_0^{\otimes 2} \otimes \sigma_z \otimes \sigma_0 + \sigma_0^{\otimes 3} \otimes \sigma_z\right)$ operator [66].

the further requirement to the function to be convex or concave can be dropped if the eigenvalue modification is applied. Nevertheless, please be aware of the non-generality of this estimator and therefore the necessity of additional assumptions.

4.6.7 Non-generality of this method

As argued in Sec. 4.3.1, general quantum state estimators delivering always physical and unbiased results are not possible. Therefore, these strict constraints (physicality and unbiasedness) are relaxed. The here presented estimator is not able to deliver always physical density operators and can in some cases be also biased. Consider for example a pure state that is measured with finite statistics. Theoretically, one would expect all but one eigenvalues to vanish and the one left-over eigenvalue to be unity. In contrast, finite numbers of counts can also cause one large eigenvalue to be above unity, while the average of all other eigenvalues is below zero. Applying the quantum state estimator $\hat{\rho}_{\mathrm{MOD}}$ in the used variant does not lead to a physical state in this case[14]. Consequently, too pure states with low statistics are problematic. Nevertheless, the spectral distribution of the reconstructed state clearly reveals that the estimation did not succeed.

There are two possible ways to fail. Either the reported state does not fulfil the physicality constraint of $\hat{\rho}_{\mathrm{MOD}} \geq 0$ or the estimates are not unbiased. While the first case was already discussed, the latter one can occur if the contribution of the state, one is interested in, is very low and the amount of admixed white noise is large. If the state $\rho = p|\mathrm{GHZ}_4\rangle\langle\mathrm{GHZ}_4| + (1-p)\rho_{\mathrm{WN}}$ is used with $p = 0.04$ yielding $F_0 = 0.1$, the eigenvalues of ρ are $\lambda_1 = \cdots = \lambda_{15} = 0.06$ and $\lambda_{16} = 0.10$. The statistical scatter easily causes the eigenvalue λ_{16} to be within the interval of modification, leading to a result, where all eigenvalues are averaged over, such that $\lambda_1 = \cdots = \lambda_{16} = 1/2^4$. In this case, the

---

[14] Recall that the adaptive estimator mentioned in Sec. 4.6.3 would deliver a physical result - the maximally mixed state. Nevertheless, the result may be far from the theoretical state in this case.

estimator consequently delivers the maximally mixed state instead of being unbiased. The information (λ_{16}) cannot be distinguished from the noise (λ_1 to λ_{15}), leading to a loss of information. The same problem might occur if a state with one large contribution and a few minor contributions is considered. If the estimator is not engineered for this case by choosing for example a proper value for p, the information contained in these low contributions can be lost. One example for such a situation will be discussed later on when testing the estimator with an experimental state (Sec. 4.6.8).

Certificate for estimator

The evaluated quantities of the quantum states estimated by modifying the eigenvalues is in good agreement with the theoretical predictions. Nevertheless, it was argued that this method cannot give good results for all quantum states. For that purpose, one is interested in a certificate expressing the quality of the quantum state estimation. It is still a leftover question if such a certificate can be constructed to validate the resulting state after modifying the eigenvalues. Different approaches are conceivable. One could evaluate the quality of the state estimation for example by the sum of the squared changes of the eigenvalues to determine the influence of *MOD*. This value should consider the number of qubits and the statistics such that large modifications for states with good statistics are ranked as being more disputable than large modifications of states with low statistics. For a start, only states are considered that are reliably able to be modified like the states belonging to the mentioned noise model of Eq. 4.44.

4.6.8 Experimental verification

Finally, the method of modifying eigenvalues to obtain a supposable physical density operator is tested by using an experimental state. In Sec. 3, the *no correlation* state was analysed. For preparation, a Dicke state $|D_4^{(2)}\rangle$ was created and measured, which will now be used to test the quantum state estimator $\hat{\rho}_{\mathrm{MOD}}$. Tab. 4.4 shows the eigenvalues of the 4 qubit Dicke state with 2 excitations and higher order noise, which was already presented in Sec. 3.4.2. The eigenvalues *LIN* are obtained from the directly linearly inverted state such that also negative eigenvalues occur. Without further assumptions besides the general noise model of Eq. 4.44, the eigenvalues *MOD* are estimated by the here proposed estimator. One averages over all low eigenvalues such that the higher order noise, which is contained in the eigenvalues λ_{14} and λ_{15}, is lost. By assuming that the admixture of white noise is only very small, the eigenvalues of MOD_b are obtained. Eigenvalues λ_1 up to λ_{13} are averaged to be around $1.6 \cdot 10^{-5}$, while λ_{14} and λ_{15} preserve the higher order noise present in the state. In all cases, λ_{16} is not changed, whereby the fidelity of the estimated state with respect to the theoretical $|D_4^{(2)}\rangle$ state is expected to be reported unbiasedly.

Consequently, the most naive ansatz of averaging over all small eigenvalues gives already a good result. Nonetheless, a slight modification of the estimator still improves the results. Using the further assumption that the contribution of the admixed white noise is very low one can finally obtain a physical state with great accordance to the measured state. The fidelities achieved by the different state estimators with respect to the target state and states corresponding to higher order noise are listed in Tab. 4.5. The best results for the fidelities are indeed obtained when using the MOD_b estimator with further

Eigenvalue	LIN	MOD	MOD$_b$	ML	FLS
λ_1	-0.0232	0.0049	0.0000	0.0000	0.0000
λ_2	-0.0161	0.0049	0.0000	0.0000	0.0000
λ_3	-0.0138	0.0049	0.0000	0.0000	0.0000
λ_4	-0.0106	0.0049	0.0000	0.0000	0.0000
λ_5	-0.0069	0.0049	0.0000	0.0000	0.0000
λ_6	-0.0036	0.0049	0.0000	0.0000	0.0000
λ_7	0.0018	0.0049	0.0000	0.0000	0.0000
λ_8	0.0047	0.0049	0.0000	0.0000	0.0000
λ_9	0.0078	0.0049	0.0000	0.0010	0.0012
λ_{10}	0.0096	0.0049	0.0000	0.0023	0.0028
λ_{11}	0.0119	0.0049	0.0000	0.0063	0.0065
λ_{12}	0.0183	0.0049	0.0000	0.0097	0.0098
λ_{13}	0.0206	0.0049	0.0000	0.0124	0.0129
λ_{14}	0.0275	0.0049	0.0275	0.0153	0.0155
λ_{15}	0.0460	0.0049	0.0460	0.0296	0.0303
λ_{16}	0.9261	0.9261	0.9261	0.9235	0.9210

TABLE 4.4: The eigenvalues of the measured state of Sec. 3.4.2 are compared for different state estimators. *LIN* labels the eigenvalues of the state directly obtained by interpreting the measured frequencies as probabilities, while *MOD* and *MOD$_b$* label the quantum state estimators based on modification of the eigenvalues. *MOD* averages over all eigenvalues $\lambda_i \in [\lambda_1; 1/16 - \lambda_1]$. *MOD$_b$* is constructed by being aware of coloured noise. Therefore, *MOD$_b$* averages only eigenvalues with $\lambda_i \in [\lambda_1; p/16 - \lambda_1]$ with $p = 0.02$, assuming only 2% of white noise in the theoretical state. All choices of $p \in [0; 0.056]$ would reproduce these eigenvalues, while larger values for p are also considering λ_{14} and probably λ_{15} to average over.

Fidelity target state	LIN	MOD	MOD$_b$	ML	FLS
$\vert D_4^{(0)}\rangle$	0.0001	0.0049	0.0005	0.0004	0.0005
$\vert D_4^{(1)}\rangle$	0.0334	0.0087	0.0343	0.0263	0.0268
$\vert D_4^{(2)}\rangle$	0.9195	0.9194	0.9195	0.9173	0.9148
$\vert D_4^{(3)}\rangle$	0.0102	0.0070	0.0114	0.0112	0.0114
$\vert D_4^{(4)}\rangle$	-0.0016	0.0049	0.0011	0.0004	0.0005

TABLE 4.5: Comparison of the fidelities of the reconstructed state with respect to the Dicke states with 4 qubits. The method *MOD$_b$* gives the best values for the fidelity with respect to the desired target state $D_4^{(2)}$ as well as with respect to the higher orders $D_4^{(1)}$ and $D_4^{(3)}$. The negative fidelity $F(\hat{\rho}_{\mathrm{LIN}}, |D_4^{(4)}\rangle) = -0.0016$ is due to the unphysicality of $\hat{\rho}_{\mathrm{LIN}}$.

	Fisher info. (J_z)	Entropy	Negativity	Purity
LIN	0.9280	$0.4622 + 0.3365i$	0.7590	0.8628
MOD	0.0244	0.6690	0.7470	0.8580
MOD$_b$	0.1916	0.4555	0.7429	0.8605
ML	0.1416	0.5679	0.7453	0.8542
FLS	0.1492	0.5842	0.7425	0.8498
BND$_{\mathrm{ML}}$	≥ 0.2002	≤ 0.4933	≥ 0.7202	≥ 0.8433
BND$_{\mathrm{MOD}_b}$	≥ 0.0217	≤ 0.5392	≥ 0.7289	≥ 0.8414
BND$_{\mathrm{ML}}$	$\geq_{68.3\%} -0.0641$	$\leq_{68.3\%} 0.5364$	$\geq_{68.3\%} 0.6950$	$\geq_{68.3\%} 0.8224$
BND$_{\mathrm{MOD}_b}$	$\geq_{68.3\%} 0.0148$	$\leq_{68.3\%} 0.5839$	$\geq_{68.3\%} 0.7118$	$\geq_{68.3\%} 0.8202$

TABLE 4.6: Comparison of non-linear measures for the experimentally prepared state for different state reconstruction methods. Here, not only the boundary values *BND* for the linear approximation are listed, but also the weakening effect of the inequality due to the Hoeffding calculus is shown. All error bounds are computed for 68.3% confidence level. The linear operator L of BND_{ML} is constructed by using the maximum likelihood fit, while BND_{MOD_b} employs the reconstruction of MOD_b. The measured data was probabilistically splitted up into two new sets of data. One of those data sets was employed via a state reconstruction (*ML* or MOD_b) to construct the linear operator L while the other set was used with linear inversion to find the lower/upper bound by $\mathrm{tr}(L\hat{\rho}_{\mathrm{LIN}})$. $N_S = 920$ is used for evaluating C_L for the Hoeffding bound (see Eq. 4.43). Please note that the state $\hat{\rho}_{\mathrm{LIN}}$ is not physical. Therefore, meaningless results are expected when it is directly used for function evaluation, cf. for example the Fisher information or the entropy.

assumptions. Tab. 4.6 lists the values for non-linear functions obtained by different methods for state estimation. While the negativity and the purity are similar to each other for the four methods delivering a physical state, the Fisher information and the entropy differ more from each other. Strikingly, the simple linear inversion *LIN* delivers a value for the Fisher information that is far off the other values. Furthermore, its entropy is a complex number away from any physical interpretation. The values given for the one-sided bounds, i.e. obtained by means of the linear approximation, are weakened when using the Hoeffding calculus. Therefore, for instance the value for the entropy is bounded by ≤ 0.4933 when using BND$_{\mathrm{ML}}$. If the 68.3% confidence interval is calculated according to Hoeffding, this bound is given by $\leq_{68.3\%} 0.5364$. Tab. 4.6 illustrates that the maximum likelihood and the free least-squares results are not in agreement with the directly obtained one-sided bounds for the Fisher information and the entropy when using the maximum likelihood state for the construction of the linear operator. After weakening the bounds by using the 68.3% confidence interval by means of the Hoeffding calculus, maximum likelihood and free least-squares fulfil the inequalities for the Fisher information. In contrast, the eigenvalue modification is in agreement with the bounds when using the state of $\hat{\rho}_{\mathrm{MOD}_g}$ for constructing the operator. Please note that the boundaries are obtained by using two sets of data. The inital set of about 1840 counts per basis setting is probabilistically splitted into two sets. The first set with about 920 counts per basis setting is used to construct a linear operator by means of a maximum likelihood fit or the modification of eigenvalues. Afterwards, the second set with also about 920 counts per basis setting leads to a linearly inverted state $\hat{\rho}_{\mathrm{LIN}}$ that delivers the bound by calculating $\mathrm{tr}(L\hat{\rho}_{\mathrm{LIN}})$. This ensures that the obtained measurement data are not used twice.

Thus, the quantum state estimator using modified eigenvalues gives a state whose fidelities with respect to the Dicke state $|D_4^{(2)}\rangle$ and the contributions of higher orders are in accordance with the linearly evaluated fidelities. Furthermore, non-linear measures are reported in agreement with the bounds obtained by the presented linear approximation. Nevertheless, note that the experimental data offers rather good statistics.

4.7 Conclusion

In this chapter, two methods were introduced that allow estimating non-linear functions of states. Using the measured data to construct a density operator may fail in many cases since the obtained matrix does not fulfil the physicality constraint of $\hat{\rho}_{\mathrm{LIN}} \geq 0$. The result lacks a physical meaning. Eigenvalues that - for a physical density operator - can be interpreted as probabilities are not longer bounded between zero and unity if the result is not positive semi-definite. Evaluating measures like the entropy where the eigenvalues of the state are used consequently leads to meaningless conclusions.

Here, it was shown that procedures based on fitting the unphysical result to the state $\hat{\rho}_{\mathrm{fit}}$ with $\hat{\rho}_{\mathrm{fit}} \geq 0$, in best agreement with the obtained tomographic data, leads to biased statements. Instead, linear functions like the fidelity with respect to a pure state, can be evaluated directly on the possibly unphysical result. Furthermore, this chapter explains and illustrates by examples how non-linear functions can be linearly approximated if they are either convex or concave. This procedure works for any type of state, but is intrinsically restricted to convex and concave functions. Thereby, one is able to find lower or upper bounds for the quantity of interest even if the reconstructed state does not represent a physical state.

Since this method is only applicable for convex or concave functions, one is still interested in a reliable quantum state estimator. One method for a particularly chosen noise model was proposed. This procedure, based on modification of the eigenvalues, is not a general method, but provides good results for the considered measures. The quantum states reconstructed by modifying the eigenvalues of the unphysical result allow to obtain bounds on the evaluated functions. Here, possibly present systematical deviations are in most cases negligible compared to any statistical scatter. This easy state reconstruction method is a promising tool for estimating non-convex and non-concave functions of quantum states where the linear approximation cannot be used. Nevertheless, it is crucial to keep in mind that even an unbiased quantum state estimator may give biased results for non-linear functions.

Chapter 5

Conclusion

This thesis is composed of two mostly mutually independent topics. Chapter 3 presents novel results on the observation of genuine n-partite entanglement even if no n-partite correlations are present. Naively, one would expect that a genuinely entangled quantum state causes inevitably correlations between all involved parties. In contrast, this study gives a counterexample of a state with genuine tripartite entanglement, but without any correlations between the outcomes of the three parties. This firstly considered quantum state is experimentally prepared and analysed to prove the mentioned phenomenon. To show that this effect is not special for the very particular chosen state, a whole class of states is theoretically analysed afterwards. Nevertheless, a generalisation of this feature to states with an even number of qubits is still missing. Probably, it is possible to find a state with even number of qubits that is the incoherent sum of more than two pure states and is genuinely entangled as well even if it does not show full correlations. If it is already sufficient to mix two states for even numbers of qubits to obtain a state with these properties, cannot be answered so far.

Furthermore, this thesis deals in chapter 4 with the reconstruction of quantum states by means of measured data. The experimental data delivers mostly an unphysical result if used straightforward to build up a quantum state. In this case, a multitude of functions cannot be evaluated in a meaningful way. Therefore, it is common practise to apply fitting methods that deliver the (physical) quantum state that is in best agreement with the measured data. These maximum likelihood and free least-squares procedures are studied in detail and found to give biased results. Furthermore, the commonly used methods to report error bars are highly questionable since they do not sample error information of the underlying quantum state, but instead are based on an already biased quantum state. Consequently, another method is presented that gives linear approximations of the considered functions. Thereby, unbiased information about non-linear measures can be extracted. Furthermore, confidence intervals can be determined easily and without any further assumptions. Unfortunately, this procedure gives only one-sided bounds and is not always applicable since the linearisation is only possible for convex or concave functions. The linear approximation is based on a physical quantum state. Although all choices for this state to construct a linear approximation are valid, it is still an open question how the best approximation can be found.

Finally, another method to estimate a quantum state was presented that is based on further assumptions. Exemplarily, a predefined noise model was used. The constructed quantum state estimator lacks the generality of the presented fitting methods, but provides promising results if the assumptions are properly chosen.

Appendix A

Additional calculations

A.1 Derivation of constraints bounding the correlations

If the density matrix is parametrised in terms of its eigenvalues and eigenstates, it is obvious, that the eigenvalues may not be larger than one and cannot go below zero. Instead of using its eigendecomposition, it is often helpful to describe a state ρ in terms of correlations, i.e. one uses the correlation tensor as for example given in Eq. 2.18. The constraint onto the eigenvalues $\{\lambda_i\}_i$ with $\lambda_i \in [0,1]$ for $i = 1, \ldots, 2^n$ and $\sum_{i=1}^{2^n} \lambda_i = 1$ leads to another constraint onto the correlations. Its derivation will be given in Eq. A.1, which is just based on the fact, that the state's purity cannot exceed the value 1. Please note, that here all 4^n correlations are taken into account, not only the 3^n full correlations.

$$\begin{aligned}
\operatorname{tr}\left(\rho^2\right) &= \operatorname{tr}\left(\frac{1}{2^{2n}} \sum_{\mu_1,\ldots,\mu_n=0}^{3} T_{\mu_1\ldots\mu_n} \bigotimes_{i=1}^{n} \sigma_{\mu_i} \sum_{\nu_1,\ldots,\nu_n=0}^{3} T_{\nu_1\ldots\nu_n} \bigotimes_{j=1}^{n} \sigma_{\nu_j}\right) \\
&= \frac{1}{2^{2n}} \sum_{\mu_1,\ldots,\mu_n=0}^{3} \sum_{\nu_1,\ldots,\nu_n=0}^{3} T_{\mu_1\ldots\mu_n} T_{\nu_1\ldots\nu_n} \operatorname{tr}\left(\bigotimes_{i=1}^{n} \sigma_{\mu_i} \bigotimes_{j=1}^{n} \sigma_{\nu_j}\right) \\
&= \frac{1}{2^{2n}} \sum_{\mu_1,\ldots,\mu_n=0}^{3} \sum_{\nu_1,\ldots,\nu_n=0}^{3} T_{\mu_1\ldots\mu_n} T_{\nu_1\ldots\nu_n} \delta_{\vec{\mu},\vec{\nu}} 2^n \\
&= \frac{1}{2^n} \sum_{\mu_1,\ldots,\mu_n=0}^{3} T^2_{\mu_1\ldots\mu_n} \leq 1 \\
&\Rightarrow \sum_{\mu_1,\ldots,\mu_n=0}^{3} T^2_{\mu_1\ldots\mu_n} \leq 2^n .
\end{aligned} \tag{A.1}$$

A.2 Parity function g

A.2.1 Definition

The function $g(\mu_1, \dots, \mu_n; i_1, \dots, i_n)$ defines the signatures of the contributing number of counts to the desired correlation as described in Eq. 2.80. Therefore, the parity function may be defined as

$$g(\mu_1, \dots, \mu_n; i_1, \dots, i_n) = (-1)^{N(\mu_1, \dots, \mu_n; i_1, \dots, i_n)}, \tag{A.2}$$

where $N(\mu_1, \dots, \mu_n; i_1, \dots, i_n)$ counts, how many elements of $i_1, \dots, i_n$ are 1, if the corresponding measurement direction is $\{x, y, z\}$, i.e. $N(\vec{\mu}, \vec{i}) = \sum_{j=1}^{n} (\delta_{\mu_j,x} + \delta_{\mu_j,y} + \delta_{\mu_j,z}) i_j$, where i_j is 0, if the chosen eigenvector belongs to the eigenvalue 1 of the matrix σ_{μ_j}, and 1, if it corresponds to the eigenvalue -1 of the Pauli matrix of the jth qubit. If instead only full basis settings are considered, the arguments $\mu_1, \dots, \mu_n$ can be skipped. Then, Eq. A.2 simplifies to

$$g(i_1, \dots, i_n) = (-1)^{\sum_j i_j}. \tag{A.3}$$

A.2.2 Calculating partial derivatives

In Sec. 4.6.1 one is interested in the partial derivatives of the correlations and the respective numbers of counts. It holds

$$\begin{aligned}
\frac{\partial T_S}{\partial c_j^S} &= \frac{\partial \frac{\sum_i g(i_1, \dots, i_n) c_i^S}{\sum_i c_i^S}}{\partial \tilde{c}_j^S} \\
&= \frac{g(j_1, \dots, j_n) \sum_i c_i^S - \sum_i g(i_1, \dots, i_n) c_i^S}{(\sum_i c_i^S)^2} \\
&= -2 \frac{\sum_{i;\text{odd}} c_i^S \delta_{1,g(j_1, \dots, j_n)} + \sum_{i;\text{even}} c_i^S \delta_{-1,g(j_1, \dots, j_n)}}{\left(\sum_i c_i^S\right)^2},
\end{aligned} \tag{A.4}$$

where *even* and *odd* denote the parity of the outcomes to sum over. Therefore, if one is for example interested in $\partial T_{11} / \partial c_{00}^{11}$, evaluating the g function to $g(j_1, \dots, j_n) = g(0, 0) = 1$ leads to

$$\partial T_{11} / \partial c_{00}^{11} = -2 \frac{\sum_{i;\text{odd}} c_i^{11}}{\left(\sum_i c_i^{11}\right)^2} = -2 \frac{c_{01}^{11} + c_{10}^{11}}{\left(\sum_i c_i^{11}\right)^2}. \tag{A.5}$$

A.3 Predictability

The predictability (Sec. 3.2) will be further illuminated by the state considered in chapter 3: $\rho_{\text{NC},3} = \frac{1}{8}(|001\rangle\langle 001| + |110\rangle\langle 110| + \dots)$. Since writing all terms would be lengthy and of no additional information, here only the excerpt of the matrix elements is given that are relevant for the following example. Consider now first the case of full knowledge, i.e. the third participant of an experiment knows the measurement outcome of both the other parties, who are responsible for qubits 1 and 2, respectively. Assume, the first two measurements were resulting in 0. Already from this information the third party can directly infer the results of his measurement, since $Pr_{|\psi\rangle_1=|0\rangle,|\psi\rangle_2=|0\rangle}(|\psi\rangle_3 = |0\rangle) \propto |\langle 000|\rho_{\mathrm{NC},3}|000\rangle|^2 = 0$ and $Pr_{|\psi\rangle_1=|0\rangle,|\psi\rangle_2=|0\rangle}(|\psi\rangle_3 = |1\rangle) \propto |\langle 001|\rho_{\mathrm{NC},3}|001\rangle|^2 = \alpha$. The proportionality is given instead of equality, since the conditional probability has to be normalised by marginalising over all cases, that are in agreement with the condition, such that $Pr_B(A) \equiv Pr(A|B) = Pr(A,B)/Pr(B)$. Here, this causes a renormalisation with α. Thus, the third party knows the measurement outcome of $|\psi\rangle_3$ to be $|1\rangle$ with probability $\frac{\alpha}{\alpha} = 1$.

Compare now this case of full knowledge with the case of knowing the correlation between two qubits only to infer the outcome of the last one. Then, for the case of correlated qubits 1 and 2,

$$Pr_{|\psi\rangle_1=|\psi\rangle_2}(|\psi\rangle_3 = |0\rangle) \propto \underbrace{|\langle 000|\rho_{\mathrm{NC},3}|000\rangle|^2}_{=0} + |\langle 110|\rho_{\mathrm{NC},3}|110\rangle|^2 = \alpha, \tag{A.6}$$

while

$$Pr_{|\psi\rangle_1=|\psi\rangle_2}(|\psi\rangle_3 = |1\rangle) \propto |\langle 001|\rho_{\mathrm{NC},3}|001\rangle|^2 + \underbrace{|\langle 111|\rho_{\mathrm{NC},3}|111\rangle|^2}_{=0} = \alpha. \tag{A.7}$$

Therefore, knowing the first two qubits to have same outcomes, but not knowing, what concrete outcomes they have, does not yield any knowledge about the third qubit, while knowing the individual outcomes can allow predicting the third qubit's results.

A.4 Tracing out one qubit of a Dicke state

The state $\rho_{\mathrm{NC},3} = \frac{1}{2}(|\mathrm{W}_3\rangle\langle \mathrm{W}_3| + |\overline{\mathrm{W}}_3\rangle\langle\overline{\mathrm{W}}_3|)$ that is given in Sec. 3.4.1 has to be prepared experimentally. Since preparing the states $|\mathrm{W}_3\rangle$ and $|\overline{\mathrm{W}}_3\rangle$ and mix them afterwards can be cumbersome, a more straightforward way is to prepare a Dicke state $|D_4^{(2)}\rangle$ and trace out one qubit. For completeness, it will be shown in Eq. A.8 that $\mathrm{tr}_4(|D_4^{(2)}\rangle\langle D_4^{(2)}|) = \rho_{\mathrm{NC},3}$ holds.

$$\begin{aligned}
\mathrm{tr}_4\left(|D_4^{(2)}\rangle\langle D_4^{(2)}|\right) =& \frac{1}{6}\big(\mathrm{tr}_4\left(|0011\rangle\langle 0011|\right) + \mathrm{tr}_4\left(|0011\rangle\langle 0101|\right) \\
&+ \mathrm{tr}_4\left(|0011\rangle\langle 0110|\right) + \cdots + \mathrm{tr}_4\left(|1100\rangle\langle 1100|\right)\big) \\
=& \frac{1}{6}\left(|001\rangle\langle 001| + |001\rangle\langle 010| + \cdots + |110\rangle\langle 110|\right) \\
=& \frac{1}{2}\Big[\frac{1}{3}(|001\rangle + |010\rangle + |100\rangle)(\langle 001| + \langle 010| + \langle 100|) \\
&+ \frac{1}{3}(|110\rangle + |101\rangle + |011\rangle)(\langle 110| + \langle 101| + \langle 011|)\Big] \\
=& \frac{1}{2}(|\mathrm{W}_3\rangle\langle \mathrm{W}_3| + |\overline{\mathrm{W}}_3\rangle\langle\overline{\mathrm{W}}_3|) \\
=& \rho_{\mathrm{NC},3}.
\end{aligned} \tag{A.8}$$

A.5 Optimisation over bi-separable states

This section derives how the genuine tripartite entanglement is detected, as it is used in Sec. 3.3.2. The derivation is based on ideas of Tomasz Paterek and will be explained in [13]. Recall the entanglement criterion as used in Sec. 3.4.5 where

$$\max_{\hat{T}^{\text{bi-prod}}} \left(\hat{T}, \hat{T}^{\text{bi-prod}}\right)_G < \left(\hat{T}, \hat{T}\right)_G \tag{A.9}$$

holds if the state ρ with the correlation tensor $\hat{T}$ is genuinely n-partite entangled. Define the left-hand side as $L = \max_{\hat{T}^{\text{bi-prod}}} \left(\hat{T}, \hat{T}^{\text{bi-prod}}\right)_G$ and the right-hand side as $R = \left(\hat{T}, \hat{T}\right)_G$. Now, the derivation of L is needed to show the genuine tripartite entanglement of the mixture of the generalised $|\mathrm{W}_3^g\rangle$ state and its antistate, as described in Sec. 3.3.2. The used metric is chosen such that $G_{\Pi(XX0)} = G_{\Pi(YY0)} = G_{\Pi(ZZ0)} = 1$ holds for Π denoting all permutations. The choice of the metric significantly differs from the usual choice $G^{\mu_1,\dots,\mu_n,\nu_1,\dots,\nu_n} = \delta^{\mu_1,\nu_1} \cdots \delta^{\mu_n,\nu_n}$ (see Sec. 2.4.1). This metric is motivated by the fact that $\rho_{\text{NC},3}$ has no full correlations but necessarily non-full correlations.

The value for the right-hand side of Eq. A.9 is found to be $R = 3$ because

$$\begin{aligned}
R = &\left(\hat{T}, \hat{T}\right)_G \\
=& T_{XX0}^2 + T_{YY0}^2 + T_{X0X}^2 + T_{Y0Y}^2 + T_{0XX}^2 + T_{0YY}^2 + T_{ZZ0}^2 + T_{Z0Z}^2 + T_{0ZZ}^2 \\
=& 2\left(\sin(2\beta)\sin(\alpha)\right)^2 + 2\left(\sin(2\beta)\cos(\alpha)\right)^2 + 2\left(\sin^2(\beta)\sin(2\alpha)\right)^2 \\
&+ \left(\cos(2\alpha)\sin^2(\beta) - \cos^2(\beta)\right)^2 + \left(-\cos(2\alpha)\sin^2(\beta) - \cos^2(\beta)\right)^2 + \cos^2(2\beta) \\
=& 3.
\end{aligned} \tag{A.10}$$

The calculation of the value for the left-hand side L of Eq. A.9 is more laborious thus it is sketched, only. Use without loss of generality the bi-product state $(\cos(\theta)|00\rangle + \sin(\theta)|11\rangle) \otimes |c\rangle$ is chosen with the local Bloch vectors [50] $\vec{a}$, $\vec{b}$ and $\vec{c}$. Therefore, L is obtained to be

$$\begin{aligned}
L =& T_{XX0}\left(T_{XX0}^{\text{bi-prod}} + T_{YY0}^{\text{bi-prod}}\right) \\
&+ T_{ZZ0}T_{ZZ0}^{\text{bi-prod}} + T_{X0X}(a_X c_X + a_Y c_Y) \\
&+ T_{Z0Z}a_Z c_Z + T_{0XX}(b_X c_X + b_Y c_Y) + T_{0ZZ}b_Z c_Z.
\end{aligned} \tag{A.11}$$

With further simplifications [13] Eq. A.11 leads to a set of three times three inequalities:

$$\begin{aligned}
L_{(i),1} \leq & \quad |T_{XX0}| \\
& + \sqrt{(|T_{XX0}| + |T_{ZZ0}|)^2 + (\max(|T_{X0X}| + |T_{Z0Z}|) + \max(|T_{0XX}| + |T_{0ZZ}|))^2} \\
L_{(ii),1} \leq & \quad |T_{ZZ0}| + \sqrt{4T_{XX0}^2 + (\max(|T_{X0X}| + |T_{Z0Z}|) + \max(|T_{0XX}| + |T_{0ZZ}|))^2} \\
L_{(iii),1} \leq & \quad |T_{XX0}| + \max(|T_{X0X}| + |T_{Z0Z}|) + \max(|T_{0XX}| + |T_{0ZZ}|)
\end{aligned}$$

$L_{(i),2}$ and $L_{(i),3}$ are obtained from $L_{(i),1}$ by cyclic permutation of the qubits. Analogously, the other four boundary values can be derived from $L_{(ii),1}$ and $L_{(iii),1}$, respectively. Now, the value of L cannot exceed the maximal value of all nine boundary values, i.e.

$$L \leq \max\{L_{(i),1}; L_{(ii),1}; L_{(iii),1}; L_{(i),2}; L_{(ii),2}; L_{(iii),2}; L_{(i),3}; L_{(ii),3}; L_{(iii),3}\}, \tag{A.12}$$

This situation is depicted in Fig. 3.1 where the maximally obtained value for the bound of L is depicted for possible choice of α and β in the state parametrisation of $|\mathrm{W}_3^g\rangle$ (Eq. 3.8). As an example, consider the *no correlation* state $\rho_{\mathrm{NC},3}$. The limit value for L is obtained by Eq. A.12 to be $L \leq 7/3 \approx 2.33$ (see Fig. 3.1), which is equivalent to the value obtained for the bi-separable state $\rho^{\mathrm{bi-prod}} = (\cos(\theta)|PP\rangle - \sin(\theta)|MM\rangle) \otimes |P\rangle$ with $|P\rangle = \sqrt{\frac{1}{2}}(|0\rangle + |1\rangle)$ and $\theta = \frac{1}{2}\tan^{-1}\left(\frac{3}{4}\right)$ in Sec. 3.4.5. Again, for the full derivation, please see [13].

A.6 Hoeffding error calculation

This section demonstrates how to linearly approximate a quantity and calculate the confidence interval using the Hoeffding calculus. Consider therefore the bipartite negativity of a quantum state and construct the linear operator around the $|\phi^+\rangle$ state. To do so, calculate the partial transpose of interest of the respective state

$$\rho^{T_A} = (|\phi^+\rangle\langle\phi^+|)^{T_A} = \frac{1}{2}\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{A.13}$$

Find the eigendecomposition of the state ρ^{T_A}:

$$\underbrace{-\frac{1}{2}}_{\lambda_1}\underbrace{\begin{pmatrix} 0 \\ -1/\sqrt{2} \\ 1/\sqrt{2} \\ 0 \end{pmatrix}}_{|\psi_1\rangle}; \underbrace{\frac{1}{2}}_{\lambda_2}\underbrace{\begin{pmatrix} 0 \\ 1/\sqrt{2} \\ 1/\sqrt{2} \\ 0 \end{pmatrix}}_{|\psi_2\rangle}; \underbrace{\frac{1}{2}}_{\lambda_3}\underbrace{\begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}}_{|\psi_3\rangle}; \underbrace{\frac{1}{2}}_{\lambda_4}\underbrace{\begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}}_{|\psi_4\rangle}. \tag{A.14}$$

The projector onto the negative eigenspace has now to be build. Therefore, the eigenvector corresponding to the eigenvalue $-1/2$ is used for the linear operator L:

$$L = \begin{pmatrix} 0 \\ -1/\sqrt{2} \\ 1/\sqrt{2} \\ 0 \end{pmatrix}\begin{pmatrix} 0 & -1/\sqrt{2} & 1/\sqrt{2} & 0 \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}. \tag{A.15}$$

This projector onto ρ^{T_A}'s negative eigenspace can now be decomposed into M_i^S defined for example in Eq. 2.78. Doing so, one finds

$$\begin{aligned} L = \frac{1}{36}\Big(& -9M_{00}^{XX} + 10M_{01}^{XX} + 10M_{10}^{XX} - 9M_{11}^{XX} \\ & -9M_{00}^{YY} + 10M_{01}^{YY} + 10M_{10}^{YY} - 9M_{11}^{YY} \\ & -9M_{00}^{ZZ} + 10M_{01}^{ZZ} + 10M_{10}^{ZZ} - 9M_{11}^{ZZ} \\ & + \sum_{S'} M_{00}^{S'} + M_{01}^{S'} + M_{10}^{S'} + M_{11}^{S'}\Big) \end{aligned} \tag{A.16}$$

where the occurring summation is running over $S' \in \{XY, XZ, YX, YZ, ZX, ZY\}$. Comparing Eq. A.16 with $L = \sum_{i,S} l_i^S M_i^S$, one finds the coefficients l_i^S. Finally, in Hoeffdings inequality as given in Eq. 4.43 one has to compute $h^2 = \sum_S (\max_i l_i^S - \min_i l_i^S)^2$. For this example, it holds

$$\begin{aligned} h^2 &= \sum_S \left(\max_i l_i^S - \min_i l_i^S\right)^2 \\ &= \left(\underbrace{\frac{1}{36}(10-(-9))}_{XX} + \underbrace{\frac{1}{36}(10-(-9))}_{YY} + \underbrace{\frac{1}{36}(10-(-9))}_{ZZ} + \underbrace{\frac{1}{36}(0-0)}_{XY,\ldots,ZY}\right)^2 \\ &= \left(\frac{3\cdot 19}{36}\right)^2 \approx 2.51. \end{aligned} \tag{A.17}$$

According to $C_L = \sqrt{-\frac{h^2 \ln(1-\gamma)}{2N_S}}$ (Eq. 4.43) for $N_S = 100$ one obtains for the $\gamma = 0.683$ confidence region a value of $C_L \approx 0.120$.

A.7 Variance of matrix elements

The maximally achievable standard deviation of antidiagonal terms of density matrices will be derived here. The matrix element $\rho_{i,2^n-i}$ is composed of the respective numbers of counts when measuring $\{\sigma_x \otimes \cdots \otimes \sigma_x \otimes \sigma_x,\ \sigma_x \otimes \cdots \otimes \sigma_x \otimes \sigma_y,\ \sigma_x \otimes \cdots \otimes \sigma_y \otimes \sigma_x,\ \ldots,\ \sigma_y \otimes \cdots \otimes \sigma_y \otimes \sigma_y\}$. It holds

$$\Delta\rho_{i,2^n-i} = \sqrt{\underbrace{\sum_{i,S} \left(\frac{\partial \rho_{i,2^n-i}}{\partial c_i^S}\right)^2 \left(\Delta c_i^S\right)^2}_{2^n\cdot 2^n \text{terms}}}. \tag{A.18}$$

Assuming white noise, i.e. $c_i^S = N_S/2^n\ \forall i, S$ and same statistics N_S for each setting S Eq. A.18 can be bounded by above with

$$\Delta\rho_{i,2^n-i} \leq \sqrt{2^n 2^n \left(\frac{\partial \rho_{i,2^n-i}}{\partial c_{i'}^{S'}}\right)^2 \left(\Delta c_{i'}^{S'}\right)^2} \tag{A.19}$$

where $c_{i'}^{S'}$ denotes any c_i^S and where S' denotes a setting corresponding to a tensorial product of σ_x and σ_y. Using the partial derivative from Sec. A.2.2, one finds

$$\Delta\rho_{i,2^n-i} \leq 2^n \sqrt{\frac{\left(\Delta c_{i'}^{S'}\right)^2}{4^n N_S^2}} = \sqrt{\frac{2^n}{N_S}}. \tag{A.20}$$

This justifies a $\frac{1}{N_S}$ decreasing behaviour of standard deviation of the matrix elements and furthermore also legitimates the assumption that the radius of the spectral probability density function decreases with $\frac{1}{N_S}$ even if Eq. A.20 is only a rough estimate.

Appendix B

Convex optimisation

The convex optimisation that is used to execute the maximum likelihood and free least-squares fit (see Sec. 4.2.3) is accomplished by a self-coded program written with C++. Here, some basic principles of the implementation will be given as well as some ideas for further modifications will be presented.

B.1 Used Implementation

In the first step, some initialisations have to be done. Therefore, the frequencies are loaded from ASCII coded files. Afterwards, the corresponding projectors are build up. Thus, the program has now a list of 6^n frequencies (f_i^S with 2^n different outcomes i in 3^n basis settings) that are obtained from measured data. Each frequency is furthermore associated to a projector, that identifies the measurement outcome it belongs to.
The used parameterisation of the density matrix ρ is chosen according to Eq. 2.18, such that the operator basis $\{S_i\}_i$ consists of all 4^n different outer products with

$$S_i = \bigotimes_{j=1}^{n} \sigma_{i_j}, \tag{B.1}$$

where i_j denotes the measurement direction of the jth qubit with $i_j \in \{0, x, y, z\}$. Since a Pauli matrix has only two entries different than zero, the tensor product of n Pauli matrices has only $\frac{1}{2^n}$ non-zero entries. Using this and the fact, that only ± 1 and $\pm i$ are possible values for the entries, a lot of memory can be saved to lower the requirements of the algorithm.
To compute the gradients and Hessian matrices according to Eqs. 4.16 and 4.17 for the maximum likelihood estimation (for FLS analogously) the term $\mathrm{tr}(S_i M_k)$ has to be evaluated often. Instead of doing this each time needed, this can be done once at the beginning. Additionally, pre-computing a lookup table to store, which entries are non-zero, causes a further increase of the program's speed, but costs additional memory.
The main steps of the algorithm to reconstruct the state consist of two nested loops. The outer loop - called the *p-loop* - controls the size of the parameter t of the barrier term as given in Eq. 4.11. After each of the 11 iterations t is decreased from the initial value $t_0 = 1$ to $t_{i+1} \to t_i/10$. The inner loop (*m-loop*) is responsible to ensure the convergence for the current function to optimise over. Thus, t is set in the p-loop, while the nested

m-loops converges to the respective optimum. Afterwards, t is further decreased. The m-loop first computes the current density matrix ρ out of the current set of parameters $\{x_i\}_i$. For this ρ the probabilities $\mathrm{tr}(\rho(x)M_k)$ are computed to compare them with the set of measured frequencies. Furthermore, the gradient and the Hessian matrix for the current state ρ are determined. With this knowledge, a Newton step can be performed with [97]

$$\Delta\vec{x} = -\left(\nabla^2 f(\vec{x})\right)^{-1} \vec{\nabla} f(\vec{x}). \tag{B.2}$$

While this is still a valid approach, Eq. B.2 is based on inverting the Hessian matrix with $H = \nabla^2 f(x) \in \mathrm{R}^{4^n \times 4^n}$. Instead of computing the inverse of the Hessian matrix H directly, the implementation makes use of the symmetry and positive definiteness of H, such that it is decomposed into a lower triangular matrix with $H = LL^T$. With this Cholesky decomposition [111] the equation $L\vec{w} = -\vec{\nabla} f(\vec{x})$ is solved. With this, $\vec{w} = -L^{-1}\vec{\nabla} f(\vec{x})$ holds, which is used to solve $L^T\Delta\vec{x} = \vec{w}$. Therefore [97],

$$\Delta\vec{x} = \left(L^T\right)^{-1} w = -\left(L^T\right)^{-1} L^{-1}\vec{\nabla} f(\vec{x}) = -H^{-1}\vec{\nabla} f(\vec{x}) \tag{B.3}$$

is equivalent to Eq. B.2. For further details, including the computational effort, please see [97] and for information concerning the concrete implementation [112].

After applying the Newton step $\vec{x}_i + \Delta\vec{x}$, the density matrix $\rho = \rho(\vec{x})$ is assumed to be closer to the optimal point. Depending on the shape of the problem $\vec{x}_i + \Delta\vec{x}$ can now be a unphysical state, while $\rho(\vec{x}_i)$ was not. To ensure, that one does not leave the space of $\rho \geq 0$, the physicality has to be checked additionally. Therefore, if $\rho(\vec{x}_i + \Delta\vec{x})$ has negative eigenvalues, not the full Newton step is taken into account, but instead its step size is decreased. Thus, this is done for increasing $j = 0, 1, 2, \ldots$, until $\rho(\vec{x}_i + (\frac{1}{2})^j\Delta\vec{x})$ gives a physical state. Finally, the step is accepted as $\vec{x}_{i+1} = \vec{x}_i + \frac{1}{2}^j\Delta\vec{x}$. This method can be augmented by using *backtracking line search* methods [97] to find a good step size and update then $\vec{x}_{i+1} = \vec{x}_i + \alpha\Delta\vec{x}$. This Newton method is employed in the m-loop 10 times at maximum. If the gradient undergoes a given threshold beforehand, the loop can be left, since optimal point has been reached good enough.

To overcome possible numerical problems due to finite accuracy of the used variables with double precision, final steps with quadruple precision can be executed if the gradient could not be minimised sufficiently. The quadruple precision is used up to five times for the last value of the t parameter in the last step of the p-loop. A clearly structured output informs the user if each iteration of the p-loop ended because either the gradient was sufficiently small, the number of maximal iterations was exceeded or the size of the Newton step was too small.

The current implementation of the algorithm is optimised for systems up to six qubits, such that reconstructing six qubit quantum states takes approximately up to 6 to 8 minutes using a quad core CPU of type AMD Phenom$^{\mathrm{tm}}$ *II X*4 840 clocked with 3.20 GHz and with 8 Gb RAM running openSuSE “Asparagus” (ver. 12.1). Obviously, the speed highly depends on the state to reconstruct.

B.2 Other methods for optimisation

Other approaches besides the described Newton method are possible. For the Newton step as given in Eq. B.2, one has to compute the Hessian matrix of the state ρ and finally invert it either directly or use its Cholesky decomposition. This can be avoided by using *Quasi-Newton* methods [113, 114] like the *Broyden-Fletcher-Goldgarb-Shanno (BFGS) method* [115]. Instead of evaluating the Hessian matrix for each state $\rho(\vec{x}_i)$ anew, an approximation to its inverse is computed once and afterwards only updated. A direct implementation of this would cause problems each time when the t parameter is decreased by an order of magnitude. Hence, a better approach to implement the BFGS method could be to decrease t more often, but with less change each time. Thus, the problem to solve does not change dramatically for each p-iteration, but is varied only slightly. Therefore, continuously updating H^{-1} could be possible. The memory requirements can be reduced by using memory optimised versions like the *L-BFGS* method [116].
Another possibility to avoid a Cholesky decomposition or directly inverting the Hessian matrix is by using a *conjugate gradient* method [117]. This method shows its advantages best when compared with a *steepest descent* method as shown in Fig. B.1. Steepest descent methods cause a continuous change of the direction of descent due to perpetual alteration of the gradient. This can lead to a bouncing behaviour of the optimisation. The conjugate gradient methods compute the direction for the next step in dependence of the previous direction to avoid this jiggling. While different versions of conjugate gradient methods like the Fletcher-Reeves (FR), the Polak-Ribière (PR) and the Hesteness-Stiefel (HS) exist, a general implementation can be found in [112]. Since conjugate gradient methods do not need any second derivatives, these methods may be a way to prepare the quantum state reconstruction algorithm to handle states with more than six qubits as well.

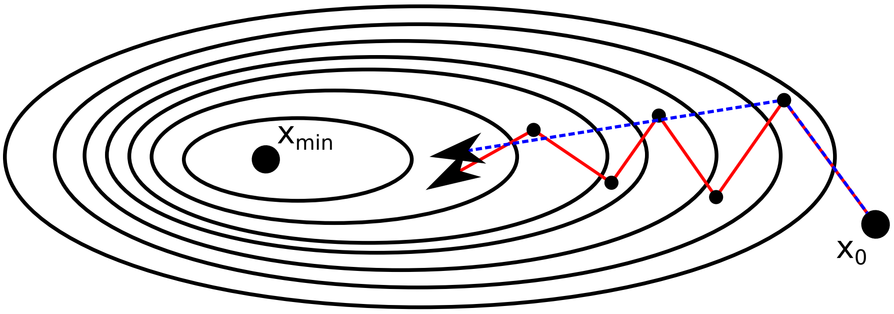

Figure B.1: Comparison of conjugate gradient and steepest descent methods [112]. The contour represents the function to optimise over. The algorithm may start at x_0 and is looking for $x_{\min}$. While the steepest descent methods use the gradient of the current evaluation point x_i to find the next x_{i+1} (red curve), conjugate gradient methods (blue) find the next x_{i+1} by using a direction composed by the gradients evaluated at x_0, x_1, …, x_i.

Appendix C

Publication list

[50]
Wiesław Laskowski, Christian Schwemmer, Daniel Richart, *Lukas Knips*, Tomasz Paterek and Harald Weinfurter
Optimized state independent entanglement detection based on geometrical threshold criterion
Physical Review A
Status: *accepted*

[13]
Π(Wiesław Laskowski, Tomasz Paterek, Christian Schwemmer, *Lukas Knips* and Harald Weinfurter), where Π denotes a permutation
t.b.a.
Status: *in preparation*

[23]
Christian Schwemmer, *Lukas Knips*, Daniel Richart and Harald Weinfurter; Tobias Moroder, Matthias Kleinmann and Otfried Gühne
Systematic errors in current quantum state tomography tools
Status: *in preparation*

Bibliography

[1] Ryszard Horodecki, Paweł Horodecki, Michał Horodecki, and Karol Horodecki. Quantum entanglement. *Rev. Mod. Phys.*, 81:865–942, Jun 2009. doi: 10.1103/RevModPhys.81.865. URL `http://link.aps.org/doi/10.1103/RevModPhys.81.865`.

[2] Jian-Wei Pan, Zeng-Bing Chen, Chao-Yang Lu, Harald Weinfurter, Anton Zeilinger, and Marek Żukowski. Multiphoton entanglement and interferometry. *Rev. Mod. Phys.*, 84:777–838, May 2012. doi: 10.1103/RevModPhys.84.777. URL `http://link.aps.org/doi/10.1103/RevModPhys.84.777`.

[3] Roland Krischek, Christian Schwemmer, Witlef Wieczorek, Harald Weinfurter, Philipp Hyllus, Luca Pezzé, and Augusto Smerzi. Useful multiparticle entanglement and sub-shot-noise sensitivity in experimental phase estimation. *Phys. Rev. Lett.*, 107:080504, Aug 2011. doi: 10.1103/PhysRevLett.107.080504. URL `http://link.aps.org/doi/10.1103/PhysRevLett.107.080504`.

[4] V. Giovannetti, S. Lloyd, and L. Maccone. Advances in quantum metrology. *Nature Photonics*, 5:222–229, April 2011. doi: 10.1038/nphoton.2011.35. URL `http://www.nature.com/nphoton/journal/v5/n4/pdf/nphoton.2011.35.pdf`.

[5] M. A. Nielsen and I. L. Chuang. *Quantum computation and quantum information*. Cambridge University Press, 2000.

[6] Artur Ekert and Richard Jozsa. Quantum computation and shor's factoring algorithm. *Rev. Mod. Phys.*, 68:733–753, Jul 1996. doi: 10.1103/RevModPhys.68.733. URL `http://link.aps.org/doi/10.1103/RevModPhys.68.733`.

[7] Andrew M. Childs and Wim van Dam. Quantum algorithms for algebraic problems. *Rev. Mod. Phys.*, 82:1–52, Jan 2010. doi: 10.1103/RevModPhys.82.1. URL `http://link.aps.org/doi/10.1103/RevModPhys.82.1`.

[8] Nicolas Gisin, Grégoire Ribordy, Wolfgang Tittel, and Hugo Zbinden. Quantum cryptography. *Rev. Mod. Phys.*, 74:145–195, Mar 2002. doi: 10.1103/RevModPhys.74.145. URL `http://link.aps.org/doi/10.1103/RevModPhys.74.145`.

[9] Sebastian Nauerth, Florian Moll, Markus Rau, Christian Fuchs, Joachim Horwath, Stefan Frick, and Harald Weinfurter. Air-to-ground quantum communication. *Nature Photonics*, 7:382–386, 2013. doi: 10.1038/nphoton.2013.46. URL `http://dx.doi.org/10.1038/nphoton.2013.46`.

[10] Otfried Gühne and Géza Tóth. Entanglement detection. *Physics Reports*, 474(16):1 – 75, 2009. ISSN 0370-1573. doi: http://dx.doi.org/10.1016/j.physrep.2009.02.004. URL `http://www.sciencedirect.com/science/article/pii/S0370157309000623`.

[11] Dagomir Kaszlikowski, Aditi Sen(De), Ujjwal Sen, Vlatko Vedral, and Andreas Winter. Quantum correlation without classical correlations. *Phys. Rev. Lett.*, 101: 070502, Aug 2008. doi: 10.1103/PhysRevLett.101.070502. URL `http://link.aps.org/doi/10.1103/PhysRevLett.101.070502`.

[12] Wiesław Laskowski, Marcin Markiewicz, Tomasz Paterek, and Marcin Wieśniak. Incompatible local hidden-variable models of quantum correlations. *Phys. Rev. A*, 86:032105, Sep 2012. doi: 10.1103/PhysRevA.86.032105. URL `http://link.aps.org/doi/10.1103/PhysRevA.86.032105`.

[13] Tomasz Paterek, Wiesław Laskowski, Christian Schwemmer, Lukas Knips, and Harald Weinfurter. Experimental demonstration of tripartite entanglement without tripartite correlations. title may be changed. *in preparation*, 2013.

[14] Matteo Paris and Jaroslav Řeháček, editors. *Quantum State Estimation (Lecture Notes in Physics, vol.649)*, volume 649 of *Lecture Notes in Physics*. Springer, 2004. doi: 10.1007/b98673. URL `http://www.springerlink.com/openurl.asp?genre=issue&issn=1616-6361&volume=649`.

[15] Daniel F. V. James, Paul G. Kwiat, William J. Munro, and Andrew G. White. Measurement of qubits. *Phys. Rev. A*, 64:052312, Oct 2001. doi: 10.1103/PhysRevA.64.052312. URL `http://link.aps.org/doi/10.1103/PhysRevA.64.052312`.

[16] Z. Hradil. Quantum-state estimation. *Phys. Rev. A*, 55:R1561–R1564, Mar 1997. doi: 10.1103/PhysRevA.55.R1561. URL `http://link.aps.org/doi/10.1103/PhysRevA.55.R1561`.

[17] K. Banaszek, G. M. D'Ariano, M. G. A. Paris, and M. F. Sacchi. Maximum-likelihood estimation of the density matrix. *Phys. Rev. A*, 61:010304, Dec 1999. doi: 10.1103/PhysRevA.61.010304. URL `http://link.aps.org/doi/10.1103/PhysRevA.61.010304`.

[18] Robin Blume-Kohout. Hedged maximum likelihood quantum state estimation. *Phys. Rev. Lett.*, 105:200504, Nov 2010. doi: 10.1103/PhysRevLett.105.200504. URL `http://link.aps.org/doi/10.1103/PhysRevLett.105.200504`.

[19] David Gross, Yi-Kai Liu, Steven T. Flammia, Stephen Becker, and Jens Eisert. Quantum state tomography via compressed sensing. *Phys. Rev. Lett.*, 105:150401, Oct 2010. doi: 10.1103/PhysRevLett.105.150401. URL `http://link.aps.org/doi/10.1103/PhysRevLett.105.150401`.

[20] Matthias Christandl and Renato Renner. Reliable quantum state tomography. *Phys. Rev. Lett.*, 109:120403, Sep 2012. doi: 10.1103/PhysRevLett.109.120403. URL `http://link.aps.org/doi/10.1103/PhysRevLett.109.120403`.

[21] Tobias Moroder, Philipp Hyllus, Géza Tóth, Christian Schwemmer, Alexander Niggebaum, Stefanie Gaile, Otfried Gühne, and Harald Weinfurter. Permutationally invariant state reconstruction. *New Journal of Physics*, 14(10):105001, 2012. URL `http://stacks.iop.org/1367-2630/14/i=10/a=105001`.

[22] B. Efron and R. Tibshirani. Bootstrap Methods for Standard Errors, Confidence Intervals, and Other Measures of Statistical Accuracy. *Statistical Science*, 1(1):54–75, 1986. URL `http://projecteuclid.org/DPubS/Repository/1.0/Disseminate?view=body&id=pdf_1&handle=euclid.ss/1177013815`.

[23] Schwemmer, Knips, Richart, Moroder, Kleinmann, Gühne, and Weinfurter. Systematic errors in current quantum state tomography tools. *in preparation*, 2013.

[24] Wassily Hoeffding. Probability inequalities for sums of bounded random variables. *Journal of the American Statistical Association*, 58(301):13–30, 1963. doi: 10.1080/01621459.1963.10500830. URL `http://www.tandfonline.com/doi/abs/10.1080/01621459.1963.10500830`.

[25] C. E. Shannon. A mathematical theory of communication. *The Bell System Technical Journal*, 27:379–423, 7-10 1948.

[26] Benjamin Schumacher. Quantum coding. *Phys. Rev. A*, 51:2738–2747, Apr 1995. doi: 10.1103/PhysRevA.51.2738. URL `http://link.aps.org/doi/10.1103/PhysRevA.51.2738`.

[27] P. A. M. Dirac. A new notation for quantum mechanics. *Mathematical Proceedings of the Cambridge Philosophical Society*, 35:416–418, 6 1939. ISSN 1469-8064. doi: 10.1017/S0305004100021162. URL `http://journals.cambridge.org/article_S0305004100021162`.

[28] Nikolai Kiesel. Experiments on Multiphoton Entanglement. *PhD Thesis, Ludwig-Maximilians-Universität München*, 2007. URL `http://xqp.physik.uni-muenchen.de/publications/files/theses_phd/phd_kiesel.pdf`.

[29] F. Bloch. Nuclear induction. *Phys. Rev.*, 70:460–474, Oct 1946. doi: 10.1103/PhysRev.70.460. URL `http://link.aps.org/doi/10.1103/PhysRev.70.460`.

[30] Pavel Trojek. Efficient Generation of Photonic Entanglement and Multiparty Quantum Communication. *PhD Thesis, Ludwig-Maximilians-Universität München*, 2007. URL `http://xqp.physik.uni-muenchen.de/publications/files/theses_phd/phd_trojek.pdf`.

[31] Robert Gilmore. *Lie groups, Lie algebras, and some of their applications*. John Wiley & Sons, Inc., 1974. ISBN 0471301795.

[32] Asher Peres. *Quantum Theory: Concepts and Methods*. Kluwer Academic, 1995. ISBN 0792336321. URL `http://cdn.preterhuman.net/texts/science_and_technology/physics/Quantum_theory/Quantum%20Theory%20-%20Concepts%20&%20Methods%20-%20A.%20Peres.pdf`.

[33] Christian Schmid. Multi-Photon entanglement and applications in quantum information. *PhD Thesis, Ludwig-Maximilians-Universität München*, 2008. URL `http://xqp.physik.uni-muenchen.de/publications/files/theses_phd/phd_schmid.pdf`.

[34] E. Schrödinger. Die gegenwärtige Situation in der Quantenmechanik. *Naturwissenschaften*, 23(50):844–849, 1935. ISSN 0028-1042. doi: 10.1007/BF01491987. URL `http://dx.doi.org/10.1007/BF01491987`.

[35] J. Yin, Y. Cao, H.-L. Yong, J.-G. Ren, H. Liang, S.-K. Liao, F. Zhou, C. Liu, Y.-P. Wu, G.-S. Pan, Q. Zhang, C.-Z. Peng, and J.-W. Pan. Bounding the speed of ‘spooky action at a distance’. *ArXiv e-prints*, March 2013.

[36] Reinhard F. Werner. Quantum states with einstein-podolsky-rosen correlations admitting a hidden-variable model. *Phys. Rev. A*, 40:4277–4281, Oct 1989. doi: 10.1103/PhysRevA.40.4277. URL `http://link.aps.org/doi/10.1103/PhysRevA.40.4277`.

[37] Albert Einstein. Physik und Realität. *Journal of the Franklin Institute*, 221(3):313 – 347, 1936. ISSN 0016-0032. doi: http://dx.doi.org/10.1016/S0016-0032(36)91045-1. URL `http://www.sciencedirect.com/science/article/pii/S0016003236910451`.

[38] John S. Bell. On the Einstein-Podolsky-Rosen Paradox. *Physics*, 1:195–200, 1964.

[39] Gregor Weihs, Thomas Jennewein, Christoph Simon, Harald Weinfurter, and Anton Zeilinger. Violation of bell's inequality under strict einstein locality conditions. *Phys. Rev. Lett.*, 81:5039–5043, Dec 1998. doi: 10.1103/PhysRevLett.81.5039. URL `http://link.aps.org/doi/10.1103/PhysRevLett.81.5039`.

[40] John S. Bell. Speakable and Unspeakable in Quantum Mechanics. 1987. URL `http://philosophyfaculty.ucsd.edu/faculty/wuthrich/GSSPP09/Files/BellJohnS1981Speakable_BertlmannsSocks.pdf`.

[41] David Lindley. What's wrong with quantum mechanics? *Phys. Rev. Focus*, 16:10, Sep 2005. doi: 10.1103/PhysRevFocus.16.10. URL `http://link.aps.org/doi/10.1103/PhysRevFocus.16.10`.

[42] P. G. Kwiat and L. Hardy. The mystery of the quantum cakes. *American Journal of Physics*, 68(1):33–36, 2000. doi: 10.1119/1.19369. URL `http://dx.doi.org/10.1119/1.19369`.

[43] H. Aschauer, J. Calsamiglia, M. Hein, and Hans J. Briegel. Local invariants for multi-partite entangled states allowing for a simple entanglement criterion. *Quantum Inf. Comput.*, 4:383–395, 2004. URL `http://www.rintonpress.com/xqic4/qic-4-5/383-395.pdf`.

[44] Piotr Badziag, Časlav Brukner, Wiesław Laskowski, Tomasz Paterek, and Marek Żukowski. Experimentally friendly geometrical criteria for entanglement. *Phys. Rev. Lett.*, 100:140403, Apr 2008. doi: 10.1103/PhysRevLett.100.140403. URL `http://link.aps.org/doi/10.1103/PhysRevLett.100.140403`.

[45] W. Dür, G. Vidal, and J. I. Cirac. Three qubits can be entangled in two inequivalent ways. *Phys. Rev. A*, 62:062314, Nov 2000. doi: 10.1103/PhysRevA.62.062314. URL `http://link.aps.org/doi/10.1103/PhysRevA.62.062314`.

[46] Charles H. Bennett, Sandu Popescu, Daniel Rohrlich, John A. Smolin, and Ashish V. Thapliyal. Exact and asymptotic measures of multipartite pure-state entanglement. *Phys. Rev. A*, 63:012307, Dec 2000. doi: 10.1103/PhysRevA.63.012307. URL `http://link.aps.org/doi/10.1103/PhysRevA.63.012307`.

[47] M. B. Plenio and S. Virmani. An introduction to entanglement measures. *eprint arXiv:quant-ph/0504163*, April 2005.

[48] Wiesław Laskowski, Marcin Markiewicz, Tomasz Paterek, and Marek Żukowski. Correlation-tensor criteria for genuine multiqubit entanglement. *Phys. Rev. A*, 84:062305, Dec 2011. doi: 10.1103/PhysRevA.84.062305. URL `http://link.aps.org/doi/10.1103/PhysRevA.84.062305`.

[49] Erhard Schmidt. Zur Theorie der linearen und nichtlinearen Integralgleichungen. *Mathematische Annalen*, 63(4):433–476, 1907. ISSN 0025-5831. doi: 10.1007/BF01449770. URL http://dx.doi.org/10.1007/BF01449770.

[50] Wiesław Laskowski, Christian Schwemmer, Daniel Richart, Lukas Knips, Tomasz Paterek, and Harald Weinfurter. Optimized state-independent entanglement detection based on a geometrical threshold criterion. *Phys. Rev. A*, 88:022327, Aug 2013. doi: 10.1103/PhysRevA.88.022327. URL http://link.aps.org/doi/10.1103/PhysRevA.88.022327.

[51] Roland Krischek. Experimental multi-photon entanglement and quantum enhanced metrology. *PhD Thesis, Ludwig-Maximilians-Universität München*, 2011. URL http://xqp.physik.uni-muenchen.de/publications/files/theses_phd/phd_krischek.pdf.

[52] Asher Peres. Separability criterion for density matrices. *Phys. Rev. Lett.*, 77: 1413–1415, Aug 1996. doi: 10.1103/PhysRevLett.77.1413. URL http://link.aps.org/doi/10.1103/PhysRevLett.77.1413.

[53] G. Vidal and R. F. Werner. Computable measure of entanglement. *Phys. Rev. A*, 65:032314, Feb 2002. doi: 10.1103/PhysRevA.65.032314. URL http://link.aps.org/doi/10.1103/PhysRevA.65.032314.

[54] Karol Życzkowski, Paweł Horodecki, Anna Sanpera, and Maciej Lewenstein. Volume of the set of separable states. *Phys. Rev. A*, 58:883–892, Aug 1998. doi: 10.1103/PhysRevA.58.883. URL http://link.aps.org/doi/10.1103/PhysRevA.58.883.

[55] A. Uhlmann. The transition probability in the state space of a $\star$-algebra . *Reports on Mathematical Physics*, 9(2):273 – 279, 1976. ISSN 0034-4877. doi: 10.1016/0034-4877(76)90060-4. URL http://www.sciencedirect.com/science/article/pii/0034487776900604.

[56] Alexei Gilchrist, Nathan K. Langford, and Michael A. Nielsen. Distance measures to compare real and ideal quantum processes. *Phys. Rev. A*, 71:062310, Jun 2005. doi: 10.1103/PhysRevA.71.062310. URL http://link.aps.org/doi/10.1103/PhysRevA.71.062310.

[57] Jennifer L. Dodd and Michael A. Nielsen. Simple operational interpretation of the fidelity of mixed states. *Phys. Rev. A*, 66:044301, Oct 2002. doi: 10.1103/PhysRevA.66.044301. URL http://link.aps.org/doi/10.1103/PhysRevA.66.044301.

[58] Ingemar Bengtsson and Karol Zyczkowski. *Geometry of Quantum States: An Introduction to Quantum Entanglement*. Cambridge University Press, January 2008. ISBN 052189140X. URL http://www.amazon.com/exec/obidos/redirect?tag=citeulike07-20&path=ASIN/052189140X.

[59] Jerzy Dajka, Jerzy Łuczka, and Peter Hänggi. Distance between quantum states in the presence of initial qubit-environment correlations: A comparative study. *Phys. Rev. A*, 84:032120, Sep 2011. doi: 10.1103/PhysRevA.84.032120. URL http://link.aps.org/doi/10.1103/PhysRevA.84.032120.

[60] John Neumann. *Mathematische Grundlagen der Quantenmechanik*. Springer, Berlin New York, 1955. ISBN 3540592075.

[61] A. Einstein, B. Podolsky, and N. Rosen. Can quantum-mechanical description of physical reality be considered complete? *Phys. Rev.*, 47:777–780, May 1935. doi: 10.1103/PhysRev.47.777. URL http://link.aps.org/doi/10.1103/PhysRev.47.777.

[62] Michel Le Bellac. *Quantum Physics*. Cambridge University Press, 2006. doi: 10.2277/0521852773.

[63] D. M. Greenberger, M. Horne, and A. Zeilinger. *Bell's Theorem, Quantum Theory and Conceptions of the Universe*. Kluwer Academic, Dordrecht, 1989.

[64] L. Novo, T. Moroder, and O. Gühne. Genuine multiparticle entanglement of permutationally invariant states. *ArXiv e-prints*, February 2013. URL http://arxiv.org/abs/1302.4100.

[65] R. H. Dicke. Coherence in spontaneous radiation processes. *Phys. Rev.*, 93:99–110, Jan 1954. doi: 10.1103/PhysRev.93.99. URL http://link.aps.org/doi/10.1103/PhysRev.93.99.

[66] Géza Tóth. Detection of multipartite entanglement in the vicinity of symmetric dicke states. *J. Opt. Soc. Am. B*, 24(2):275–282, Feb 2007. doi: 10.1364/JOSAB.24.000275. URL http://josab.osa.org/abstract.cfm?URI=josab-24-2-275.

[67] Wiesław Laskowski, Marcin Markiewicz, Tomasz Paterek, and Ryszard Weinar. Entanglement witnesses with variable number of local measurements. *Phys. Rev. A*, 88:022304, Aug 2013. doi: 10.1103/PhysRevA.88.022304. URL http://link.aps.org/doi/10.1103/PhysRevA.88.022304.

[68] M. Hein, J. Eisert, and H. J. Briegel. Multiparty entanglement in graph states. *Phys. Rev. A*, 69:062311, Jun 2004. doi: 10.1103/PhysRevA.69.062311. URL http://link.aps.org/doi/10.1103/PhysRevA.69.062311.

[69] Wei-Bo Gao, Xiao-Qi Zhou, Jin Zhang, Tao Yang, and Jian-Wei Pan. A proof-of-principle experiment of eliminating photon-loss errors in cluster states. *New Journal of Physics*, 10(5):055003, 2008. URL http://stacks.iop.org/1367-2630/10/i=5/a=055003.

[70] Hans J. Briegel and Robert Raussendorf. Persistent entanglement in arrays of interacting particles. *Phys. Rev. Lett.*, 86:910–913, Jan 2001. doi: 10.1103/PhysRevLett.86.910. URL http://link.aps.org/doi/10.1103/PhysRevLett.86.910.

[71] P. Walther, K. J. Resch, T. Rudolph, E. Schenck, H. Weinfurter, V. Vedral, M. Aspelmeyer, and A. Zeilinger. Experimental one-way quantum computing. *Nature*, 434(7030):169–176, 2005. URL http://www.nature.com/nature/journal/v434/n7030/full/nature03347.html.

[72] Robert Raussendorf and Hans J. Briegel. A one-way quantum computer. *Phys. Rev. Lett.*, 86:5188–5191, May 2001. doi: 10.1103/PhysRevLett.86.5188. URL http://link.aps.org/doi/10.1103/PhysRevLett.86.5188.

[73] John A. Smolin. Four-party unlockable bound entangled state. *Phys. Rev. A*, 63:032306, Feb 2001. doi: 10.1103/PhysRevA.63.032306. URL http://link.aps.org/doi/10.1103/PhysRevA.63.032306.

[74] Elias Amselem and Mohamed Bourennane. Experimental four-qubit bound entanglement. *Nature Physics*, 5:748–752, 2009. doi: 10.1038/nphys1372. URL http://www.nature.com/nphys/journal/v5/n10/abs/nphys1372.html.

[75] Philipp Hyllus, Wiesław Laskowski, Roland Krischek, Christian Schwemmer, Witlef Wieczorek, Harald Weinfurter, Luca Pezzé, and Augusto Smerzi. Fisher information and multiparticle entanglement. *Phys. Rev. A*, 85:022321, Feb 2012. doi: 10.1103/PhysRevA.85.022321. URL http://link.aps.org/doi/10.1103/PhysRevA.85.022321.

[76] F. Mezzadri. How to generate random matrices from the classical compact groups. *ArXiv Mathematical Physics e-prints*, September 2006. URL http://arxiv.org/abs/math-ph/0609050.

[77] Eugene Hecht. *Optics (5th Edition)*. Addison-Wesley, 5 edition, 2002. ISBN 9783486588613.

[78] B. Saleh and M. Teich. *Fundamentals of Photonics*. John Wiley & Sons, Inc., 1991.

[79] Witlef Wieczorek. Multi-Photon Entanglement. *PhD Thesis, Ludwig-Maximilians-Universität München*, 2009. URL http://xqp.physik.uni-muenchen.de/publications/files/theses_phd/phd_wieczorek.pdf.

[80] Pieter Kok and Samuel L. Braunstein. Postselected versus nonpostselected quantum teleportation using parametric down-conversion. *Phys. Rev. A*, 61:042304, Mar 2000. doi: 10.1103/PhysRevA.61.042304. URL http://link.aps.org/doi/10.1103/PhysRevA.61.042304.

[81] Harald Weinfurter and Marek Żukowski. Four-photon entanglement from down-conversion. *Phys. Rev. A*, 64:010102, Jun 2001. doi: 10.1103/PhysRevA.64.010102. URL http://link.aps.org/doi/10.1103/PhysRevA.64.010102.

[82] T.W. Hansch and B. Couillaud. Laser frequency stabilization by polarization spectroscopy of a reflecting reference cavity. *Optics Communications*, 35(3):441 – 444, 1980. ISSN 0030-4018. doi: http://dx.doi.org/10.1016/0030-4018(80)90069-3. URL http://www.sciencedirect.com/science/article/pii/0030401880900693.

[83] Th Udem, R. Holzwarth, and T. W. Hansch. Optical frequency metrology. *Nature*, 416(6877):233–237, March 2002. ISSN 0028-0836. doi: 10.1038/416233a. URL http://dx.doi.org/10.1038/416233a.

[84] Robert Szipöcs, Kárpát Ferencz, Christian Spielmann, and Ferenc Krausz. Chirped multilayer coatings for broadband dispersion control in femtosecond lasers. *Opt. Lett.*, 19(3):201–203, Feb 1994. doi: 10.1364/OL.19.000201. URL http://ol.osa.org/abstract.cfm?URI=ol-19-3-201.

[85] Roland Krischek, Witlef Wieczorek, Akira Ozawa, Nikolai Kiesel, Patrick Michelberger, Thomas Udem, and Harald Weinfurter. Ultraviolet enhancement cavity for ultrafast nonlinear optics and high-rate multiphoton entanglement experiments. *Nature Photonics*, 4:170–173, 2010. doi: 10.1038/nphoton.2009.286. URL http://www.nature.com/nphoton/journal/v4/n3/full/nphoton.2009.286.html.

[86] Witlef Wieczorek, Roland Krischek, Nikolai Kiesel, Patrick Michelberger, Géza Tóth, and Harald Weinfurter. Experimental entanglement of a six-photon symmetric dicke state. *Phys. Rev. Lett.*, 103:020504, Jul 2009. doi: 10.1103/PhysRevLett.103.020504. URL `http://link.aps.org/doi/10.1103/PhysRevLett.103.020504`.

[87] H. Häffner, W. Hänsel, C. F. Roos, J. Benhelm, D. Chek-Al-Kar, M. Chwalla, T. Krber, U. D. Rapol, M. Riebe, P. O. Schmidt, C. Becher, O. Gühne, W. Dür, and R. Blatt. Scalable multiparticle entanglement of trapped ions. *Nature*, 438:643–646, 2005. doi: 10.1038/nature04279. URL `http://www.nature.com/nature/journal/v438/n7068/full/nature04279.html`.

[88] N. Kiesel, C. Schmid, G. Tóth, E. Solano, and H. Weinfurter. Experimental observation of four-photon entangled dicke state with high fidelity. *Phys. Rev. Lett.*, 98:063604, Feb 2007. doi: 10.1103/PhysRevLett.98.063604. URL `http://link.aps.org/doi/10.1103/PhysRevLett.98.063604`.

[89] A. Fedorov, L. Steffen, M. Baur, M. P. da Silva, and A. Wallraff. Implementation of a Toffoli gate with superconducting circuits. *Nature*, 481(7380):170–172, January 2012. ISSN 0028-0836. doi: 10.1038/nature10713. URL `http://dx.doi.org/10.1038/nature10713`.

[90] K. J. Resch, P. Walther, and A. Zeilinger. Full characterization of a three-photon greenberger-horne-zeilinger state using quantum state tomography. *Phys. Rev. Lett.*, 94:070402, Feb 2005. doi: 10.1103/PhysRevLett.94.070402. URL `http://link.aps.org/doi/10.1103/PhysRevLett.94.070402`.

[91] Stefanie Barz, Elham Kashefi, Anne Broadbent, Joseph F. Fitzsimons, Anton Zeilinger, and Philip Walther. Demonstration of blind quantum computing. *Science*, 335(6066):303–308, 2012. doi: 10.1126/science.1214707. URL `http://www.sciencemag.org/content/335/6066/303.abstract`.

[92] K.R.W. Jones. Principles of quantum inference. *Annals of Physics*, 207(1):140 – 170, 1991. ISSN 0003-4916. doi: http://dx.doi.org/10.1016/0003-4916(91)90182-8. URL `http://www.sciencedirect.com/science/article/pii/0003491691901828`.

[93] Max Born. Zur Quantenmechanik der Stoßvorgänge. *Zeitschrift für Physik*, 37(12):863–867, 1926. ISSN 0044-3328. doi: 10.1007/BF01397477. URL `http://dx.doi.org/10.1007/BF01397477`.

[94] Tobias Moroder, Matthias Kleinmann, Philipp Schindler, Thomas Monz, Otfried Gühne, and Rainer Blatt. Certifying systematic errors in quantum experiments. *Phys. Rev. Lett.*, 110:180401, Apr 2013. doi: 10.1103/PhysRevLett.110.180401. URL `http://link.aps.org/doi/10.1103/PhysRevLett.110.180401`.

[95] S. Glancy, E. Knill, and M. Girard. Gradient-based stopping rules for maximum-likelihood quantum-state tomography. *New Journal of Physics*, 14(9):095017, 2012. URL `http://stacks.iop.org/1367-2630/14/i=9/a=095017`.

[96] Jaroslav Řeháček, Zdeněk Hradil, E. Knill, and A. I. Lvovsky. Diluted maximum-likelihood algorithm for quantum tomography. *Phys. Rev. A*, 75:042108, Apr 2007. doi: 10.1103/PhysRevA.75.042108. URL `http://link.aps.org/doi/10.1103/PhysRevA.75.042108`.

[97] Stephen Boyd and Lieven Vandenberghe. *Convex Optimization*. Cambridge University Press, mar 2004. ISBN 0521833787. URL `http://www.stanford.edu/~boyd/cvxbook/bv_cvxbook.pdf`.

[98] B. Efron and R. J. Tibshirani. An Introduction to the Bootstrap. 1993.

[99] Nicholas Metropolis and S. Ulam. The monte carlo method. *Journal of the American Statistical Association*, 44(247):335–341, 1949. doi: 10.1080/01621459.1949.10483310. URL `http://www.tandfonline.com/doi/abs/10.1080/01621459.1949.10483310`.

[100] E.E. Bassett, J.M. Bremner, I.T. Jolliffe, B. Jones, B.J.T. Morgan, and P.M. North. *Statistics: Problems and Solutions*. World Scientific Publishing Company Incorporated, 1986. ISBN 981024293X. URL `http://books.google.de/books?id=GMIa_8ZGX9kC`.

[101] C.B. Lang and N. Pucker. *Mathematische Methoden in der Physik*. Spektrum Akademischer Verlag, 2005. ISBN 9783827415585. URL `http://books.google.de/books?id=gXqrAAAACAAJ`.

[102] M. Fisz. *Wahrscheinlichkeitsrechnung und mathematische Statistik*. Hochschulbücher für Mathematik, Bd. 40. VEB Deutscher Verlag der Wissenschaften, 1958.

[103] D. Petz and C. Ghinea. Introduction to Quantum Fisher Information. In R. Rebolledo and M. Orszag, editors, *Quantum Probability and Related Topics*, pages 261–281, January 2011. doi: 10.1142/9789814338745_0015.

[104] Marcus P. da Silva, Olivier Landon-Cardinal, and David Poulin. Practical characterization of quantum devices without tomography. *Phys. Rev. Lett.*, 107:210404, Nov 2011. doi: 10.1103/PhysRevLett.107.210404. URL `http://link.aps.org/doi/10.1103/PhysRevLett.107.210404`.

[105] Steven T. Flammia and Yi-Kai Liu. Direct fidelity estimation from few pauli measurements. *Phys. Rev. Lett.*, 106:230501, Jun 2011. doi: 10.1103/PhysRevLett.106.230501. URL `http://link.aps.org/doi/10.1103/PhysRevLett.106.230501`.

[106] Eugene P. Wigner. On the distribution of the roots of certain symmetric matrices. *Annals of Mathematics*, 67(2):pp. 325–327, 1958. ISSN 0003486X. URL `http://www.jstor.org/stable/1970008`.

[107] M. L. Mehta. *Random matrices*. Academic Press, Boston, 1991. ISBN 0124880517.

[108] Michael K.-H. Kiessling and Herbert Spohn. A note on the eigenvalue density of random matrices. *Communications in Mathematical Physics*, 199(3):683–695, 1999. ISSN 0010-3616. doi: 10.1007/s002200050516. URL `http://dx.doi.org/10.1007/s002200050516`.

[109] Freeman J. Dyson. A brownian-motion model for the eigenvalues of a random matrix. *Journal of Mathematical Physics*, 3(6):1191–1198, 1962. doi: 10.1063/1.1703862. URL `http://link.aip.org/link/?JMP/3/1191/1`.

[110] Friedrich Liese. *Statistical decision theory estimation, testing, and selection*. Springer, New York, 2008. ISBN 9780387731940.

[111] Gene Golub. *Matrix computations*. Johns Hopkins University Press, Baltimore, 1996. ISBN 9780801854149.

[112] William Press. *Numerical Recipes: The Art of Scientific Computing*. Cambridge University Press, Cambridge, UK New York, 2007. ISBN 0521880688. URL `http://www.nr.com`.

[113] C. G. Broyden. Quasi-Newton Methods and their Application to Function Minimisation. *Mathematics of Computation*, 21(99):368–381, 1967. doi: 10.2307/2003239. URL `http://dx.doi.org/10.2307/2003239`.

[114] J. F. Bonnans. *Numerical optimization theoretical and practical aspects*. Springer, Berlin New York, 2006. ISBN 9783540354451. URL `http://link.springer.com/book/10.1007/978-3-540-35447-5/page/1`.

[115] David Luenberger. *Linear and nonlinear programming*. Springer, New York, 2008. ISBN 9780387745022. URL `http://link.springer.com/book/10.1007/978-0-387-74503-9/page/1`.

[116] R. Byrd, P. Lu, J. Nocedal, and C. Zhu. A limited memory algorithm for bound constrained optimization. *SIAM Journal on Scientific Computing*, 16(5):1190–1208, 1995. doi: 10.1137/0916069. URL `http://epubs.siam.org/doi/abs/10.1137/0916069`.

[117] R. Fletcher and C. M. Reeves. Function minimization by conjugate gradients. *The Computer Journal*, 7(2):149–154, 1964. doi: 10.1093/comjnl/7.2.149. URL `http://comjnl.oxfordjournals.org/content/7/2/149.abstract`.