

---

# Efficient Quantum State Analysis and Entanglement Detection

Lukas Martin Knips

---



München 2019



---

# Efficient Quantum State Analysis and Entanglement Detection

Lukas Martin Knips

---

Dissertation  
an der Fakultät für Physik  
der Ludwig-Maximilians-Universität  
München

vorgelegt von  
Lukas Martin Knips  
aus Bonn

München, den 21. Mai 2019

Erstgutachter: Prof. Dr. Harald Weinfurter

Zweitgutachter: Prof. Dr. Otfried Gühne

Tag der mündlichen Prüfung: 18. Juli 2019

## Zusammenfassung

Verschränkung ist eines der grundlegendsten Phänomene der Quantentheorie und lässt sich nur unzureichend durch einen klassischen Formalismus beschreiben. Verschränkung spielt nicht nur möglicherweise eine Rolle in natürlichen Prozessen, sondern stellt insbesondere eine sehr mächtige Ressource für eine Vielzahl an Anwendungen bereit. So beruhen Quanteninformationsprotokolle wie Quantenzustandsteleportation, *superdense coding* sowie diverse Ansätze für *quantum computation* darauf. Dennoch ist es gerade in größeren Systemen, die aus mehr als zwei Teilchen bestehen, äußerst schwierig, Verschränkung nachzuweisen. In der hier vorliegenden Arbeit werden verschiedene Methoden vorgestellt, mit denen Vielteilchenverschränkung effizient detektiert werden kann. Diese Methoden werden experimentell untersucht und anschließend in Hinblick auf ihre Verwendbarkeit diskutiert.

Darüber hinaus werden Methoden für Quantenzustands- und Quantenprozessomographie eingeführt, mit deren Hilfe experimentelle Aufbauten genauestens untersucht werden können. Die dabei typischen Herausforderungen werden aufgezeigt, wobei Methoden präsentiert werden, die neue Einblicke in experimentell gewonnene Daten erlauben.

Abschließend wird das sogenannte Messproblem beleuchtet, das die unvermeidbare Störung eines Quantensystems durch einen Messprozess beschreibt. Es werden verschiedene Methoden für Quantenmessungen diskutiert, wobei schließlich die optimalen Messinstrumente hergeleitet werden. Mit diesen kann Information über einen Zustand gewonnen werden, während die dadurch verursachte Störung gleichzeitig minimiert wird. Diese optimalen Instrumente werden mit anderen Messschemata verglichen, die beispielsweise auf dem optimalen Quantenkloner beruhen.

## Abstract

Entanglement is one of the most fascinating features of quantum theory, as it fundamentally distinguishes quantum systems from classical systems. It not only might play a role in natural processes, but provides an especially powerful resource for a plethora of quantum information protocols, including quantum state teleportation, superdense coding as well as quantum computation schemes. However, for larger systems involving more than two parties, it is a challenging task to verify and characterize entanglement. In this thesis, efficient means for detecting genuine multipartite entanglement in multiqubit quantum systems will be derived, demonstrated, and discussed.

Beyond that, methods for quantum state and process tomography will be introduced, which allow to analyze and investigate prepared multipartite states or even the performance of whole setups. Some typical obstacles for reliable state tomography will be pointed out, together with new means to understand experimentally obtained data.

Finally, the measurement problem, which describes the unavoidable backaction of a measurement process to the quantum state under investigation, will be illustrated. Together with the discussion of different methods of quantum measurements, the optimal measurement instruments for binary qubit measurements will be derived. These optimal measurement instruments allow to perform a quantum measurement, providing a tunable amount of information about the quantum state, while avoiding any unnecessary disturbance to the state. This set of optimal instruments will be compared to other measurement schemes, such as measurements based on the optimal quantum cloning protocol.

# Contents

	Page
Zusammenfassung	v
Abstract	vi
Table of Contents	vii
List of Figures	x
List of Tables	xii
<b>1 Introduction</b>	<b>1</b>
<b>2 Fundamentals</b>	<b>5</b>
2.1 Introduction . . . . .	6
2.2 Theoretical Foundations . . . . .	6
2.2.1 Quantum States . . . . .	6
2.2.1.1 Qubit . . . . .	6
2.2.1.2 Density Matrix . . . . .	8
2.2.1.3 Multiqubit Systems . . . . .	10
2.2.1.4 Qutrits, Ququads and Qudits . . . . .	11
2.2.1.5 Unitary Transformation . . . . .	12
2.2.1.6 Partial Trace and Purification . . . . .	12
2.2.2 Entanglement . . . . .	13
2.2.2.1 k-separability and Entanglement Intactness . . . . .	14
2.2.2.2 k-producibility and Entanglement Depth . . . . .	14
2.2.2.3 Local Operations and Classical Communication . . . . .	15
2.2.2.4 Positive Partial Transpose . . . . .	17
2.2.2.5 Entanglement Witness . . . . .	17
2.2.2.6 Bell Inequality . . . . .	19
2.2.3 Distances and Measures . . . . .	20
2.2.3.1 Fidelity . . . . .	20
2.2.3.2 Bures Distance and Bures Angle . . . . .	22
2.2.3.3 Hilbert-Schmidt Norm . . . . .	22

2.2.3.4	Trace Distance . . . . .	23
2.2.4	Prominent Quantum States and Their Properties . . . . .	23
2.2.4.1	Maximally Mixed State . . . . .	24
2.2.4.2	Bell States . . . . .	24
2.2.4.3	Greenberger-Horne-Zeilinger States . . . . .	25
2.2.4.4	Graph and Cluster States . . . . .	25
2.2.4.5	Dicke and W States . . . . .	26
2.2.4.6	Qutrit and Qudit States . . . . .	27
2.2.5	Measurements and Quantum Instruments . . . . .	28
2.2.5.1	General Measurements . . . . .	28
2.2.5.2	Quantum Channels . . . . .	30
2.2.5.3	Quantum Instruments . . . . .	31
2.2.6	Statistical and Mathematical Concepts . . . . .	31
2.2.6.1	Distribution Functions . . . . .	31
2.2.6.2	Statistical Moments . . . . .	33
2.2.6.3	Hypothesis Testing and p-values . . . . .	34
2.2.6.4	Significance, Confidence, Power, and Errors . . . . .	36
2.2.6.5	Likelihood . . . . .	36
2.2.6.6	Maximum Likelihood . . . . .	37
2.2.6.7	Point Estimators . . . . .	38
2.2.6.8	Frequentistic and Bayesian inference . . . . .	39
2.2.6.9	Confidence Intervals . . . . .	39
2.2.6.10	Credibility Intervals . . . . .	40
2.3	Experimental Framework . . . . .	40
2.3.1	Linear Optics . . . . .	41
2.3.1.1	Waveplates . . . . .	41
2.3.1.2	Variable Phase Retardance . . . . .	43
2.3.1.3	Beam Splitters . . . . .	43
2.3.2	Non-linear Optics . . . . .	44
2.3.2.1	Spontaneous Parametric Down Conversion . . . . .	44
<b>3</b>	<b>Quantum State Tomography</b> . . . . .	<b>47</b>
3.1	Introduction . . . . .	48
3.2	Methods of Quantum State Tomography . . . . .	48
3.2.1	Formalism . . . . .	48
3.2.2	Overcomplete Pauli Tomography . . . . .	50
3.2.3	Informationally Complete Tomography . . . . .	51
3.2.4	Symmetric Informationally Complete Tomography . . . . .	52
3.2.5	Efficient Specialized Protocols . . . . .	52
3.2.5.1	Permutationally Invariant State Tomography . . . . .	53
3.2.5.2	Compressed Sensing . . . . .	53
3.2.5.3	Matrix Product State Tomography . . . . .	54
3.2.5.4	Adaptive State Tomography . . . . .	54

3.3	Enforcing Physicality . . . . .	55
3.3.1	Understanding Unphysical Results . . . . .	55
3.3.2	Constrained Optimization . . . . .	56
3.3.3	Bias in Various Experiments . . . . .	60
3.3.4	[P1] Systematic Errors in Current Qu. State Tomography Tools . . . . .	61
3.4	Multiqubit State Tomography with Finite Data . . . . .	71
3.4.1	Wigner Semicircle Distribution . . . . .	71
3.4.2	Catalan Numbers . . . . .	73
3.4.3	Illustration of Proof - Example of Fourth Moment . . . . .	75
3.4.4	Degenerate Eigenvalues . . . . .	79
3.4.5	Error Estimation . . . . .	79
3.4.6	[P2] Multiqubit State Tomography with Finite Data . . . . .	85
3.5	Conclusion . . . . .	102
<b>4</b>	<b>Characterizing Entanglement</b>	<b>103</b>
4.1	Introduction . . . . .	104
4.2	Efficient Multipartite Entanglement Detection . . . . .	105
4.2.1	Need for Efficient Methods . . . . .	105
4.2.2	Commuting and Anticommuting Operators . . . . .	106
4.2.3	Nonlinear Witness for Bell States . . . . .	108
4.2.4	Nonlinear Witness for Three-Qubit GHZ State . . . . .	110
4.2.5	[P3] Multipartite Entanglement Detection with Minimal Effort . . . . .	113
4.2.6	Application for Higher Dimensional Systems . . . . .	124
4.2.7	[P4] Quadratic Entanglement Criteria for Qutrits . . . . .	126
4.3	Entanglement and Random Measurements . . . . .	131
4.3.1	Distribution of Expectation Values . . . . .	132
4.3.1.1	Obtaining the Purity . . . . .	134
4.3.2	Certifying Genuine Multipartite Entanglement . . . . .	135
4.3.3	[P5] Multipartite Entanglement Analysis From Random Correlations . . . . .	140
4.4	Entanglement without Correlations . . . . .	148
4.4.1	Entanglement and Correlations . . . . .	148
4.4.2	Correlations and Predictability . . . . .	150
4.4.3	Postulates for Correlations . . . . .	151
4.4.4	Even and Odd Number of Qubits . . . . .	152
4.4.5	Obtaining a State without Correlations . . . . .	152
4.4.6	[P6] Gen. Multipart. Ent. without Multipartite Correlations . . . . .	154
4.4.7	[P7] Gen. $N$ -part. Ent. without $N$ -partite Correlation Functions . . . . .	172
4.5	Conclusion . . . . .	179
<b>5</b>	<b>Quantum Measurements</b>	<b>181</b>
5.1	Introduction . . . . .	182
5.2	Weak Measurements and Weak Values . . . . .	183
5.2.1	Weak Amplification . . . . .	186

5.2.2	Direct Tomography . . . . .	186
5.2.3	Weak Value Controversies . . . . .	186
5.2.4	Weak Value for Mixed Pre- and Postselection . . . . .	187
5.2.5	Universal Description of Local Interactions . . . . .	187
5.3	Quantum Instruments and Universal Cloning . . . . .	188
5.3.1	Uncertainty Principle and Robertson Inequality . . . . .	189
5.3.2	Heisenberg Microscope . . . . .	190
5.3.3	Measurement-Disturbance Tradeoff Relations . . . . .	191
5.3.3.1	Measures . . . . .	192
5.3.3.2	Universal Optimal Asymmetric Quantum Cloning . . . . .	192
5.3.3.3	Coherent Swap . . . . .	195
5.3.3.4	Optimal Instruments . . . . .	195
5.3.4	[P10] Procedures for the Optimal Meas.-Disturbance Tradeoff . . . . .	198
5.4	Conclusion . . . . .	209
<b>6</b>	<b>Conclusions</b>	<b>211</b>
	<b>List of Publications</b>	<b>215</b>
	<b>Bibliography</b>	<b>219</b>
	<b>Acknowledgments</b>	<b>245</b>
	<b>Index</b>	<b>245</b>

# List of Figures

2.1	Bloch Sphere. . . . .	7
2.2	Hierarchy of Separable States. . . . .	15
2.3	Entanglement Classes of GHZ and W States. . . . .	16
2.4	Entanglement Witness. . . . .	18
2.5	Relation of Bell Inequalities and Witnesses. . . . .	21
2.6	Graph States. . . . .	27
2.7	Distribution Functions. . . . .	32
2.8	Waveplate-Induced Transformations. . . . .	42
2.9	Non-Polarizing and Polarizing Beam Splitter. . . . .	44
2.10	Down Conversion with Type I Phase Matching. . . . .	45
2.11	Type II Down Conversion. . . . .	46
3.1	Implementation of Tomography Schemes. . . . .	51
3.2	Comparison of Tomography Methods. . . . .	58
3.3	Tomographically Accessible Quantum States. . . . .	59
3.4	Wigner Semicircle Distribution. . . . .	71
3.5	Slope at Eigenvalue Cutoff. . . . .	73
3.6	Non-Crossing and Crossing Partitions. . . . .	74
3.7	Non-Crossing and Crossing Partitions of Four Operators. . . . .	78
3.8	Distribution of Hilbert-Schmidt Norm. . . . .	83
4.1	Correlations of Bell State. . . . .	109
4.2	Correlations of 3-Qubit GHZ State. . . . .	112
4.3	Scheme for Random Measurements. . . . .	131
4.4	Probability Distributions for Randomly Measured Correlations. . . . .	134
4.5	Distribution of Biseparable 3-Qubit States. . . . .	137
4.6	Distribution of Biseparable 4-Qubit States. . . . .	138
4.7	Correlations and Predictability. . . . .	150
5.1	Standard Measurement. . . . .	183
5.2	Weak Measurement. . . . .	184
5.3	Concept of Weak Measurements. . . . .	185
5.4	Weak Value with Finite Coupling. . . . .	189

5.5	Universal Optimal Asymmetric Quantum Cloning. . . . .	193
5.6	Universal Quantum Cloning Circuit Model. . . . .	193
5.7	Coherent Swap Operation. . . . .	195
5.8	Measurement-Disturbance Relation with Optimal Instruments. . . . .	196

# List of Tables

3.1	Obtaining Unphysical Results. . . . .	55
3.2	Statistics of State Tomography. . . . .	72
4.1	Correlations of 3-Qubit GHZ State. . . . .	110
4.2	Correlations of $ W_3\rangle$ . . . . .	148
4.3	Correlations of $ \overline{W}_3\rangle$ . . . . .	149
4.4	Correlations of $\frac{1}{2} ( W_3\rangle\langle W_3  +  \overline{W}_3\rangle\langle \overline{W}_3 )$ . . . . .	149



# Chapter 1

## Introduction

In the late 19<sup>th</sup> century and beginning of the 20<sup>th</sup> century first experimental observations have been made which could not be explained using classical physics and thus required a more comprehensive theory. At that time, the radiation of a black body was extensively studied by several physicists, including Robert Kirchhoff, Josef Stefan, Ludwig Boltzmann, Wilhelm Wien, Lord Rayleigh, Max Planck and others. Although earlier measurements showed a good agreement between experiment and the contemporary theory, new results around the turn of the century indicated that those theories could have been only approximations [1]. In order to understand those deviations, Max Planck started to revise his theory, which was initially based solely on Maxwell's laws [2], by introducing some quantized "energy elements" with energy  $h\nu$  [3, 4], for which he introduced the new fundamental constant  $h$ . This finally led to a consistent description of black body radiation and to a solution of the "ultraviolet catastrophe" of the Rayleigh-Jeans law.

The idea of energy quanta was then adopted by Albert Einstein in 1905 which provided a novel perspective on the duality of particle and wave properties of light [5]. Subsequently, he also derived Planck's radiation law based on the approach of stimulated and spontaneous emission processes [6]. The quantization of angular momentum allowed Niels Bohr to derive the Balmer formula for the hydrogen atom, with it the Paschen and Lyman series, and to explain the stability of its energy levels [1, 7]. In the 1920s, the particle property of light quanta was shown anew by Arthur Compton in the scattering process of electrons and photons [8]. Shortly after, the wave property of matter had been hypothesized by Louis de Broglie [9] and could be experimentally confirmed in 1927 by Clinton Davisson and Lester Germer [10]. In 1925, Heisenberg derived a consistent description of quantum theory [11], which laid the foundation for the formulation of matrix mechanics [12, 13]. Inspired by de Broglie's theory, Schrödinger postulated in 1926 an equation of motion describing the evolution of matter waves, which is known nowadays as the *Schrödinger equation* [14]. Around the same time, Werner Heisenberg introduced his famous uncertainty relation, setting a lower bound to the simultaneous knowledge of two conjugate variables such as position and momentum [15].

Although this first period of quantum theory was strikingly fruitful, the implications as to the *meaning* of this new theory were not seriously challenged until the famous debate

between Niels Bohr and Albert Einstein, in which the role of randomness in quantum theory as well as the arbitrariness of choosing between either the particle and the wave properties were central themes [16]. In 1935, the discussion culminated in Albert Einstein, Boris Podolsky, and Nathan Rosen questioning the completeness of quantum theory [17]. While the discussion about hidden variables and completeness of quantum mechanics somewhat continued in the following years [18–20], it remained a mere theoretical debate. After almost 30 years, John Bell ingeniously found a means to probe for the validity of local hidden variable concepts [21]. He derived an inequality, which is based on only a few reasonable assumptions. It suffices that the underlying theory is local (effects propagate with at most luminal speed, allowing a space-like separation of measurements) as well as it includes the so-called elements of reality (measured outcomes are real properties of the system before the measurement). By experimentally violating this inequality, the invalidity of at least one of the assumptions can be shown. A few years after their introduction, first experimental demonstrations of a violation of Bell’s inequality have been given by Stuart Freedman and John Clauser [22] and subsequently by Clauser [23, 24], by Fry and Thompson [24, 25], and by several others, see [24] for a review. Another famous experiment was conducted by Alain Aspect together with Philippe Grangier and Gérard Roger [26]. However, those experimental tests could only provide first evidence against local hidden variables, as severe loopholes have not been closed. In 1998, the *locality loophole* has been closed [27] by sufficiently separating the two observers of the Bell test, while in 2001 the *detection loophole* could be closed for the first time [28]. It took almost another 15 years, until both of these loopholes have been closed at the same time [29–33], allowing finally to rule out local hidden variables without major loopholes.

The feature of *entanglement* (“Verschränkung”) was first introduced by Erwin Schrödinger. He begins one of his publications with the words [19]

When two systems, of which we know the states by their respective representatives, enter into temporary physical interaction due to known forces between them, and when after a time of mutual influence the systems separate again, then they can no longer be described in the same way as before [...]. I would not call that one but rather the characteristic trait of quantum mechanics, the one that enforces its entire departure from classical lines of thought. By the interaction the two representatives [...] have become entangled.

He continues by stating that

the best possible knowledge of a whole does not necessarily include the best possible knowledge of all its parts, even though they may be entirely separated and therefore virtually capable of being “best possibly known”, i.e., of possessing, each of them, a representative of its own.

Nowadays, entanglement is used as one of the key resources for a plethora of different quantum information protocols. As such, it powers versions of quantum key distribution schemes [34] for exchanging secure encryption keys and it enables superdense coding [35, 36]

---

allowing to transmit 2 bits by classically transferring only a single bit. Furthermore, Quantum state teleportation [37] using solely classical communication as well as entanglement swapping [38], which allows to entangle particles that never directly interacted, use entanglement as a key resource. Beyond that, quantum computation [39], for example relying on highly entangled quantum states [40], promises a great benefit for some computational problems which are hard to solve using classical algorithms [41]. This period of arising means to exploit and utilize quantum theory for quantum technology is sometimes referred to as the “second quantum revolution” in correspondence with the first one at the beginning of the 20<sup>th</sup> century, when the theory itself was developed.

Although entanglement is fairly well understood for bipartite systems, quantum systems involving more parties still raise several questions. In the present thesis, multipartite quantum systems are studied experimentally that contain up to six entangled qubits. Characterizing such a state and revealing its entanglements is a hard task, which requires specialized methods. If beyond that, one is interested in a complete description of the state, one needs to find the wavefunction of the system. In experimental situations, where noise cannot be avoided completely, or when subsystems of an entangled state are to be described, one usually refers to the density matrix formalism in order to fully describe the present state. Determining all free parameters of the density matrix is called *quantum state tomography*, which in general is an intricate process with several challenges, for example, due to the finite statistics of the measurement. To finally obtain a sufficiently accurate and precise estimate requires lots of sampling, which might require a long measurement time, if possible at all. This motivates the search for more efficient tools for tomography which are also sought to be less sensitive to statistical noise.

This work is structured as follows. In the following chapter 2, fundamental concepts such as quantum states, entanglement, measurements and the required experimental tools will be introduced and discussed. Those theoretical and experimental tools provide the basis for all subsequent chapters. In chapter 3, the challenge of quantum state and process tomography will be discussed. Along with publications that were published in the course of this thesis, the issues of finite statistics and systematic errors in state tomography will be discussed. In chapter 4, multipartite entanglement will be illuminated from various perspectives. First, different means how to reveal and certify genuine multipartite entanglement for qubit and qutrit systems are introduced. Then, the connection between entanglement and correlations will be challenged, using, on one hand, a scenario of measurements without local reference frames and, on the other hand, a quantum state without correlations. This so-called *no correlation* state is proven to be genuinely  $n$ -partite entangled without containing any  $n$ -partite correlations. Finally, in chapter 5, different types of quantum measurements are to be scrutinized. There, first, for pre- and postselected systems, the concept of *weak measurements* will be introduced, which can reduce the measurement induced backaction at the cost of information gain (per measured particle). After investigating fundamental issues in the context of *weak values*, the tradeoff relation between the information gained from a measurement and the thereby induced disturbance is studied. Using the approach of quantum instruments, a tight measurement disturbance tradeoff relation will be derived and experimentally confirmed. The optimal measurement

schemes will be derived, which outperform other reasonable schemes. Finally, in chapter 6, a brief summary of the present thesis will be given.

# Chapter 2

## Fundamentals

### Contents

---

<b>2.1</b>	<b>Introduction</b>	<b>6</b>
<b>2.2</b>	<b>Theoretical Foundations</b>	<b>6</b>
2.2.1	Quantum States	6
2.2.2	Entanglement	13
2.2.3	Distances and Measures	20
2.2.4	Prominent Quantum States and Their Properties	23
2.2.5	Measurements and Quantum Instruments	28
2.2.6	Statistical and Mathematical Concepts	31
<b>2.3</b>	<b>Experimental Framework</b>	<b>40</b>
2.3.1	Linear Optics	41
2.3.2	Non-linear Optics	44

---

## 2.1 Introduction

This chapter paves the way for the more specialized topics discussed in the subsequent parts of this thesis. Here, basic concepts or concepts of relevance for multiple parts of this thesis will be introduced, while topics with a more limited scope will be reviewed in the subsequent chapters.

First, a broad overview of fundamental theoretical concepts, which are the building blocks of quantum theory and thus also elementary ingredients of quantum state analysis, will be given. Afterwards, a discussion of the experimental tools which are needed for the state preparation and detection of photonic quantum states follows.

## 2.2 Theoretical Foundations

### 2.2.1 Quantum States

A quantum state is used to describe the current state of a particular quantum system. Contrary to classical physics, in quantum mechanics, a system can be in a superposition state of various basis states. In mathematical terms, for describing a pure quantum state, a state vector  $|\psi\rangle$  is sufficient. If, however, a statistical (classical) mixture of different pure quantum states is to be described, one has to resort to the *density matrix* or *density operator* formalism. In the following sections, first the concept of a single quantum bit will be introduced. Based on this, a generalization to statistical mixtures will be given. After extending the formalism to multiqubit systems, a short excursion to higher dimensional systems such as *qutrits* is given.

#### 2.2.1.1 Qubit

The term *quantum bit* or *qubit* [42] refers to a two level system with the ground state  $|0\rangle$  and the excited state  $|1\rangle$  being the eigenstates of, say, the Hamilton operator. While a classical system such as a classical bit can only occupy either the ground state or the excited state, a quantum system can also exist in a superposition of such [43].

In general, all (pure) two dimensional ( $d = 2$ ) states can be described by superpositions of  $|0\rangle$  and  $|1\rangle$ , with proper coefficients  $\alpha \in \mathbb{C}$ ,  $\beta \in \mathbb{C}$  using

$$|\psi\rangle = \alpha|0\rangle + \beta|1\rangle \in \mathcal{H}_2, \quad (2.1)$$

where  $\mathcal{H}_2$  denotes the two-dimensional Hilbert space. In order to obtain a normalized state, i.e.,  $\langle\psi|\psi\rangle = 1$ , the condition

$$|\alpha|^2 + |\beta|^2 = 1 \quad (2.2)$$

has to be fulfilled by the two complex amplitudes  $\alpha$  and  $\beta$ . A more elegant definition of a pure single qubit state incorporating this constraint is thus

$$|\psi\rangle = \cos\left(\frac{\theta}{2}\right)|0\rangle + e^{i\varphi}\sin\left(\frac{\theta}{2}\right)|1\rangle \quad (2.3)$$

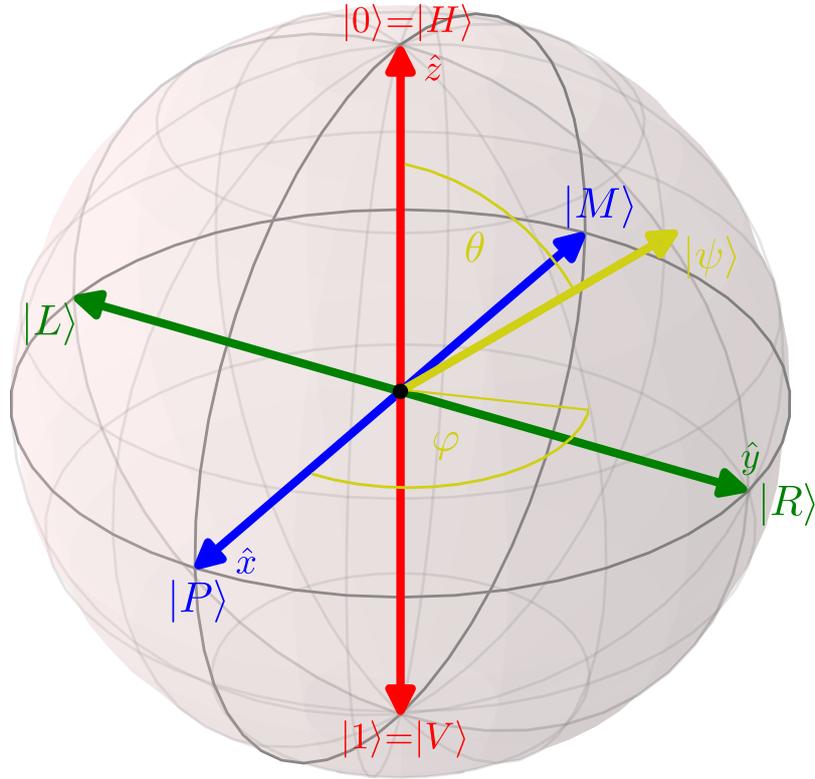


FIGURE 2.1: The Bloch sphere picture proves beneficial to represent the space of all states of a single qubit system. The ground state and the excited state are at the poles, while superpositions with different phases are located at the equator. Pure states are found on the surface, while mixed states lie inside of the Bloch sphere. A pure state  $|\psi\rangle$  (yellow vector) can be parametrized in terms of two angles,  $\theta$  and  $\varphi$ . Here, the states along the axes are identified with names used in the context of the polarization degree of freedom.

with the real parameters  $\theta \in [0, \pi]$  and  $\varphi \in [0, 2\pi]$ . Eq. (2.3) directly corresponds to a spherical representation of the qubit, the so-called *Bloch sphere* representation [44], see Fig. 2.1. Coherent superposition states as given in Eqs. (2.1) and (2.3) correspond to vectors lying at the sphere's surface. Using the parametrization of Eq. (2.3), the Cartesian coordinates corresponding to the respective states are given by

$$x = \sin \theta \cos \varphi, \quad (2.4a)$$

$$y = \sin \theta \sin \varphi, \quad (2.4b)$$

$$z = \cos \theta. \quad (2.4c)$$

States which fulfill the normalization condition of Eq. (2.2) are called *pure states*.

Qubits can be implemented using various systems [44]. Examples of realizations include spin-1/2 systems [45], where the spin directly corresponds to the eigenstates, atomic two

level systems [44], pairs of quantum dots [46], Rydberg atoms [47], Josephson junctions employing the electric charge as the degree of freedom [48], or superconducting loops using the magnetic flux [48]. In the context of photons, which are used throughout this thesis, one mainly resorts to its polarization degree of freedom, such that, e.g., horizontal polarization corresponds to the ground state and vertical polarization to the excited state in the two level picture [49, 50]. Alternatively, in interferometric setups, two possible paths and their superpositions are also used to encode a qubit. Even more options, also for higher dimensional quantum states, are possible when using the frequency or the mode structure of the light.

A description mathematically equivalent to the Bloch sphere is the Poincaré sphere [51], which is used in the specific context of polarization states of electromagnetic waves when considering the circularly polarized states as ground and excited state.

### 2.2.1.2 Density Matrix

While Eq. (2.3) is a useful parametrization of a single pure qubit state, it is not sufficient to describe general systems which can also be in an incoherent mixture of different pure states. There, the density matrix notation is introduced with the state of a system described by

$$\varrho = \sum_i p_i |\psi_i\rangle\langle\psi_i|, \quad (2.5)$$

where the states  $|\psi_i\rangle$  are incoherently added with mixing probabilities  $p_i$ ,  $\sum_i p_i = 1$ . Decomposing  $\varrho$  according to Eq. 2.5 into pure states  $|\psi_i\rangle$  is in general not unique. The eigendecomposition into  $\varrho = \sum_i \lambda_i |\psi'_i\rangle\langle\psi'_i|$  with orthogonal eigenstates  $|\psi'_i\rangle$  and the corresponding eigenvalues  $\lambda_i$  gives the probabilities of successful projection onto the respective eigenstates. For a pure state, Eq. (2.5) simplifies to  $\varrho = |\psi\rangle\langle\psi|$ .

The state of a single qubit can be written in the Pauli operator basis as

$$\varrho = \frac{1}{2} (\sigma_0 + T_x \sigma_x + T_y \sigma_y + T_z \sigma_z) \quad (2.6)$$

with the Bloch vector  $\vec{T} = (T_x, T_y, T_z)^T$ , the Pauli matrices  $\sigma_x$ ,  $\sigma_y$ , and  $\sigma_z$ , and the identity matrix  $\sigma_0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ . The Pauli matrices [43] are traceless Hermitian matrices, defined as

$$\sigma_x \equiv \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y \equiv \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_z \equiv \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \quad (2.7)$$

with the eigenstates

$$\sigma_x |+\rangle = + |+\rangle, \quad (2.8a)$$

$$\sigma_x |-\rangle = - |-\rangle, \quad (2.8b)$$

$$\sigma_y |\odot\rangle = + |\odot\rangle, \quad (2.8c)$$

$$\sigma_y |\oslash\rangle = - |\oslash\rangle, \quad (2.8d)$$

$$\sigma_z |0\rangle = + |0\rangle, \quad (2.8e)$$

$$\sigma_z |1\rangle = - |1\rangle, \quad (2.8f)$$

where  $|\pm\rangle \equiv \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$ ,  $|\circlearrowleft\rangle \equiv \frac{1}{\sqrt{2}}(|0\rangle + i|1\rangle)$ , and  $|\circlearrowright\rangle \equiv \frac{1}{\sqrt{2}}(|0\rangle - i|1\rangle)$  are superpositions of  $|0\rangle$  and  $|1\rangle$ . Please note that some references in the literature define  $\sigma_y$  as the conjugate of the above.

In the notation of polarization states, the correspondences  $|P\rangle \equiv |+\rangle$  (*plus* or *diagonally polarized*),  $|M\rangle \equiv |-\rangle$  (*minus* or *antidiagonally polarized*),  $|R\rangle \equiv |\circlearrowright\rangle$  (*right circularly polarized*),  $|L\rangle \equiv |\circlearrowleft\rangle$  (*left circularly polarized*),  $|H\rangle \equiv |0\rangle$  (*horizontally polarized*),  $|V\rangle \equiv |1\rangle$  (*vertically polarized*) will be used.

Eq. (2.6) can also be stated as

$$\varrho = \frac{1}{2} \sum_{\mu=0}^3 T_{\mu} \sigma_{\mu} \quad (2.9)$$

for  $\sigma_{\mu} \in \{\sigma_0, \sigma_x, \sigma_y, \sigma_z\}$  and with the parameter  $T_0 \equiv 1$ . Throughout this thesis, the Pauli matrices as well as Bloch vector elements may be labeled fully equivalently using alphanumeric indices  $\{0, x, y, z\}$  or using numeric indices only, for example  $\sigma_1 \equiv \sigma_x$ .

The diagonal entries of the density matrix in a given basis are called *populations*, as they describe the probability to find the system in this particular eigenstate. On the contrary, off-diagonal elements define the coherences between the different basis states and are therefore denoted as *coherences*.

The normalization of the state  $\varrho$ ,

$$\text{tr}(\varrho) = \text{tr}\left(\frac{1}{2} \sum_{\mu=0}^3 T_{\mu} \sigma_{\mu}\right) = \frac{1}{2} T_0 \text{tr}(\sigma_0) = T_0 \equiv 1, \quad (2.10)$$

is ensured by the constant parameter  $T_0 \equiv 1$ . For a given state  $\varrho$ , the Bloch vector components  $T_{\mu}$  can be obtained using  $T_{\mu} = \text{tr}(\varrho \sigma_{\mu})$ .

Recalling Eqs. (2.4), one realizes that the Cartesian coordinates directly correspond to the Bloch vector coordinates,

$$x = T_x = \text{tr}(\varrho \sigma_x) = \langle \sigma_x \rangle, \quad (2.11a)$$

$$y = T_y = \text{tr}(\varrho \sigma_y) = \langle \sigma_y \rangle, \quad (2.11b)$$

$$z = T_z = \text{tr}(\varrho \sigma_z) = \langle \sigma_z \rangle, \quad (2.11c)$$

where  $\langle o \rangle$  denotes the expectation value of the corresponding operator  $o$  with respect to the quantum state  $\varrho$ . Contrary to pure states, mixed states do not lie on the surface of the Bloch sphere, but are distinguished by a shorter Bloch vector length,  $|\vec{T}|_2 = \sqrt{x^2 + y^2 + z^2} < 1$ .

In the case of a single qubit, the Bloch vector length  $|\vec{T}|_2$  can easily be obtained by means of the purity  $\mathcal{P} \equiv \text{tr}(\varrho^2)$ ,

$$\begin{aligned} \mathcal{P} = \text{tr}(\varrho^2) &= \frac{1}{4} \sum_{\mu, \nu=0}^3 T_{\mu} T_{\nu} \text{tr}(\sigma_{\mu} \sigma_{\nu}) = \frac{1}{4} \sum_{\mu, \nu=0}^3 T_{\mu} T_{\nu} 2\delta_{\mu, \nu} \\ &= \frac{1}{2} \sum_{\mu=0}^3 T_{\mu}^2 = \frac{1}{2} (1 + T_x^2 + T_y^2 + T_z^2) = \frac{1}{2} \left(1 + |\vec{T}|_2^2\right), \end{aligned} \quad (2.12)$$

where  $\delta_{\alpha,\beta}$  denotes the Kronecker delta.

In summary, one-qubit states can thus be represented by points on the Bloch sphere if they are pure or somewhere inside the Bloch sphere if they are mixed states. The maximally mixed state

$$\varrho_{\text{mms}} \equiv \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad (2.13)$$

lies at the origin of the sphere, evidently with purity  $\mathcal{P} = 0$ .

### 2.2.1.3 Multiqubit Systems

The description of the single qubit quantum state, see Eq. (2.1), can be extended to multiqubit systems. Instead of a single qubit living in  $\mathcal{H}_2$ ,  $n$  qubits from  $\mathcal{H}_2$  each, can be combined, such that

$$|\psi\rangle = \sum_{i_1, \dots, i_n=0}^1 c_{i_1, \dots, i_n} |i_1\rangle \otimes \dots \otimes |i_n\rangle \in \mathcal{H}_2^{\otimes n}, \quad (2.14)$$

i.e., a general pure state can be written as the superposition of all combinations of tensor products of the single qubit basis states, where now the  $n$  qubits of the state can be distributed to up to  $n$  different parties or observers. Similarly to the case of a single qubit, the amplitudes have to fulfill a normalization constraint,  $\sum_{i_1, \dots, i_n=0}^1 |c_{i_1, \dots, i_n}|^2 = 1$ . It is important to note that the state in Eq. (2.14) can in general not be factorized and thus cannot be written as a product of individual qubit states,

$$|\psi_{\text{prod}}\rangle = \left( \sum_{i_1=0}^1 c_{i_1}^{(1)} |i_1\rangle \right) \otimes \dots \otimes \left( \sum_{i_n=0}^1 c_{i_n}^{(n)} |i_n\rangle \right), \quad (2.15)$$

where  $n$  normalization conditions  $|c_0^{(j)}|^2 + |c_1^{(j)}|^2 = 1$  are to be satisfied. A state of the form of Eq. 2.15 is called product state. For a more detailed analysis of different classes of states, see Sec. 2.2.2.

Since Eq. 2.14 is limited to pure states, a more general expression has to be found to imbed incoherent mixtures of pure states. In analogy to Eq. 2.9, a general mixed multiqubit state can be described using the tensor product of Pauli matrices, which form the operator basis. For  $n$  qubits, the summation thus contains  $4^n$  entries,  $\sigma_0$ ,  $\sigma_x$ ,  $\sigma_y$ , and  $\sigma_z$  for each qubit, i.e.

$$\varrho = \frac{1}{2^n} \sum_{\mu_1, \dots, \mu_n} T_{\mu_1, \dots, \mu_n} \sigma_{\mu_1} \otimes \sigma_{\mu_2} \otimes \dots \otimes \sigma_{\mu_n} \quad (2.16)$$

with  $\sigma_{\mu_j} \in \{\sigma_0, \sigma_x, \sigma_y, \sigma_z\}$ ,  $1 \leq j \leq n$ , and the correlation tensor  $T$ .  $T_{\mu_1, \dots, \mu_n}$  are called

correlations tensor elements or, for short, correlations. According to

$$\begin{aligned}
& \text{tr}(\sigma_{\nu_1} \otimes \sigma_{\nu_2} \otimes \cdots \otimes \sigma_{\nu_n} \varrho) = \\
& \frac{1}{2^n} \sum_{\mu_1, \dots, \mu_n} T_{\mu_1, \dots, \mu_n} \text{tr}(\sigma_{\nu_1} \otimes \sigma_{\nu_2} \otimes \cdots \otimes \sigma_{\nu_n} \cdot \sigma_{\mu_1} \otimes \sigma_{\mu_2} \otimes \cdots \otimes \sigma_{\mu_n}) = \\
& \frac{1}{2^n} \sum_{\mu_1, \dots, \mu_n} T_{\mu_1, \dots, \mu_n} \underbrace{\text{tr}(\sigma_{\nu_1} \sigma_{\mu_1})}_{2 \cdot \delta_{\nu_1, \mu_1}} \cdots \underbrace{\text{tr}(\sigma_{\nu_n} \sigma_{\mu_n})}_{2 \cdot \delta_{\nu_n, \mu_n}} = \\
& T_{\nu_1, \dots, \nu_n}, \tag{2.17}
\end{aligned}$$

the correlations can be obtained as expectation values of the respective Pauli matrices. Correlations with  $T_{\mu_1, \dots, \mu_n}$ , where  $\mu_j \neq 0$  for all  $j$ , i.e., all indices are  $x$ ,  $y$ , or  $z$ , are so-called *full correlations*, whereas those, which are expectation values of operators containing at least one  $\sigma_0$ , are named *non-full correlations*, or throughout this thesis more seldom *marginal correlations*. In order to experimentally obtain the value of a full correlation, all particles have to be measured in the respective basis, e.g., by projections onto the eigenstates of the basis. Hence, the outcomes of all  $n$  single qubit measurements are relevant. However, for example, the non-full correlation of  $\sigma_z \otimes \sigma_z \otimes \sigma_0$  is obtained by considering the outcomes of only the first two particles measured in the  $\sigma_z$  basis, whereas the third particles's result is not taken into account. *Local correlations* such as  $\sigma_z \otimes \sigma_0 \otimes \sigma_0$  are calculated based on a single qubit measurements and are sometimes also called *local Bloch vector*. Accordingly to the normalization in the single qubit case, Eq. (2.10), the correlation  $T_{0,0,\dots,0} \equiv T_{00\dots 0} \equiv 1$  is fixed due to the constraint  $\text{tr}(\varrho) = 1$ .

A density matrix has to obey the following requirements:

- the trace has to be normalized,  $\text{tr}(\varrho) = 1$ ,
- $\varrho$  has to be Hermitian,  $\varrho^\dagger = \varrho$ ,
- the eigenvalues have to be non-negative, i.e., the density matrix has to be positive semi-definite,  $\varrho \geq 0$ .

Therefore, the eigenvalues, which can be seen as the probabilities to obtain the corresponding eigenstate, have to fulfill  $\lambda_i \in [0, 1]$ . The construction according to Eq. (2.16) ensures the Hermiticity requirement, whereas the normalization is taken care of by  $T_{00\dots 0} \equiv 1$ . The last requirement, however, can cause some trouble, see Chapter 3.

#### 2.2.1.4 Qutrits, Ququads and Qudits

As introduced in Sec. 2.2.1.1, qubits correspond to states of two level systems. Yet, a generalization to higher dimensional systems is also possible, allowing the occupation of, for example, the eigenstates  $|0\rangle$ ,  $|1\rangle$ , and  $|2\rangle$  for a ternary quantum system or, for short, *qutrit*. Correspondingly, a *ququad* (sometimes *ququart*) is composed of the four eigenstates  $|0\rangle$ ,  $|1\rangle$ ,  $|2\rangle$ , and  $|3\rangle$ , whereas the general term *qudit* refers to a  $d$ -dimensional system based

upon the states  $\{|0\rangle, \dots, |d-1\rangle\}$ . Contrary to a qubit, which lives in a two-dimensional Hilbert space  $\mathcal{H}_2$ , a qudit relies on a  $d$ -dimensional Hilbert space  $\mathcal{H}_d$ . Although by far most parts of this thesis rely on qubit systems only, the extension to higher dimensional systems will be of relevance for some generalizations of proposed methods. As mentioned in Sec. 2.2.1.1, higher dimensional systems can be implemented using different frequencies or various optical modes of the light or for example using a time-bin encoding.

### 2.2.1.5 Unitary Transformation

Two  $n$ -qubit states  $\varrho_1 \in \mathcal{H}_2^{\otimes n}$  and  $\varrho_2 \in \mathcal{H}_2^{\otimes n}$  with the same eigenvalues can be transformed into each other using an unitary transformation  $U \in \mathbb{C}^{2^n \times 2^n}$  with  $\varrho_1 \rightarrow U\varrho_1U^\dagger = \varrho_2$ .  $U$  is unitary if and only if its conjugate transpose  $U^\dagger$  is also its inverse, i.e.,  $U^\dagger U = UU^\dagger = \mathbb{1}$ . Please note that both conditions are necessary, namely  $U^\dagger$  has to be simultaneously the left and right inverse [51]. Unitary transformations are isospectral, i.e., they do not change the spectrum of eigenvalues, as can be shown easily. If the Hamiltonian of a system is Hermitian, as for closed systems, the evolution of it can be described by a unitary matrix.

### 2.2.1.6 Partial Trace and Purification

A possible way to obtain statistically mixed states is by preparing a pure entangled state with larger number of particles. Subsequently, some of the particles can be traced out, i.e., only the remaining particles are considered to be a state with fewer particles. For an  $(n = n_A + n_B)$ -qubit state  $|\psi\rangle = |\psi_A\rangle \otimes |\psi_B\rangle$  with  $|\psi_A\rangle \in A \equiv \mathcal{H}_2^{\otimes n_A}$  and  $|\psi_B\rangle \in B \equiv \mathcal{H}_2^{\otimes n_B}$ , the partial trace over  $B$  results in

$$\varrho_A = \text{tr}_B (|\psi_A\rangle\langle\psi_A| \otimes |\psi_B\rangle\langle\psi_B|) = |\psi_A\rangle\langle\psi_A| \underbrace{\text{tr} (|\psi_B\rangle\langle\psi_B|)}_1 = |\psi_A\rangle\langle\psi_A|, \quad (2.18)$$

which is again a pure state, now with  $n_A$  particles. However, if the initial state is not a product state, the general parametrization from Eq. (2.16) may be used. For the partial trace of the last  $n_B$  qubits, one obtains

$$\begin{aligned} \varrho_A &= \text{tr}_B (\varrho) \\ &= \frac{1}{2^n} \sum_{\mu_1, \dots, \mu_n} T_{\mu_1, \dots, \mu_n} \text{tr}_B \left( \sigma_{\mu_1} \otimes \dots \otimes \sigma_{\mu_{n_A}} \otimes \sigma_{\mu_{n_A+1}} \otimes \dots \otimes \sigma_{\mu_n} \right) \\ &= \frac{1}{2^n} \sum_{\mu_1, \dots, \mu_n} T_{\mu_1, \dots, \mu_n} \underbrace{\sigma_{\mu_1} \otimes \dots \otimes \sigma_{\mu_{n_A}}}_{\text{subsystem A}} \underbrace{\text{tr} \left( \sigma_{\mu_{n_A+1}} \otimes \dots \otimes \sigma_{\mu_n} \right)}_{2^{n_B}} \\ &= \frac{1}{2^{n_A}} \sum_{\mu_1, \dots, \mu_{n_A}} T_{\mu_1, \dots, \mu_{n_A}, 0, \dots, 0} \sigma_{\mu_1} \otimes \dots \otimes \sigma_{\mu_{n_A}}. \end{aligned} \quad (2.19)$$

Consider the maximally entangled Bell state  $|\psi^-\rangle = \frac{1}{\sqrt{2}}(|01\rangle - |10\rangle)$  (see Sec. 2.2.4.2), which is a pure two-qubit state with correlations  $T_{00} = 1$  and  $T_{11} = T_{22} = T_{33} = -1$ . After tracing out the, say, second particle, one obtains

$$\varrho_A = \text{tr}_B (|\psi^-\rangle\langle\psi^-|) = \frac{1}{2} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad (2.20)$$

which is the maximally mixed state. Thus, tracing out one particle of this pure two-qubit state results in the maximally mixed one-qubit state.

The inverse process is called *purification*. A canonical way how to construct a pure state  $|\Psi\rangle$  with  $2n$  qubits from the mixed state  $\varrho_A$  with  $n$  qubits relies on the eigendecomposition, i.e.,  $\varrho_A = \sum_i \lambda_i |\psi_i\rangle\langle\psi_i|$ . The pure state is then given by  $|\Psi\rangle = \sum_i \sqrt{\lambda_i} |\psi_i\rangle_A \otimes |\psi_i\rangle_B$ . However,  $|\Psi\rangle$  is not unique, as different pure states can result in the same mixed state by tracing out some particles [51]. Furthermore, the eigendecomposition is not unique if some eigenvalues are degenerate.

Please note that the usual density matrix formalism cannot discriminate between *proper* and *improper* mixtures [52], i.e., one cannot distinguish a proper mixture where the mixedness is due to some inherently statistical uncertainty (if existing at all) from an improper one which is only mixed due to entanglement to external systems which are out of reach. Taking the external system of an improper mixture into account, the joint system of the measured and the external one can be described by a pure state, i.e., some information about the measured system might be “somewhere else”.

### 2.2.2 Entanglement

According to Eq. (2.14), any pure two-qubit state can be written as

$$|\psi\rangle = \sum_{i_1, i_2=0}^1 c_{i_1, i_2} |i_1\rangle \otimes |i_2\rangle = c_{0,0}|00\rangle + c_{0,1}|01\rangle + c_{1,0}|10\rangle + c_{1,1}|11\rangle \quad (2.21)$$

with properly normalized amplitudes  $c_{i_1, i_2}$ . The abbreviation  $|00\rangle = |0\rangle \otimes |0\rangle$  will be used throughout this thesis. If it is not possible to find a decomposition into the product form

$$|\psi_{\text{prod}}\rangle = \left( \sum_{i_1=0}^1 c_{i_1} |i_1\rangle \right) \otimes \left( \sum_{i_2=0}^1 c_{i_2} |i_2\rangle \right), \quad (2.22)$$

the state  $|\psi\rangle$  is not a product state, but *entangled*. While this formulation is rather trivial for the case of a pure two-qubit state, the description and the detection of entanglement turn out to be more complicated for mixed  $n$ -qubit states. In this section, the terminology of separability and producibility are introduced. Furthermore, some fundamental entanglement detection schemes will be explained.

### 2.2.2.1 $k$ -separability and Entanglement Intactness

Contrary to a pure two-qubit state, which can either be entangled or a product state, for mixed states one refers to the concept of separability. A mixed two-qubit state is called separable, if it can be written in the form

$$\varrho_{\text{sep}} = \sum_i p_i (\varrho_A)_i \otimes (\varrho_B)_i, \quad (2.23)$$

i.e., the convex combination of states  $\varrho_A$  from subsystem  $A$  and  $\varrho_B$  from subsystem  $B$ , respectively [53]. In turn, a mixed two-qubit state, which is not separable, is entangled.

A state of the form of Eq. (2.23), which is separable into the states of two subsystems, is called biseparable. In general,  $\varrho$  is said to be  $k$ -separable if one can find a convex decomposition with

$$\varrho_{k\text{-sep}} = \sum_i p_i \bigotimes_{j=1}^k \varrho_{G_j^{(i)}} \quad (2.24)$$

where  $G_j^{(i)}$  denotes a set of indices for a corresponding  $k$ -fold partition with  $\mathcal{G} = \bigcup_{j=1}^k G_j^{(i)}$  (with mutually disjoint  $G_j^{(i)}$  and  $G_{j'}^{(i)}$ ) denoting the set of all indices. Please note that for the different summation terms labeled by  $i$ , different partitions can be chosen. For example, the state

$$\varrho_{\text{bi-sep}} = \frac{1}{2} (\varrho_{AB} \otimes \varrho_C + \varrho_A \otimes \varrho_{BC}) \quad (2.25)$$

with non-separable  $\varrho_{AB}$  and  $\varrho_{BC}$  is bi-separable because it is the convex combination of bi-separable states [54], see Fig. 2.2.

A  $k$ -separable state is also  $k'$ -separable with  $k' < k$ , for example, a tri-separable ( $k = 3$ ) state is also bi-separable ( $k = 2$ ). A state which is  $k$ -separable, but not  $(k + 1)$ -separable, is said to have an *entanglement intactness* of  $k$  [55]. The mixture of states with  $k_1$  and  $k_2$  entanglement intactness, respectively, has at least  $\min\{k_1, k_2\}$  entanglement intactness, i.e., the entanglement intactness of a mixture cannot be smaller than the intactness of the respective constituent with the smaller intactness.

A state with  $n$  qubits which is not bi-separable (and thus not  $k$ -separable for any  $k \geq 2$ ) is called *genuinely  $n$ -partite entangled* [56]. Please note that the mixture of two states which are bi-separable with respect to different bipartitions can never contain genuine  $n$ -partite entanglement, see Fig. 2.2. At the other end of the Hierarchy, a state with  $n$  qubits which is  $n$ -separable, is called fully separable.

### 2.2.2.2 $k$ -producibility and Entanglement Depth

In line with the concept of separability, the notion of producibility may be introduced. If a state  $\varrho$  can be formed using an entangled state with entanglement between  $m$  parties,  $\varrho$  is said to be  $m$ -producible. Such an  $m$ -producible state can of course also be formed by means of states with entanglement between  $m'$  (with  $m' > m$ ) parties, thus, it is also  $m + 1$ -producible. For example, the state  $\varrho_{\text{bi-sep}}$  from Eq. (2.25) is 2-producible because

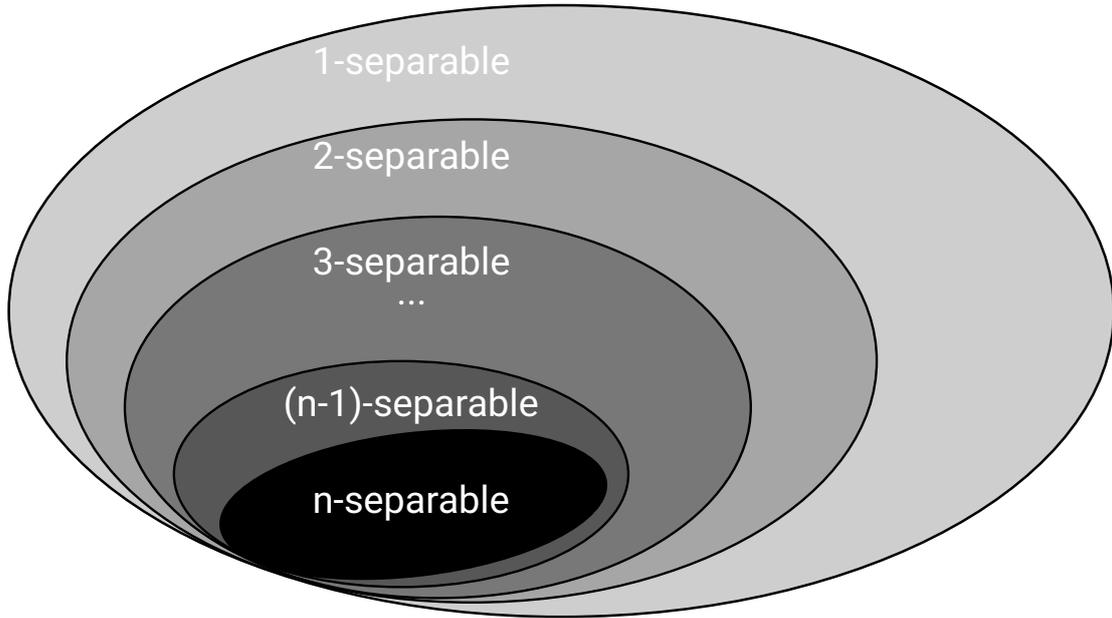


FIGURE 2.2: Hierarchy of separable states. The  $n$ -separable states are a convex subset of the convex set of the  $(n - 1)$ -separable states. Thus, by incoherently mixing two  $k$ -separable states, the resulting state is at least  $k$ -separable. The set of 1-separable states which are not 2-separable (or bi-separable) is the set of *non-separable states* or *genuinely  $n$ -partite entangled states*.

it can be produced using states with entanglement between 2 parties. On the other hand, with the same argument it is also 3-producible.

In order to avoid this ambiguity, the concept of *entanglement depth* is used [57]. A state with entanglement depth  $m$  is  $m$ -producible, but not  $(m - 1)$ -producible. The mixture of two states with entanglement depths  $m_1$  and  $m_2$ , respectively, can have at most entanglement depth of  $\max\{m_1, m_2\}$ .

To produce a genuinely  $n$ -partite entangled state, one needs a resource with at least  $n$ -partite entanglement. Thus, this state is  $n$ -producible and has an entanglement depth of  $n$ .

### 2.2.2.3 Local Operations and Classical Communication

In order to introduce some classification in multipartite entanglement the generic picture of  $n$  particles distributed to  $n$  parties has been used. There, only local measurements of the parties together with classical communication between them are allowed. All transformations of a quantum state which can be described by the set of local operations on any of the involved subsystems including local measurements together with classical information are encapsulated by the term of local operations and classical communication (LOCC) [43]. A channel with solely local operations and classical communication is thus a channel which can be implemented by  $n$  separate parties performing individual actions

on their respective subsystem. All communication between the parties is restricted to be classical such that no quantum information may be exchanged [58]. If two systems cannot be transformed into each other using local operations and classical communication, those systems are said to be *incomparable* and contain a different type of entanglement [59]. In contrast, if they are LOCC equivalent, they can be transformed into each other using local unitary transformations [60]. If only  $|\psi\rangle$  can be transformed into another state  $|\phi\rangle$ , but not vice versa,  $|\psi\rangle$  results in a larger value for some entanglement monotones [61].

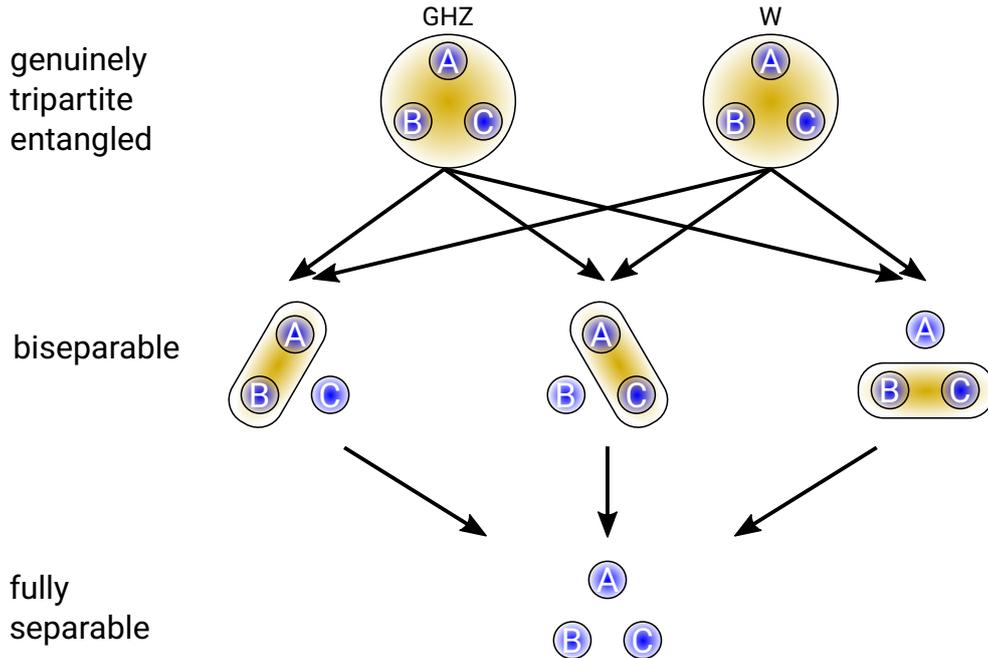


FIGURE 2.3: Entanglement classes of three-qubit states. Two classes of genuinely tripartite entangled states exist, which cannot be transformed into each other using the most general type of local operations and classical communications, which allow a finite probability of success [62]. However, bi-separable and tri-separable states can be generated from those, as indicated by arrows. Hence, all states can access the class of tri-separable states. Adapted from [62].

If, however, one accepts that a transformation may only succeed with a finite probability as, e.g., when using local filtering [63, 64], the class of transformations is denoted as *stochastic LOCC* (SLOCC). An overview for the convertibility of three-qubit states under SLOCC is shown in Fig. 2.3. The  $|\text{GHZ}\rangle$  state and the  $|\text{W}\rangle$  state, see Sec. 2.2.4, respectively, constitute two mutually distinct classes of SLOCC equivalent states, which cannot be transformed into each other using SLOCC [62]. However, from both types of states, all possible bi- and tri-separable states can be obtained by means of SLOCC. The inverse operation is not possible, i.e., a state with entanglement between all three parties cannot be formed from a bi-separable state using local operations and classical communication only.

### 2.2.2.4 Positive Partial Transpose

In general it can be a demanding task to infer whether a given state is entangled. If the density matrix  $\varrho$  is known and the system is composed of 2 qubits (or a qubit and a qutrit), the criterion of partial transposition can be used [65, 66]. For a single qubit state  $\varrho_A \in \mathcal{H}_2$ , its transpose

$$\varrho_A^T = \varrho = \frac{1}{2} \sum_{\mu=0}^3 T_\mu \sigma_\mu^T \quad (2.26)$$

is still a valid state with  $\varrho_A^T \geq 0$ . Because of  $\sigma_x = \sigma_x^T$ ,  $\sigma_y = -\sigma_y^T$ , and  $\sigma_z = \sigma_z^T$ , only the chirality of the state will be changed, i.e., the right-circular component becomes the left-circular one and vice-versa. If now the partial transpose, i.e., the transpose to a subsystem, is applied, a product state  $\varrho = \varrho_A \otimes \varrho_B$  always remains physical. Let  $\varrho_A$  and  $\varrho_B$  have the eigenvalues  $\lambda_{1,2}^{(A)}$  and  $\lambda_{1,2}^{(B)}$ , respectively, which all are non-negative. The eigenvalues of  $\varrho_A^{TA}$  remain  $\lambda_{1,2}^{(A)}$ . Then,

$$\varrho \rightarrow \varrho^{TA} = \varrho_A^{TA} \otimes \varrho_B \quad (2.27)$$

has the four non-negative eigenvalues  $\lambda_1^{(A)}\lambda_1^{(B)}$ ,  $\lambda_1^{(A)}\lambda_2^{(B)}$ ,  $\lambda_2^{(A)}\lambda_1^{(B)}$ , and  $\lambda_2^{(A)}\lambda_2^{(B)}$ . Thus,  $\varrho^{TA} \geq 0$ . Therefore, separable states always have a positive partial transpose (PPT). If the partial transpose of this 2-qubit state is negative, the state cannot be separable and is thus entangled. Please note that states for more than  $d = 6$  dimensions, i.e., systems with more dimensions than two qubits or a qubit-qutrit, can have a positive partial transpose while being entangled. Therefore, the PPT criterion is only sufficient and necessary for  $d \leq 6$  dimensional systems [65, 66], while it is a sufficient criterion for systems with  $d > 6$ .

### 2.2.2.5 Entanglement Witness

While the PPT criterion is a powerful tool applicable to any state, it is limited so small systems and requires full knowledge of the density matrix. For many - including experimental - purposes, one is interested in methods for entanglement verification on multiqubit (or in general multiqutrit) systems, ideally without the need of full quantum state tomography in the first place. *Entanglement witnesses* provide such means, i.e., they allow to verify entanglement with limited measurement effort [65, 67–69]. While there are different types of entanglement witnesses including non-linear witnesses, for this introduction we focus on the elementary case of projection-based linear entanglement witnesses.

An operator  $\hat{\mathcal{W}} \in \mathcal{H}_A \otimes \mathcal{H}_B$  is called entanglement witness if there exists an entangled state  $\varrho_{\text{entangled}} \in \mathcal{H}_A \otimes \mathcal{H}_B$  such that

$$\text{tr} \left( \varrho_{\text{entangled}} \hat{\mathcal{W}} \right) < 0, \quad (2.28)$$

while for all separable states  $\varrho_{\text{separable}} = \sum_i \varrho_A^{(i)} \otimes \varrho_B^{(i)}$  (with  $\varrho_A \in \mathcal{H}_A$ ,  $\varrho_B \in \mathcal{H}_B$ )

$$\text{tr} \left( \varrho_{\text{separable}} \hat{\mathcal{W}} \right) \geq 0 \quad (2.29)$$

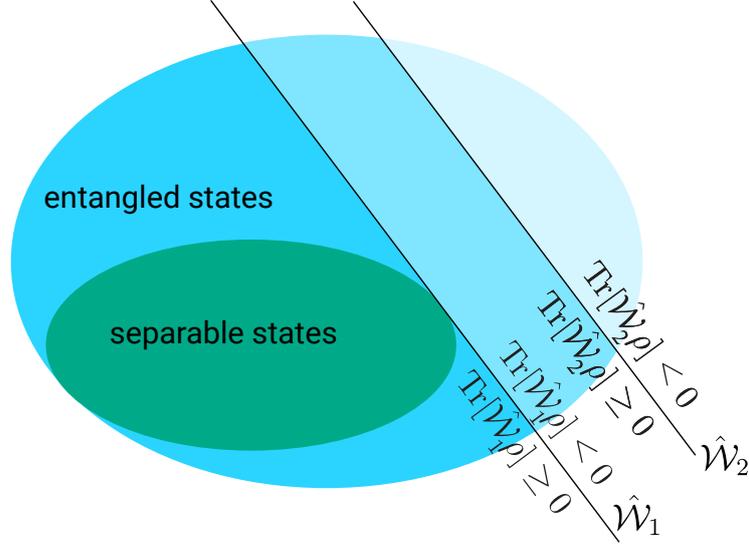


FIGURE 2.4: Entanglement witness. The witness  $\hat{\mathcal{W}}_1$  is optimal (with respect to some reference state), while  $\hat{\mathcal{W}}_2$  certifies entanglement only for a subset of the set of detected states of  $\hat{\mathcal{W}}_1$ . The separable states (green) are a convex subset of all states with  $\varrho \geq 0$ . A witness divides the set of all states into two subsets - a subset containing only entangled states (lighter blue areas) and a subset containing all separable states and some entangled states.

holds.

A standard ansatz for a (linear, fidelity based) entanglement witnesses is given by

$$\hat{\mathcal{W}} = \alpha \mathbb{1} - |\chi\rangle\langle\chi|. \quad (2.30)$$

For  $\alpha$  being the largest (squared) overlap of the reference state  $|\chi\rangle$  with the set of separable states, i.e.,  $\alpha = \max_{\varrho_{\text{separable}}} \langle\chi|\varrho_{\text{separable}}|\chi\rangle$ ,  $\hat{\mathcal{W}}$  certifies entanglement for  $\langle\hat{\mathcal{W}}\rangle < 0$ .

Please note that all separable states fulfill  $\text{tr}(\varrho_{\text{separable}}\hat{\mathcal{W}}) \geq 0$ , while only some entangled states result in  $\text{tr}(\varrho_{\text{entangled}}\hat{\mathcal{W}}) < 0$ . Thus, for a state  $\varrho$ , one can deduce

$$\text{tr}(\varrho\hat{\mathcal{W}}) < 0 \quad \Rightarrow \quad \varrho \text{ is entangled,} \quad (2.31)$$

$$\text{tr}(\varrho\hat{\mathcal{W}}) \geq 0 \quad \Rightarrow \quad \varrho \text{ could be entangled or separable.} \quad (2.32)$$

If a witness  $\hat{\mathcal{W}}_1$  certifies entanglement for the set of states  $\mathcal{S}_1$  and  $\hat{\mathcal{W}}_2$  for the set  $\mathcal{S}_2$ ,  $\hat{\mathcal{W}}_1$  is said to be finer than  $\hat{\mathcal{W}}_2$  if  $\mathcal{S}_2 \subseteq \mathcal{S}_1$ , i.e., the set of detected states by  $\hat{\mathcal{W}}_2$  is a subset of those detected by  $\hat{\mathcal{W}}_1$ . If there is no witness finer than  $\hat{\mathcal{W}}_1$ ,  $\hat{\mathcal{W}}_1$  is an optimal witness [68], see Fig. 2.4. Because the set of entangled states (all states excluding separable ones) is not convex, different optimal entanglement witnesses exist.

### 2.2.2.6 Bell Inequality

In 1935, Einstein, Podolsky, and Rosen questioned the completeness of quantum mechanics [17]. If two observers,  $A$  and  $B$ , are given the two particles of a position-momentum entangled bipartite quantum system, both can separately decide to measure position ( $x$ ) or momentum ( $p$ ) of their particles. Due to the entanglement, the position measurement of  $A$  would allow to deduce the position of  $B$ 's particle, as well as  $B$ 's momentum measurement would allow to deduce the corresponding momentum of  $A$ 's particle. As those measurements would allow to obtain  $x_A$  and  $p_A$  with certainty, those results are considered to be some elements of reality. If instead of position and momentum, the frequently considered example of a bipartite state with vanishing total spin,  $|\psi^-\rangle = \frac{1}{\sqrt{2}}(|\downarrow\uparrow\rangle - |\uparrow\downarrow\rangle)$ , is used, a measurement of the  $x$  component of the first particle could allow to deduce the  $x$  component of the second particle with certainty. If now the second particle's spin is not measured along  $x$  direction, but along  $y$  (or  $z$ ) direction, according to EPR's reasoning two non-commensurable observables have been measured, which is not allowed by quantum theory. Einstein, Podolsky and Rosen stated that [17]

if, without in any way disturbing a system, we can predict with certainty (i.e., with probability equal to unity) the value of a physical quantity, there exists an element of physical reality corresponding to this physical quantity.

Furthermore, they required that [17]

every element of the physical reality must have a counterpart in the physical theory.

Combining these assumptions, they concluded that [17]

it is possible to assign two different wave functions [...] to the same reality,

which either contradicts their reality statement or quantum mechanics cannot be complete.

An allegedly potential solution to preserve the reality assumption is to enrich quantum theory by additional parameters, which may be inaccessible (hidden) to the experimenter, but could allow to predict the measurement outcomes with certainty. In his seminal paper [21], Bell derived an inequality allowing to experimentally test for such local hidden variables. Under the assumption that space-like separated measurements are independent of each other (*locality*) and that a measurement reads system properties which are independent of the performed measurement [70] and definite (*realism*)<sup>1</sup>, Bell could derive an inequality that constrains possible outcomes of measurements on any system. Thus, any system that can be described by local hidden variables, i.e., some not necessarily empirically accessible, but local parameters containing all information for determining the measurement outcomes, has to produce outcomes which obey Bell's inequality. The observation of a violation in a physical system consequently leads to the conclusion that local

<sup>1</sup>Please note that different meanings for the requirement of realism exist. For a discussion of the conditions of Bell's theorem, see, e.g., [70, 71].

hidden variable models are not sufficient for describing the system and thus (at least) one of the assumptions is wrong. Hence, one has to either abandon the locality or the reality assumption - or both. A few years later, Clauser, Horne, Shimony, and Holt (CHSH) derived another Bell-type inequality [72] which is more experimentally friendly than Bell's original inequality. For two observers  $A$  and  $B$  with measurement settings  $\{\alpha_1, \alpha_2\}$  and  $\{\beta_1, \beta_2\}$ , the CHSH inequality reads

$$\mathcal{S} = |E(\alpha_1, \beta_1) - E(\alpha_1, \beta_2)| + |E(\alpha_2, \beta_1) + E(\alpha_2, \beta_2)| \leq 2, \quad (2.33)$$

where  $E(\alpha, \beta) = (c_{\uparrow\uparrow} - c_{\uparrow\downarrow} - c_{\downarrow\uparrow} + c_{\downarrow\downarrow}) / (c_{\uparrow\uparrow} + c_{\uparrow\downarrow} + c_{\downarrow\uparrow} + c_{\downarrow\downarrow})$  is the correlation of the outcomes  $\{c_{\uparrow\uparrow}, c_{\uparrow\downarrow}, c_{\downarrow\uparrow}, c_{\downarrow\downarrow}\}$  when  $A$  ( $B$ ) measures in setting given by angle  $\alpha$  ( $\beta$ ). Optimizing over the respective angles and input states, local hidden variable models are constrained by the value 2, while predictions with quantum theory allow values for the  $\mathcal{S}$  parameter up to the Tsirelson<sup>2</sup> bound [73] of  $2\sqrt{2}$ .

First experimental tests appeared already soon after the derivation of the inequality [22, 26]. Yet, those early experiments did not allow to close the so-called *locality loophole* or the *detection loophole* [74]. Only recently, first experiments managed to space-like separate both observers and, at the same time, measure with sufficient detection efficiency, simultaneously excluding both loopholes [29–33]. As for any separable state one can easily find a corresponding local hidden variable model, a violation of a Bell-type inequality also detects entanglement. However, compared to entanglement witnesses, Bell inequalities are less fine and can thus be considered to also be non-optimal witnesses [75]. For illustration, see Fig. 2.5.

### 2.2.3 Distances and Measures

Both for theoretical work on quantum information and for experimental verification of results, it is necessary to assess the quality of experimental results and hence to quantify the mutual distance between different states, distributions or channels, respectively. It is thus expedient to introduce the most basic measures used throughout this thesis.

#### 2.2.3.1 Fidelity

The *fidelity* was the first measure used in quantum information and describes and quantifies the overlap of two states. It can be thought of as the probability of success to find the system relative to a specific target or reference state [51]. For a pure reference state, a definition of the fidelity between the state to be tested  $\varrho$  and the reference state  $|\psi\rangle$  reads

$$F(\varrho, |\psi\rangle) \equiv \langle \psi | \varrho | \psi \rangle. \quad (2.34)$$

As such, it can be also used, e.g., to identify the state preparation quality.

---

<sup>2</sup>Due to transcription, it is sometimes referred to the ‘‘Cirelson bound’’ or ‘‘Cirel’son bound’’.

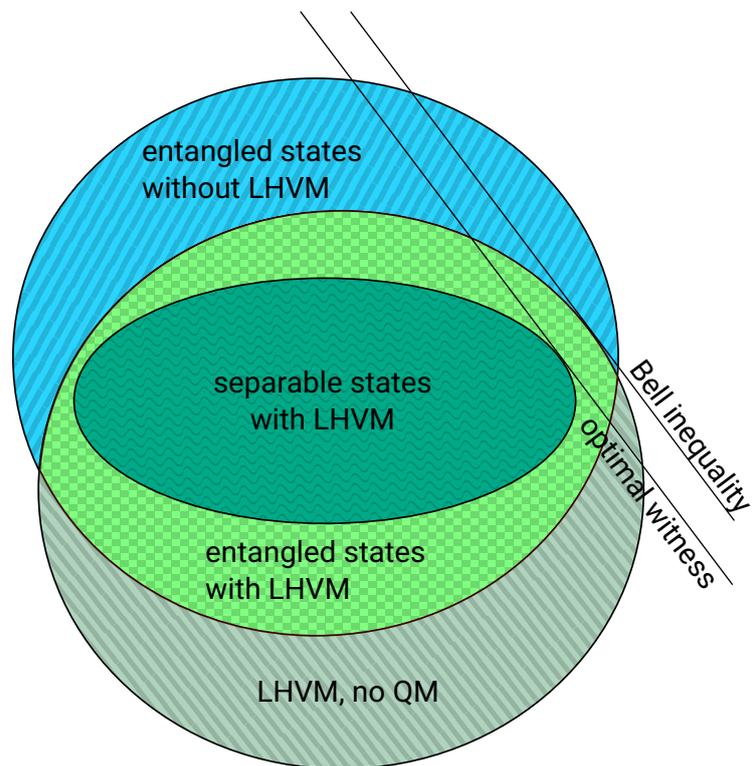


FIGURE 2.5: Models for possible measurement outcomes and the relation of Bell inequalities and witnesses. A Bell inequality is a tight bound to the set of measurement outcomes which can be explained using local hidden variable models (LHVM). However, there exist entangled quantum states, for which LHVM can be found [53]. Thus, in general an optimal witness is tighter (finer) to the set of separable states and thus detects more entangled states compared to a Bell inequality as a witness can be a tight bound to the measurement outcomes of separable states. Additionally, local hidden variable models can give probability distributions in disagreement with quantum theory. Figure adapted from [75].

If, however, the reference state is not pure, the definition of Eq. (2.34) cannot be used. Instead, the fidelity of the two (generally) mixed states  $\varrho$  and  $\sigma$  is then defined as

$$F(\varrho, \sigma) \equiv \left[ \text{tr} \sqrt{\sqrt{\varrho} \sigma \sqrt{\varrho}} \right]^2, \quad (2.35)$$

where  $\sqrt{\varrho}$  indicates the matrix root, i.e., the matrix  $\sqrt{\varrho}$  which solves  $\sqrt{\varrho} \sqrt{\varrho} = \varrho$ . This expression is also referred to as Uhlmann's transition probability.

The fidelity complies with the following requirements [51]

- $0 \leq F \leq 1$  with  $F(\varrho, \sigma) = 1 \Leftrightarrow \varrho = \sigma$ ,
- $F$  is symmetric in its arguments,
- the pure state definition of Eq. (2.34) is a special case of the general definition of Eq. (2.35), and
- $F$  is invariant under unitary transformations of both states,  $F(U\varrho U^\dagger, U\sigma U^\dagger) = F(\varrho, \sigma)$ .

As  $F(\varrho, \varrho) = 1$ , the fidelity itself is not a metric, since for a metric  $d(x, x) = 0$  has to hold. However, the fidelity can be used to induce a metric as will be shown in the next section.

Please note that sometimes in literature another definition of the fidelity is used, which is equivalent to the root of the quantity defined here.

### 2.2.3.2 Bures Distance and Bures Angle

The squared Bures distance between two quantum states  $\varrho$  and  $\sigma$  is defined as

$$D_B^2(\varrho, \sigma) \equiv 2 \left( 1 - \sqrt{F(\varrho, \sigma)} \right). \quad (2.36)$$

Furthermore, the so-called Bures angle can be introduced which is defined as

$$D_A(\varrho, \sigma) \equiv \arccos \sqrt{F(\varrho, \sigma)}. \quad (2.37)$$

Contrary to the fidelity, Bures distance and Bures angle comply with the requirements of a metric.

### 2.2.3.3 Hilbert-Schmidt Norm

The (squared) Hilbert-Schmidt norm of an operator  $A$  is generally defined as

$$|A|_{\text{HS}}^2 \equiv \text{tr} (A^\dagger A) = \sum_{i,j} |A_{i,j}|^2, \quad (2.38)$$

where  $\{A_{i,j}\}$  are the matrix elements in an arbitrary basis. For a density operator  $\varrho$ , this expression can be further written as

$$\begin{aligned} |\varrho|_{\text{HS}}^2 &= \text{tr}(\varrho^\dagger \varrho) \\ &= \frac{1}{4^n} \sum_{\mu_1, \dots, \mu_n} \sum_{\nu_1, \dots, \nu_n} T_{\mu_1, \dots, \mu_n} T_{\nu_1, \dots, \nu_n} \text{tr}(\sigma_{\mu_1} \otimes \dots \otimes \sigma_{\mu_n} \sigma_{\nu_1} \otimes \dots \otimes \sigma_{\nu_n}) \\ &= \frac{1}{2^n} \sum_{\mu_1, \dots, \mu_n} T_{\mu_1, \dots, \mu_n}^2, \end{aligned} \quad (2.39)$$

i.e., is given by the sum of the squared correlation tensor elements. Accordingly, the squared Hilbert-Schmidt distance of the difference of two density matrices reads

$$|\varrho - \sigma|_{\text{HS}}^2 = \frac{1}{2^n} \sum_{\mu_1, \dots, \mu_n} \Delta T_{\mu_1, \dots, \mu_n}^2 \quad (2.40)$$

with  $\Delta T_{\mu_1, \dots, \mu_n} \equiv T_{\mu_1, \dots, \mu_n}^\varrho - T_{\mu_1, \dots, \mu_n}^\sigma$ . *Frobenius norm*, *Schur norm*, and *Schatten-2 norm* are other names of the Hilbert-Schmidt norm.

### 2.2.3.4 Trace Distance

Of physical interest is also the *trace distance* as it allows to quantify the highest probability that two states can be distinguished using an optimal measurement. Another name for the trace distance norm is *Schatten-1 norm*. For two (generally) mixed states  $\varrho$  and  $\sigma$ , one denotes

$$D_T(\varrho, \sigma) = \frac{1}{2} |\varrho - \sigma|_1 = \frac{1}{2} \text{tr} \left( \sqrt{(\varrho - \sigma)^2} \right) = \frac{1}{2} \sum_i |\lambda_i| \quad (2.41)$$

as the trace distance, where  $\sqrt{\cdot}$  denotes here the matrix square root, see also Sec. 2.2.3.1, and  $\{\lambda_i\}$  the set of eigenvalues of the Hermitian matrix  $(\varrho - \sigma)$ . The trace distance fulfills the requirements for distances, namely [51]

- $D_T(\varrho, \sigma) \geq 0$ ,
- $D_T(\varrho, \sigma) = 0 \Leftrightarrow \varrho = \sigma$ ,
- $D_T(\varrho, \sigma) = D_T(\sigma, \varrho)$ , and
- $D_T(\varrho, \sigma) \leq D_T(\varrho, \xi) + D_T(\xi, \sigma)$ .

Its definition is equivalent to the definition of the Kolmogorov distance of probability distributions [51].

## 2.2.4 Prominent Quantum States and Their Properties

In this section, some of the most prominent quantum states needed in this thesis will be introduced.

### 2.2.4.1 Maximally Mixed State

The *maximally mixed state*, sometimes also called *completely mixed state*, is of low value in the context of quantum information as it maximizes the von Neumann entropy. From an information theoretical point of view, the maximally mixed state can be considered as a free state [76]. Irrespective of the measurement basis, the probabilities of the possible outcomes for the maximally mixed state always follow Laplace's rule of probabilities, i.e., all possible outcomes are equally likely. The  $n$ -qubit maximally mixed state  $\varrho_{\text{mms}}$  can be written as the (normalized) identity matrix and is as such independent from the basis,

$$\varrho_{\text{mms}} = \frac{\mathbb{1}_{2^n \times 2^n}}{2^n} \rightarrow U \varrho_{\text{mms}} U^\dagger = \frac{\mathbb{1}_{2^n \times 2^n}}{2^n} = \varrho_{\text{mms}} \quad (2.42)$$

for any unitary transformation  $U$ . Thus, the eigenstates have equal probabilities, i.e., the spectral components of this state have equal weight. Therefore, if considered as noise, the state is seen as *white noise*. As, independent of the basis, all density matrix elements of the maximally mixed state are populations, the state does not carry any entanglement.

### 2.2.4.2 Bell States

The maximally entangled two-qubit states denoted as *Bell states* are given by

$$|\phi^+\rangle = \frac{1}{\sqrt{2}} (|00\rangle + |11\rangle), \quad (2.43a)$$

$$|\phi^-\rangle = \frac{1}{\sqrt{2}} (|00\rangle - |11\rangle), \quad (2.43b)$$

$$|\psi^+\rangle = \frac{1}{\sqrt{2}} (|01\rangle + |10\rangle), \quad (2.43c)$$

$$|\psi^-\rangle = \frac{1}{\sqrt{2}} (|01\rangle - |10\rangle). \quad (2.43d)$$

A peculiar property of the four Bell states is the possibility to transform them into each other by applying a single qubit unitary transformation, i.e., by a local operation. A possible cycle of Bell states when only applying a unitary transformation to the first qubit is given for example by

$$|\phi^+\rangle \rightarrow \sigma_z \otimes \mathbb{1} |\phi^+\rangle = |\phi^-\rangle \quad (2.44a)$$

$$|\phi^-\rangle \rightarrow -\sigma_x \otimes \mathbb{1} |\phi^-\rangle = |\psi^-\rangle \quad (2.44b)$$

$$|\psi^-\rangle \rightarrow \sigma_z \otimes \mathbb{1} |\psi^-\rangle = |\psi^+\rangle \quad (2.44c)$$

$$|\psi^+\rangle \rightarrow \sigma_x \otimes \mathbb{1} |\psi^+\rangle = |\phi^+\rangle. \quad (2.44d)$$

Hence, the Bell states are LOCC equivalent, see Sec. 2.2.2.3. This property is made use of for *quantum teleportation*, where one of the entangled particles of a Bell state together with a third particle in a state which is to be teleported are projectively measured in the Bell state basis. The measurement result is transmitted via a classical channel to the other

party, allowing it to transform its qubit accordingly in order to retrieve the third particle's original state [37].

As being maximally entangled, Bell states are usually the main resource for almost all tasks in quantum communication or for violating a Bell inequality, see Sec. 2.2.2.6.

### 2.2.4.3 Greenberger-Horne-Zeilinger States

States of the form

$$|\text{GHZ}\rangle = \frac{1}{\sqrt{2}} (|00\dots 0\rangle + |11\dots 1\rangle) \quad (2.45)$$

with  $|0\rangle$  and  $|1\rangle$  as basis states in an arbitrary local basis are called Greenberger-Horne-Zeilinger states [77]. The two qubit GHZ state corresponds to the Bell state  $|\phi^+\rangle$ .

States of this type show several interesting features. For example, let each particle of a three qubit GHZ state be detected by means of a projective measurement onto  $|\uparrow\rangle_j = |0\rangle_j + e^{i\phi_j}|1\rangle_j$  and  $|\downarrow\rangle_j = |0\rangle_j - e^{i\phi_j}|1\rangle_j$  with variable phase  $\phi_j$  for each particle. When considering the coincidences between all possible combinations of outcomes, one observes a dependency on the sum of the three phases,  $\phi_1 + \phi_2 + \phi_3$ , while no fringes of two-fold (or single-particle) coincidences are visible [50]. This state thus contains three particle entanglement. Interestingly, as two-fold correlations are not visible in the three qubit GHZ state, the loss of already a single particle implies the complete loss of entanglement,

$$\frac{|00\dots 0\rangle_n + |11\dots 1\rangle_n}{\sqrt{2}} \rightarrow \frac{|00\dots 0\rangle\langle 00\dots 0|_{n-1} + |11\dots 1\rangle\langle 11\dots 1|_{n-1}}{2}, \quad (2.46)$$

where the subscripts  $n$  and  $n-1$ , respectively, denote the number of qubits and  $\rightarrow$  indicates the process of losing a single particle. Although the initial state has maximal entanglement (according to the typical measures), the resulting state is the incoherent mixture of  $|00\dots 0\rangle_{n-1}$  and  $|11\dots 1\rangle_{n-1}$  and is as such purely classical.

### 2.2.4.4 Graph and Cluster States

A possible implementation for optical quantum computing is the so-called one-way quantum computing based upon graph states, states emerging from Ising-type interactions between many qubits [40, 78, 79]. Performing measurements on single qubits of a multiqubit graph state together with feed-forward control and finally conducting a sequence of one-qubit readout measurements [50, 80] allows the implementation of quantum computation algorithms [81].

The usage of the notions of “graph” and “cluster” states is not perfectly consistent. Here, a graph state is generally a state which can be described using a graph with vertices and edges, where the qubits correspond to the vertices, which have to be initialized to the state  $1/\sqrt{2}(|0\rangle + |1\rangle)$ , and the edges (connections between vertices) indicate the application

of a controlled-phase gate [80],

$$U_{\text{C-phase}} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}. \quad (2.47)$$

In a more concise notation, the  $n$ -qubit graph state  $|G_{n,E}\rangle$  with set of  $j$  edges  $E = \{E_i\}_{i=1}^j$ , where each edge  $E_i$  connects vertices  $a_i$  and  $b_i$ , reads

$$|G_{n,E}\rangle = \prod_{i=1}^j U_{\text{C-phase}}^{(a_i,b_i)} \left( \frac{|0\rangle + |1\rangle}{\sqrt{2}} \right)^{\otimes n}, \quad (2.48)$$

where  $U_{\text{C-phase}}^{(a_i,b_i)}$  denotes the application of  $U_{\text{C-phase}}$  between the qubits  $a_i$  and  $b_i$ . A different definition of graph states follows the concept of the so-called *stabilizing operators* or *stabilizers*  $\{S_i\}$ . For each vertex  $i$ , which is connected via edges to the set of vertices  $N(i)$ , one can construct the operator

$$S_i = \sigma_x^{(i)} \bigotimes_{j \in N(i)} \sigma_z^{(j)}, \quad (2.49)$$

such that the graph state is simultaneous eigenstate with eigenvalue  $+1$  to all stabilizers  $S_i$ , i.e.,  $S_i|G_{n,E}\rangle = |G_{n,E}\rangle$  for every  $i$ , which is useful to construct quantum error correcting codes [82–84]. An extension to the graph states are hypergraph states, where not only two vertices can be connected by an edge, but multiple qubits by a hyperedge with multiple qubits serving as control qubits for one target qubit [85].

The cluster state is a graph state, where the vertices are aligned in a two dimensional lattice structure with edges between (at most) nearest neighbouring vertices. Please note that also the GHZ state is a graph state, which corresponds to a star-like graph with one vertex in the centre and  $n - 1$  vertices connected via edges with the centre, see Fig. 2.6.

#### 2.2.4.5 Dicke and W States

If by some mechanism,  $e$  out of  $n$  particles are coherently excited from their ground state by equally and coherently addressing all  $n$  particles, the system ends up in a Dicke state with  $n$  qubits and  $e$  excitations [86]. The resulting state can be written as

$$|D_n^{(e)}\rangle = \binom{n}{e}^{-\frac{1}{2}} \sum_i \mathcal{P}_i (|0\rangle^{\otimes(n-e)} \otimes |1\rangle^{\otimes e}), \quad (2.50)$$

where  $\mathcal{P}_i$  denotes the  $i$ -th permutation of the states of all individual qubits. Therefore,  $|D_n^{(e)}\rangle$  is invariant under permutation of qubits. If some, say,  $e$  out of  $n$ , atoms are in the excited state, while the rest is prepared in the ground state, the rate of photo emission can be increased significantly due to this coherence. This process is called *super-radiance*

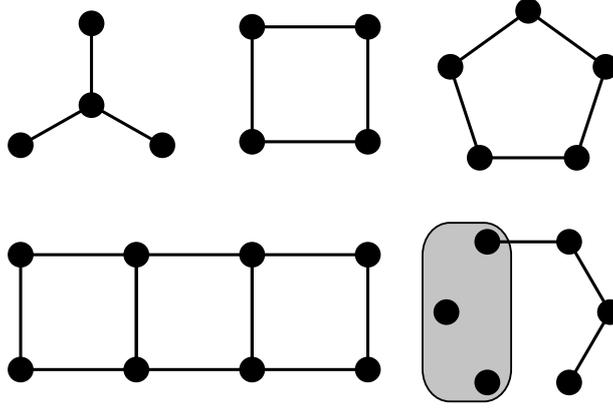


FIGURE 2.6: Different types of graph states. Each vertex corresponds to a qubit, which is initialized in state  $\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$ , while the edges denote the application of a controlled phase gate between the connected qubits. The first, star-like graph is equivalent to a GHZ state, while the square and the 2-by-4 arranged qubits are types of cluster states. The pentagon is another type of graph state. The last state is a six-qubit hypergraph state, where the hyperedge, shown as shaded region, denotes the application of a generalized phase gate.

as for a large number of atoms ( $n \gg 1$ ), the rate of emission can scale with  $n^2$  instead of  $n$  [87]. Super-radiance is maximal for systems with half excitation, i.e.,  $e = n/2$ .

A special cases of the Dicke state is the  $W$  state [62], which, for  $n$  qubits with  $e = 1$  excitation, is given by

$$|W\rangle = \frac{1}{\sqrt{n}} (|00\dots 1\rangle + \dots + |01\dots 0\rangle + |10\dots 0\rangle). \quad (2.51)$$

The state  $|\overline{W}\rangle$  contains  $e = n - 1$  excitations and therefore corresponds to  $|W\rangle$  with exchanging  $|0\rangle \leftrightarrow |1\rangle$ .

#### 2.2.4.6 Qutrit and Qudit States

Although most parts of this thesis consider compositions of  $d = 2$  dimensional systems only, a short discussion of qutrit states is of avail to understand some specialties. An example for a highly entangled qutrit state is the four qutrit GHZ state, defined as

$$|\text{GHZ}_{d=3,n=4}\rangle = \frac{1}{\sqrt{3}} \sum_{j=0}^2 |jjjj\rangle = \frac{1}{\sqrt{3}} (|0000\rangle + |1111\rangle + |2222\rangle). \quad (2.52)$$

Another example is the four qutrit cluster state written as

$$|C_{d=3,n=4}\rangle = \frac{1}{3} \sum_{i,j=0}^2 \omega^{ij} |ijij\rangle = \frac{1}{3} (|0000\rangle + |0101\rangle + \dots + \omega|1111\rangle + \dots + \omega^4|2222\rangle) \quad (2.53)$$

with  $\omega = \exp(i\pi/3)$ .

The definition for a maximally entangled state with  $n$  qudits follows directly from Eq. (2.52) and reads

$$|\Omega_{d,n}\rangle = \frac{1}{\sqrt{d}} \sum_{j=0}^{d-1} |j\rangle^{\otimes n}. \quad (2.54)$$

## 2.2.5 Measurements and Quantum Instruments

All evolutions of (closed) quantum systems are governed by the Schrödinger equation. As such they can be described by unitary transformations and are thus deterministic. However, the measurement process, i.e., the attempt to access some of the information of the quantum system and make it tangible for the classical observer, cannot be written in terms of unitary transformations (of the same closed system) and, in general, it is not possible to predict its outcomes. As it is a basic element for some of the topics discussed in this thesis, a brief introduction to the concept of measurements, quantum channels, and quantum instruments will be given in the following sections.

### 2.2.5.1 General Measurements

For a projective measurement in the basis  $\{|i\rangle\}$ , let us denote a set of projection operators  $P = \{P_i\}$  with

$$P_i \equiv |i\rangle\langle i|. \quad (2.55)$$

The projectors  $\{P_i\}$  are mutually orthogonal,

$$P_j P_k = \delta_{j,k} P_j. \quad (2.56)$$

The probability to find the system, described by the density matrix  $\varrho$ , after the measurement in the basis state  $|j\rangle$  is given by

$$p_j = \langle j|\varrho|j\rangle = \text{tr}(P_j \varrho). \quad (2.57)$$

The post-measurement state  $\tilde{\varrho}_j$  after observation of outcome  $|j\rangle$  is given by [88]

$$\tilde{\varrho}_j = |j\rangle\langle j| = \frac{P_j \varrho P_j}{\text{tr}(P_j \varrho P_j)}, \quad (2.58)$$

where the idempotency of the projector  $P_j = P_j^2$  was used in the denominator. Evidently, the post-measurement state is an eigenstate of the respective projection operator.

A straightforward extension can be derived for projections into subspaces spanned by multiple basis states. Thus, for projective measurements, the number of projection operators can be less or equal to the number of dimensions of the system.

Any physical observable, e.g., the position of a particle or the spin of an electron, is described by a Hermitian operator  $O$  (e.g.,  $x$  or  $S$ ). It can be decomposed by the spectral

decomposition theorem [80] into projection operators in the above mentioned basis of states  $\{|i\rangle\}$ ,

$$O = \sum_i o_i |i\rangle\langle i| = \sum_i o_i P_i. \quad (2.59)$$

By the linearity of the trace, one obtains for the expectation value of the observable with respect to the state  $\varrho$

$$\langle O \rangle \equiv \text{tr}(O\varrho) = \sum_i o_i \text{tr}(P_i\varrho) = \sum_i o_i p_i. \quad (2.60)$$

The expectation value is thus the sum of eigenvalues  $o_i$  of the operator  $O$  weighted by the respective probability to be observed. Eq. (2.60) is the formal expression of *Born's rule* [80, 89].

While projective measurements on the system are useful in many scenarios, it is not the most general type of measurement. In order to derive a more general class of measurements, an auxiliary system will be used as introduced by von Neumann [90, 91]. For measuring the quantum system  $\varrho$ , it can be coupled to another system, usually called *pointer*, *meter*, or *probe*, which is subsequently measured. The initially separable composite system before the coupling is in the state  $\varrho \otimes |0\rangle\langle 0|$ , where the pure state  $|0\rangle$  is used for the pointer system. According to the Naimark theorem [89], projective measurements on a subsystem of a composite system allow generalized measurements on the other subsystems. For a proper unitary evolution  $U$  describing the coupling, which entangles system and pointer, the probability for observing the outcome  $|j\rangle$  on the pointer system is obtained by [88]

$$p_j = \text{tr} \left[ \mathbb{1} \otimes P_j^{(p)} U \varrho \otimes |0\rangle\langle 0| U^\dagger \mathbb{1} \otimes P_j^{(p)} \right] = \text{tr} \left[ A_j \varrho A_j^\dagger \right] \quad (2.61)$$

where the  $P_j^{(p)}$  are rank-1 projectors acting on the pointer system (in contrast to Eq. (2.57), where they act on the system itself) and with the operators  $A_j$  describing the transformation of the actual system during the coupling. Those operators are often denoted as *effect operators* [92] or *Kraus operators* [88, 89]. As for calculating the probabilities, see Eq. (2.61), only the products  $A_j^\dagger A_j$  are relevant, one defines the *measurement operators*  $E_j$  accordingly<sup>3</sup>,

$$E_j = A_j^\dagger A_j. \quad (2.62)$$

For a complete set of (mutually orthogonal) rank-1 projection operators  $P_j^{(p)}$  acting on the pointer system, the measurement operators  $E_j$  fulfill the completeness relation

$$\sum_i E_i^\dagger E_i = \mathbb{1}. \quad (2.63)$$

By summing over all possible outcomes, this condition ensures unit probability of measurement outcomes, i.e.,  $\sum_i p_i = 1$ . The post-measurement state after observing the particular

---

<sup>3</sup>Please note that in some literature, e.g., [92], the symbol  $E_j$  is used to denote the effect operator, which, in turn, is called  $A_j$  here.

outcome  $j$  in this generalized scenario is given by

$$\tilde{\varrho}_j = \frac{A_j \varrho A_j^\dagger}{\text{tr}(A_j^\dagger A_j \varrho)}. \quad (2.64)$$

Obviously, if the measurement operators themselves are projection operators on the actual system,  $P_j = A_j^\dagger A_j$ , Eq. (2.64) simplifies to the case of Eq. (2.58). A set of measurement operators  $E = \{E_j\}$  is called *positive-operator valued measure* (POVM) [88], or sometimes *probability-operator measure* (POM), emphasizing that each element of the set corresponds to a probability operator. While this set is sufficient to calculate the probability distribution for the different outcomes, the decomposition into Kraus operators and thus the possible post-measurement states is ambiguous. It is important to note that the number of elements, i.e., the number of possible outcomes, can also be larger than the number of dimensions of the system. Furthermore, in contrast to projective measurements, the repeated measurement, described by further application of the Kraus operators in Eq. (2.64), can change the post-measurement state if they are not idempotent,  $A_j^2 \neq A_j$ .

If the projectors in Eq. (2.61) are not of rank 1, in general, a single Kraus operator is not sufficient to describe the transformation of the state. To adequately handle this case, the notion of quantum channels and quantum instruments will be introduced.

### 2.2.5.2 Quantum Channels

In quantum information, the evolution of a system often is described in terms of quantum channels. In general, a quantum channel  $T$  is a *completely positive, trace preserving linear map* from the set of density matrices in  $d_1$  dimensions to the set of density matrices in  $d_2$  dimensions.

It is obvious that a quantum channel has to be positive and trace preserving since a quantum channel shall output a positive semidefinite and normalized density matrix.  $\varrho \geq 0$  and  $\text{tr}(\varrho) = 1$  hold for the initial state, thus after transformation by the channel,  $T(\varrho) \geq 0$  and  $\text{tr}(T(\varrho)) = 1$  have to hold. The additional property of *complete* positivity comes from the requirement that composite transformations of the system by the channel together with identity transformations on auxiliary systems have to be positive independently of the dimensions of the auxiliary system. As quantum mechanics is linear, the channel cannot perform any nonlinear transformations.

The most general parametrization of a quantum channel can be given using a set of Kraus operators  $\{A_i\}$  which, for non-selective evolutions [93], fulfills  $\sum_i A_i^\dagger A_i = \mathbb{1}$ . Then, a quantum channel can be expressed as [88]

$$T(\varrho) = \sum_i A_i \varrho A_i^\dagger. \quad (2.65)$$

A quantum channel  $T$  with  $d$  dimensional input state can, according to the Choi-Jamiolkowski isomorphism [58, 94–96], be written in terms of a quantum state  $\tau$ ,

$$\tau = (T \otimes \mathbb{1}_d)(|\Omega\rangle\langle\Omega|), \quad (2.66)$$

where  $|\Omega\rangle = 1/\sqrt{d} \sum_{i=0}^{d-1} |i\rangle \otimes |i\rangle$  denotes the maximally entangled state in  $2d$  dimensions. To review that  $T$  is a valid quantum channel, it is sufficient to verify that  $\tau$  is a valid quantum state as explained in Sec. 2.2.1.3, namely that it is Hermitian, positive semi-definite, and normalized.

### 2.2.5.3 Quantum Instruments

Describing the measurement process in terms of *quantum instruments* allows to consistently formulate the transformation of the measured quantum state together with the associated measurement operator. A channel  $T$  can be decomposed according to

$$T(\varrho) = \sum_k I_k(\varrho) \quad (2.67)$$

with the set of linear maps  $\{I_k\}$ , where each map  $I_j$  can in general be decomposed using several Kraus operators,  $I_j(\varrho) = \sum_i A_i \varrho A_i^\dagger$ , where the summation takes into account the subset of Kraus operators corresponding to the map  $I_j$ . Each  $I_j$  is associated to a particular measurement outcome. Then, the measurement operators can be identified with the dual maps to the transformations of a quantum channel

$$E_j \equiv I_j^*(\mathbb{1}), \quad (2.68)$$

which can be seen from

$$\text{tr}[I_j(\varrho)] = \text{tr}[I_j(\varrho)\mathbb{1}] = \text{tr}[\varrho I_j^*(\mathbb{1})] = \text{tr}[\varrho E_j]. \quad (2.69)$$

As the quantum channel, Eq. (2.67), can in general be decomposed using several Kraus operators, see Eq. (2.65), the description of a measurement process based on quantum instruments is even more versatile than the identification of Eq. (2.62).

The condition

$$\sum_i E_i = \mathbb{1} \Leftrightarrow \sum_i I_i^*(\mathbb{1}) = \mathbb{1} \quad (2.70)$$

ensures that the channel in Eq. (2.67) is trace preserving. The set of linear maps  $I \equiv \{I_i\}$  together with the respective POVM  $E \equiv \{E_i\}$  is called quantum instrument.

## 2.2.6 Statistical and Mathematical Concepts

Several of the projects discussed in this work rely on a statistical evaluation of the measurement results. It is therefore indicated to introduce some of the fundamental concepts such as probability distributions, statistical moments, and hypothesis tests.

### 2.2.6.1 Distribution Functions

A probability distribution associates with possible outcomes of a random experiment a probability to observe this specific outcome. Using a continuous random variable, the

probability density function (pdf) can compare the relative likelihood for two different sets of outcomes. For example, consider a random variable  $X$ , which follows a normal distribution with mean 0 and variance  $\sigma^2 = 1$ ,  $X \sim \mathcal{N}(0, \sigma^2)$ . While the probability to (exactly) observe the outcome 0 vanishes, it is twice as likely to observe 0 than, say, 1.1774 as  $\Pr(X = 0)dx \approx 2\Pr(X = 1.1774)dx$  with  $\Pr(X = X_0)$  denoting the probability to observe the outcome  $X_0$  for the random variable  $X$ .

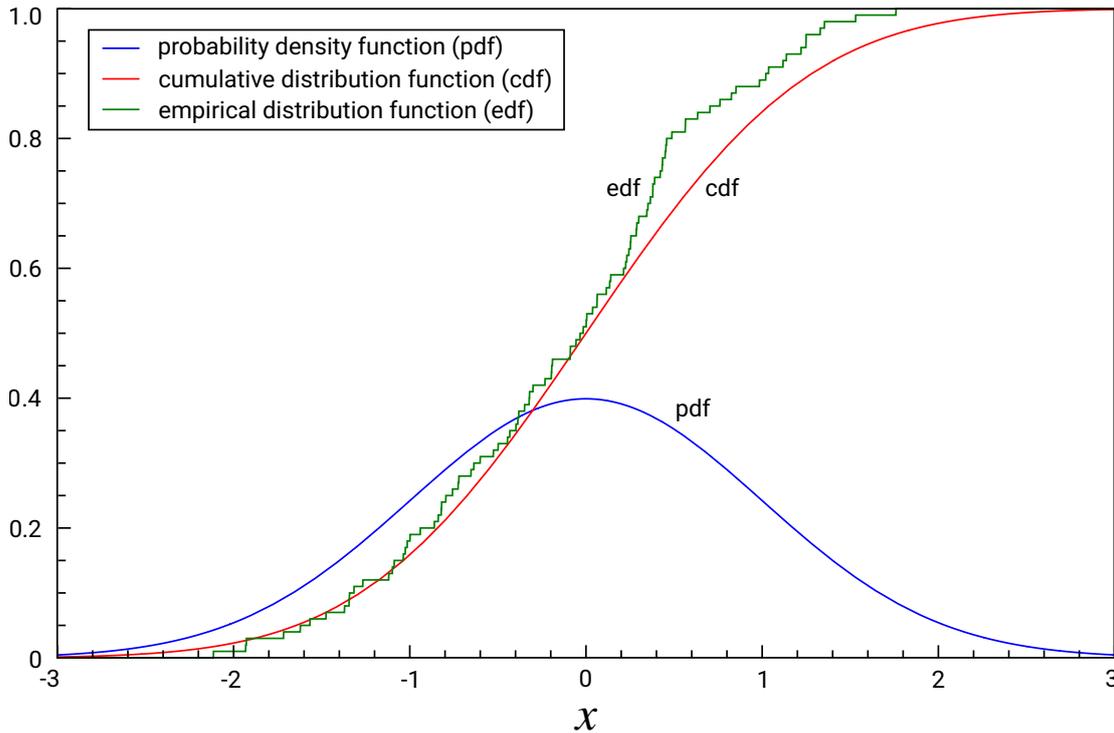


FIGURE 2.7: The probability density function (blue) and the cumulative distribution function (red) for a normal distribution with mean 0 and variance 1 together with an empirical distribution function (green) for 100 samples following the normal distribution. The empirical distribution function, which represents the set of measured data, takes a step of  $1/100$  at every observed outcome.

The probability density function is a useful tool for representing ideal distributions. For many practical purposes such as obtaining random numbers following a specific distribution or using statistical tests, however, the cumulative distribution function (cdf) is more suitable. The cumulative distribution function can be obtained as the integral over the probability density function  $f(x)$  via

$$F(x) = \int_{-\infty}^x d\tilde{x} f(\tilde{x}). \quad (2.71)$$

It is thus a measure for the probability to observe an outcome smaller than or equal to the given observation. For comparing observed distributions with theoretical distributions, the

concept of cumulative distributions is of large avail. For the observed random outcomes  $\{x_i\}_i$ , the empirical distribution function is defined as

$$\hat{F}_n(x) = \frac{|\{x_i : x_i \leq x\}|}{|\{x_i\}|}, \quad (2.72)$$

where  $|\cdot|$  denotes the cardinality of the corresponding set of elements. It thus counts how many elements are smaller or equal to the argument  $x$ .

For an example of the probability density function, the cumulative distribution function and a possible empirical distribution function, see Fig. 2.7.

### 2.2.6.2 Statistical Moments

Statistical moments quantify the spread of probability distributions. Given a probability density function  $f(x)$ , its  $k$ -th moment  $m_k^{(r)}$  can be computed by

$$m_k^{(r)} = \int_{-\infty}^{\infty} dx f(x) x^k. \quad (2.73)$$

These moments are also referred to as *raw moments* or *moments about the origin*. Shifting the density function  $f(x) \rightarrow (x - \Delta x)$ , changes all thus obtained moments. In order to describe the shape of the density distribution regardless of any such shifts, one resorts to the *central moments* (*moments about the mean*),

$$m_k^{(c)} = \int_{-\infty}^{\infty} dx f(x) (x - m_1^{(r)})^k. \quad (2.74)$$

The zeroth raw and central moments correspond to the normalization and should therefore be  $m_0^{(r)} = m_0^{(c)} = 1$  for properly normalized probability density functions. The first raw moment matches the mean value; the first central moment accordingly vanishes. The second central moment,  $m_2^{(c)}$ , is denoted as *variance*. Relating the next higher central moments with the variance, one obtains the *skewness*,  $m_3^{(c)} / (m_2^{(c)})^{3/2}$ , and the *kurtosis*,  $m_4^{(c)} / (m_2^{(c)})^2$ , which are measures for the asymmetry and concentration to the center, respectively.

Please note that for some distributions, the statistical moments are not well defined. For example, the probability density of the Cauchy distribution<sup>4</sup> is given by  $f_{x_0, \gamma}(x) = \frac{1}{\pi \gamma} \frac{\gamma^2}{(x - x_0)^2 + \gamma^2}$  with some parameters  $\gamma$  and  $x_0$ . The first moment, i.e., the mean value, is to be obtained from the integral  $\int_{-\infty}^{\infty} dx f_{x_0, \gamma}(x) x \propto \int_{-\infty}^{\infty} dx \frac{x}{(x - x_0)^2 + \gamma^2}$ . The primitive of  $f_{x_0, \gamma}(x) x$  is proportional to  $\int dx \frac{x}{(x - x_0)^2 + \gamma^2} = \frac{x_0}{\gamma} \tan^{-1} \left( \frac{x - x_0}{\gamma} \right) + \frac{1}{2} \log [(x - x_0)^2 + \gamma^2]$ , which diverges for  $x \rightarrow \pm\infty$ . Thus, the first moment (and all higher) moments do not exist for the Cauchy distribution.

<sup>4</sup>With additional intensity scaling, this distribution is in physics often denoted as Lorentz distribution used to describe spectral lines.

The so-called *moment generating function* of a probability density  $f(x)$  is defined as

$$M_X(t) = \int_{-\infty}^{\infty} dx e^{tx} f(x). \quad (2.75)$$

If all moments of  $f(x)$  are finite and  $M_X(t)$  has a positive radius of convergence, the moment generating function can be expanded as [97]

$$M_X(t) = \sum_{j=0}^{\infty} \frac{t^j m_j^{(r)}}{j!}. \quad (2.76)$$

Thus, under those conditions (finite moments and convergence), the moment generating function uniquely determines the distribution [97]. Distributions of compact support fulfill those requirements and are hence uniquely defined by their moments.

To show that for the equality of two distributions above mentioned conditions are required, i.e., that it is not sufficient that all moments of both distributions are identical, consider the following counterexample [98] with the two distributions

$$f_0(x) = \sqrt{2\pi} \frac{1}{x} \exp\left[-\frac{(\log x)^2}{2}\right], \quad (2.77)$$

$$f_1(x) = f_0(x) [1 + a \sin(2\pi \log x)], \quad (2.78)$$

which both have the  $k$ -th raw moment equal to  $m_k^{(r)} = \exp(k^2/2)$ . However, the moment generating function

$$M_X(t) = \sum_{j=0}^{\infty} \frac{t^j \exp(j^2/2)}{j!} \quad (2.79)$$

does not converge. Thus, although all moments of those distributions are equal, one indeed cannot conclude that the distributions are equal. Additionally, as mentioned above, one has to check that all moments are finite and that  $M_X(t)$  has a positive radius of convergence.

### 2.2.6.3 Hypothesis Testing and p-values

A useful tool to evaluate some assumptions about an unknown distribution based on a set of recorded data drawn from this distribution is given by hypothesis tests [99]. An appropriately chosen hypothesis test allows to exclude an initial hypothesis, the so-called “null hypothesis”, concerning the model of how that data might originate. As a guideline, the important steps for testing a statistical hypothesis are the following

1. Formulate a null hypothesis and the corresponding anti hypothesis as the negation of the null hypothesis.
2. Decide for a level of significance  $\alpha$ , which will be used as a threshold to reject the null hypothesis. Often, a value of 5% is chosen.

3. Choose a proper hypothesis test, see below.
4. Find the distribution of the so-called *test statistic* as defined by the chosen hypothesis test under the assumption that the null hypothesis is true.
5. Find the *critical test statistic*  $T^*$ , such that the probability of more extreme observations than  $T^*$  corresponds to the chosen level of significance  $\alpha$ .
6. Calculate for the actually observed data the test statistic  $T_{\text{obs}}$  and check if  $T_{\text{obs}}$  is in the region of more extreme results than  $T^*$ . Additionally, a  $p$ -value can be calculated, which indicates the probability of results at least as extreme as  $T_{\text{obs}}$ . If  $T_{\text{obs}}$  is in the region of more extreme results than  $T^*$  or, correspondingly, the  $p$ -value is below  $\alpha$ , the null hypothesis is to be rejected and the properly stated anti hypothesis has to be true. Otherwise, there is no statistical evidence to reject the null hypothesis. The test thus concludes indifferently.

Please keep in mind that not rejecting the null hypothesis does not mean that the null hypothesis has to be true. It is possible that there is solely not enough evidence for exclusion.  $\alpha$  directly states the probability of falsely rejecting the null hypothesis because it indicates what observations are considered to be extreme, although the null hypothesis produces with rate  $\alpha$  these and more extreme results. The  $p$ -value provides information about the probability of observing data  $D$  given the null hypothesis  $H_0$ ,

$$p = \Pr(D|H_0). \quad (2.80)$$

It is important to note that it does not state the probability that the null hypothesis is correct or another hypothesis is wrong,  $\Pr(D|H_0) \neq \Pr(H_0|D)$  [100].

Fortunately, many of the above given steps are already taken care of in standard data evaluation software. One is left with the decision for the correct test, a significance level  $\alpha$ , and the proper interpretation of the resulting  $p$ -value.

Because needed for several projects in this thesis, a brief overview of some hypothesis tests with corresponding test statistic will be given. These tests evaluate the null hypothesis that a data set with  $n$  observations can be described by a distribution  $F_0(x)$ . For that purpose, the empirical distribution function  $\hat{F}_n(x)$  is derived according to Eq. (2.72). The three very common tests below distinguish from each other by the definition of the respective test statistic  $T$ .

The test statistic of the *Kolmogorov-Smirnov test* [101–104] is given by

$$T_{\text{KS}} = \sup_x \left| \hat{F}_n(x) - F_0(x) \right|. \quad (2.81)$$

Thus, the largest deviation between the empirical and cumulative distribution function is evaluated.

The *Cramér-von Mises test* [104–106] uses a quadratic measure as the distance between two different distributions. In contrast to the Kolmogorov-Smirnov test, not only the

largest deviation between both distributions matters, but rather the cumulative deviation of all samples. The test statistic reads

$$T_{\text{CM}} = n \int \left( \hat{F}_n(x) - F_0(x) \right)^2 dF_0(x). \quad (2.82)$$

A small modification from the Cramér-von Mises test allows to change the weights in the distance measure of the distributions. The *Anderson-Darling test*, following the test statistic

$$T_{\text{AD}} = n \int \frac{\left( \hat{F}_n(x) - F_0(x) \right)^2}{F_0(x)(1 - F_0(x))} dF_0(x), \quad (2.83)$$

thus puts a larger weight to the tails of the distribution due to inversely weighting with  $F_0(x)$  and  $1 - F_0(x)$ . Hence, the Anderson-Darling test is very sensitive to deviations in the tails and is therefore useful if one expects specific features of the distribution  $F_0(x)$  of the null hypothesis in its tails.

#### 2.2.6.4 Significance, Confidence, Power, and Errors

For testing statistical statements using hypothesis tests, one first has to decide the strictness of the used tests and with it the acceptance rate of different kinds of errors. In the ideal case, if a hypothesis is true (false), the corresponding test should give the result *true* (*false*). However, if the hypothesis test leads to the rejection of the null hypothesis, while it actually is true (false negative), one makes an error of *first kind* (or type 1). The contrary case of a false positive is called error of *second kind* (or type 2).

The significance level, denoted with  $\alpha$  in Sec. 2.2.6.3, directly corresponds to the error rate of errors of type 1. For a level of  $\alpha = 0.05$ , one accepts that in 5% of the cases, one rejects the hypothesis, which is in fact true. Thus, one has confidence of  $1 - \alpha$  in the reported result. With  $\beta$  being the rate of error of false negatives, one can state the power of a hypothesis test by  $1 - \beta$ .

Obviously, one can construct a (trivial) hypothesis test with a confidence of 100% by using the significance level  $\alpha = 0$ . Because this test does not risk any false positives, it will never reject the null hypothesis. However, this leads to a large rate  $\beta$  of false negatives and thus to a reduction of the test's power  $1 - \beta$ . One realizes that this test with low power and no significance should be replaced by a test giving less confidence, but having higher significance and more power.

#### 2.2.6.5 Likelihood

Similar as in the hypothesis tests discussed in Sec. 2.2.6.3, where one evaluates the  $p$ -value based on a set of data, one can specify the *likelihood* of a specific parameter given a set of measurement data. Suppose a coin is given with probabilities  $p_{\text{heads}}$ ,  $p_{\text{tails}} = 1 - p_{\text{heads}}$  for the outcomes *heads* and *tails*, respectively. If these probabilities are known, one can easily

calculate the probability for observing  $n_{\text{heads}}$  times *heads* and  $n_{\text{tails}}$  times *tails* according to

$$P_{p_{\text{heads}}, p_{\text{tails}}}(n_{\text{heads}}, n_{\text{tails}}) = \binom{n_{\text{heads}} + n_{\text{tails}}}{n_{\text{heads}}} p_{\text{heads}}^{n_{\text{heads}}} p_{\text{tails}}^{n_{\text{tails}}}, \quad (2.84)$$

where the probabilities  $p_{\text{heads}}$  and  $p_{\text{tails}}$  are considered to be fixed parameters instead of random variables.

Contrary, if one has observed a specific pattern and intends to estimate possible probabilities for the different outcomes, the question has to be rephrased: given the observed data, what is the *likelihood*  $\mathcal{L}$  for a specific parameter  $p_{\text{heads}}$  describing the probability distribution? Consequently, the likelihood  $\mathcal{L}$  corresponds to the inverse question,

$$\mathcal{L}(p_{\text{heads}}, p_{\text{tails}} | n_{\text{heads}}, n_{\text{tails}}) = P_{p_{\text{heads}}, p_{\text{tails}}}(n_{\text{heads}}, n_{\text{tails}}). \quad (2.85)$$

The example in Eq. (2.85) can be generalized for  $k$  outcomes [107],

$$\mathcal{L}(p_1, \dots, p_k | n_1, \dots, n_k) = \frac{(\sum_i n_i)!}{\prod_i n_i!} \prod_i p_i^{n_i}, \quad (2.86)$$

where  $p_i$  is the probability to observe the outcome  $i$ , which was experimentally measured  $n_i$  times.

According to Bayes theorem [97],

$$P(A|B) = P(B|A) \frac{P(A)}{P(B)} \quad (2.87)$$

holds. Obviously, by comparing Eq. (2.86) with Eq. (2.87), one perceives that the term  $P(p_1, \dots, p_k) / P(n_1, \dots, n_k)$  does not appear in Eq. (2.86). As a consequence, the likelihood is not a probability and correspondingly not properly normalized, i.e., in general

$$\int dp_1 \dots dp_k \mathcal{L}(p_1, \dots, p_k | n_1, \dots, n_k) \neq 1, \quad (2.88)$$

while  $\sum_{n_1, \dots, n_k} P(n_1, \dots, n_k | p_1, \dots, p_k) = 1$  for fixed total number of outcomes,  $\sum_i n_i = N$ . This corresponds to the fact that the likelihood does not take into account that a set of probabilities  $\{p_1, \dots, p_k\}$  might be less probable than another set  $\{p'_1, \dots, p'_k\}$ , but result in a larger likelihood, i.e., the *prior* (see Sec. 2.2.6.8) is neglected. As an example, consider a series of coin tosses resulting in  $n_h = 6$  times head and  $n_t = 4$  times tails. The integral over all possible values for  $(p_h, p_t = 1 - p_h)$  results in  $1/5544$  and is by itself meaningless, while for fixed probabilities  $(p_h = 0.6, p_t = 0.4)$  the sum of the probabilities of all possible outcomes is 1.

### 2.2.6.6 Maximum Likelihood

One approach of statistical inference, i.e., obtaining unknown parameters of some distributions based on the observed data, is based on using the likelihood function as introduced in

Sec. 2.2.6.5. As the likelihood quantifies how likely some specific parameters for describing a distribution for certain observed outcomes, its maximization gives the (set of) parameters which are most likely to have resulted in the observed outcomes. Please keep in mind that this procedure does not take into account how probable those parameters are in the first place, see the discussion around Bayes' theorem, Eq. (2.87).

The *maximum likelihood estimate* (MLE) for  $k$  probabilities, which are used as the parameters of a maximization, is obtained by

$$(\hat{p}_1, \dots, \hat{p}_k) = \arg \max_{(p_1, \dots, p_k)} \mathcal{L}(p_1, \dots, p_k | n_1, \dots, n_k) \quad (2.89)$$

for the  $n_i$  being the number of observations of the outcome to which  $p_i$  is the associated probability and with  $\mathcal{L}$  defined according to Eq. (2.86). For the example of the 10 coin tosses of Sec. 2.2.6.5 with  $n_h = 6$  times head and  $n_t = 4$  times tails, the likelihood peaks for  $(p_h = 0.6, p_t = 1 - p_h = 0.4)$  with about  $\mathcal{L} \approx 5 \cdot 10^{-4}$ , while, e.g., a fair coin results in about  $\mathcal{L} \approx 4 \cdot 10^{-4}$ .

Often, one does not use the likelihood directly, but instead the negative log-likelihood,

$$\begin{aligned} \tilde{\mathcal{L}}(\{p_i\}|\{n_i\}) &\equiv -\log[\mathcal{L}(\{p_i\}|\{n_i\})] = -\log\left[\frac{(\sum_i n_i)!}{\prod_i n_i!} \prod_i p_i^{n_i}\right] \\ &= -\log\left[\left(\sum_i n_i\right)!\right] + \sum_i \log(n_i!) - \sum_i n_i \log p_i, \end{aligned} \quad (2.90)$$

because it is numerically easier to handle and factors as from the number of permutations only lead to some offset. Due to the minus sign, one has to minimize the expression of Eq. (2.90) in order to find the maximum likelihood estimate. As all but the last term are independent of the probabilities  $\{p_i\}$ , they do not need to be considered for the maximum likelihood estimate. Eq. (2.90) is similar to the so-called *Kullback-Leibler distance*<sup>5</sup> or *Kullback-Leibler divergence* [108], which as a relative entropy measures the divergence between two probability distributions [51].

The maximum likelihood estimation procedure delivers a single value for each parameter without any information regarding its significance or errors. Consequently, the maximum likelihood estimator is a *point estimator*.

### 2.2.6.7 Point Estimators

The method of point estimation is a frequently used technique for statistical inference [107], in which a physical quantity of interest is measured to subsequently obtain a single estimate as a description of the data sample [109]. The exact procedure of parameter estimation has to be decided upon, which is sometimes also called *measurement problem*. Please note that this is obviously to be distinguished from the measurement problem in quantum mechanics

<sup>5</sup>Due to the asymmetry in arguments, the Kullback-Leibler divergence is not an actual distance measure.

(where a non-unitary evolution accompanies the observation of distinct outcomes), see Sec. 2.2.5.

In contrast, interval estimators result in an estimate for the unknown parameter with additional information concerning its uncertainty. If an estimator is expected to deliver the true value for the unknown parameter, i.e., the estimates symmetrically scatter around the true, unknown parameter, it is said to be *unbiased*. Thus, the cases of over- and underestimation of the parameter will balance [107]. For special types of distributions, one can distinguish between *mean bias* and *median bias*. An estimator which is asymptotically unbiased in the limit of large sample size is called *consistent* or *almost unbiased* [110].

### 2.2.6.8 Frequentistic and Bayesian inference

For estimating quantum states - and more generally for estimating unknown parameters in any experimental setting - one can refer to different methods of statistical inference which are motivated by different concepts of randomness. This section briefly opposes the concept of objective randomness, as mostly used by frequentists, to the subjective concept. On the other hand, according to the frequentistic understanding, experiments are conducted in order to retrieve statistical events allowing to infer from their relative occurrence some unknown but constant probability [107]. Except for the choice of the used estimator, all deduction is solely due to observed events. On the other hand, according to the Bayesian approach, one initially assumes that the unknown parameter, which one is interested in, follows some known probability distribution described by the so-called *prior*. Conducting and repeating the experiment allows one to continuously update the degree of belief in different values of the parameter<sup>6</sup>. Eventually, one obtains the *posterior*, a distribution describing the final belief about the parameter [111]. Consequently, the posterior depends on the initial distribution for the prior [107]. Sometimes, one implicitly assumes a uniform prior, which might be inadequate for the respective problem. For an illustrative example, see the *two envelope problem* in, e.g., [88].

The discussed concepts of hypothesis testing and *p*-values, see Sec. 2.2.6.3, belong to the set of methods of frequentistic inference as much as the concept of *confidence intervals*, which will be introduced in the next section. In contrast, *credibility intervals*, see Sec. 2.2.6.10, are the estimated intervals in the Bayesian methodology.

### 2.2.6.9 Confidence Intervals

Assume, a single parameter  $\theta$  is to be estimated and one is interested in a method of finding intervals, which contain the true value in many instances of estimation. The true, but unknown value shall be denoted by  $\theta_0$ . In the objectivists' interpretation of probabilities, the true value  $\theta_0$  is a fixed parameter, which is not subject to any probabilities. The experimenter conducts measurements by drawing samples from some probability distribution

---

<sup>6</sup>This idea gave also rise to the understanding of quantum states as states of belief instead of states of nature. Two distinct persons might describe the same system by different states due to their different beliefs [111].

determined by  $\theta_0$ , allowing to deduce some information about its true value. In order to estimate this unknown value based on some observation  $E'$ , one may want to find some interval

$$\underline{\theta}(E') \leq \theta_0 \leq \bar{\theta}(E') \quad (2.91)$$

limiting the true value  $\theta_0$ , see [110]. Intuitively, one might formulate this expression in terms of a probability, namely demand that “the true value lies within the given bounds with a probability of, say, 0.95”. However, this is a typical misunderstanding of a confidence interval [112], as the  $\theta_0$  is a fixed, but unknown value, which is not subject to a probabilistic experiment. Therefore, the probability for the true value to lie within the interval is exactly 0 or 1 [110]. In the objective picture, one can thus not assign a probability of, say, 95 % to finding the true value within the given interval. The statement has to be phrased differently. If a confidence interval corresponding to a given confidence level is given, other intervals can be constructed from independent observations by the same procedure and means. Thus, 95 % of the confidence intervals constructed from these observations contain the true value. Please note that this is different from asserting that a single confidence interval contains the true value with 95 % probability [112].

### 2.2.6.10 Credibility Intervals

In contrast to a confidence interval, where the true value  $\theta_0$  of a parameter  $\theta$  is treated to be a fixed, but unknown value and the confidence interval’s bounds themselves are random variables, in the Bayesian mindset  $\theta_0$  is considered to be a random variable [113]. Then, based on the prior distribution of belief together with the set of observed data, one can deduce a posterior distribution [114]. Any interval of the posterior distribution with content  $1 - \alpha$  corresponds to a credibility interval which contains, according to the Bayesian framework, the true value with credibility level  $1 - \alpha$ . Hence, there are infinitely many credibility intervals  $\mathcal{C} = \{\mathcal{C}_i\}$  with

$$\int_{\mathcal{C}_i} dx f(x) = 1 - \alpha \quad (2.92)$$

for the posterior distribution  $f(x)$  and the credibility level  $1 - \alpha$ . Often it is of interest, to find the smallest credibility interval or region to a given credibility level,

$$\mathcal{C}_0 = \arg \min_{\tilde{\mathcal{C}} \in \mathcal{C}} |\tilde{\mathcal{C}}|, \quad (2.93)$$

where  $|\tilde{\mathcal{C}}|$  denotes the posterior content of  $\tilde{\mathcal{C}}$ .

## 2.3 Experimental Framework

In this section, the linear and non-linear optical devices are introduced, which are used for the experimental demonstrations presented in this thesis. First, optical devices will be

described which are used to transform the polarization of light. These, together with beam splitters, will be grouped to devices of linear optics. Afterwards, non-linear optical devices used for creation of entangled pairs of photons will be discussed.

### 2.3.1 Linear Optics

With linear optics, one usually refers to processes in which the response of the system, i.e., the transformation from the initial to the final state, does not depend on the intensity of the light. Typical elements satisfying this condition are birefringent devices such as waveplates, YVO<sub>4</sub> crystals, and other phase shifters as well as polarizing and non-polarizing beam splitters.

#### 2.3.1.1 Waveplates

The standard tools to manipulate the polarization degree of freedom are waveplates. They are made of birefringent material of accurately chosen thickness. Light propagating through it acquires different phases for the ordinary (perpendicular to the optical axis) and extraordinary (parallel to the optical axis) polarization components given by the birefringency. The axis along which the index of refraction is smaller is called *fast axis* because of its higher phase velocity compared to the axis with larger index of refraction (*slow axis*). For quartz, the index of refraction of extraordinarily polarized light,  $n_e$ , is larger than for ordinarily polarized light,  $n_o$ , therefore, the ordinary axis is the fast axis. The most prominent representatives thereof are the half-wave plate (HWP) and the quarter-wave plate (QWP), for which the two polarization components acquire a phase of  $\lambda/2$  and  $\lambda/4$ , respectively.

The half-wave plate can transform any linear polarization into any other linear polarization. Its transformation matrix reads in the basis of fast and slow axes

$$U_{\text{HWP}} \propto \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \quad (2.94)$$

In a basis rotated by  $\theta$ , the transformation becomes

$$U_{\text{HWP}}(\theta) \propto \begin{pmatrix} \cos 2\theta & \sin 2\theta \\ \sin 2\theta & -\cos 2\theta \end{pmatrix}, \quad (2.95)$$

where  $\theta$  denotes the angle between the horizon and the fast axis.

Obtaining or measuring a circularly polarized state requires additional degrees of freedom. A quarter-wave plate allows to transform linearly polarized states into states with circular polarization and vice-versa. The general transformation matrix for a quarter-wave plate in the basis of fast and slow axes is given by

$$U_{\text{QWP}} \propto \begin{pmatrix} 1 & 0 \\ 0 & \exp(i\pi/2) \end{pmatrix} \quad (2.96)$$

and becomes

$$U_{\text{QWP}}(\theta) \propto \begin{pmatrix} \cos^2 \theta + i \sin^2 \theta & (1 - i) \sin \theta \cos \theta \\ (1 - i) \sin \theta \cos \theta & \sin^2 \theta + i \cos^2 \theta \end{pmatrix} \quad (2.97)$$

in a basis rotated by the angle  $\theta$ . Examples of the transformations are shown in Fig. 2.8.

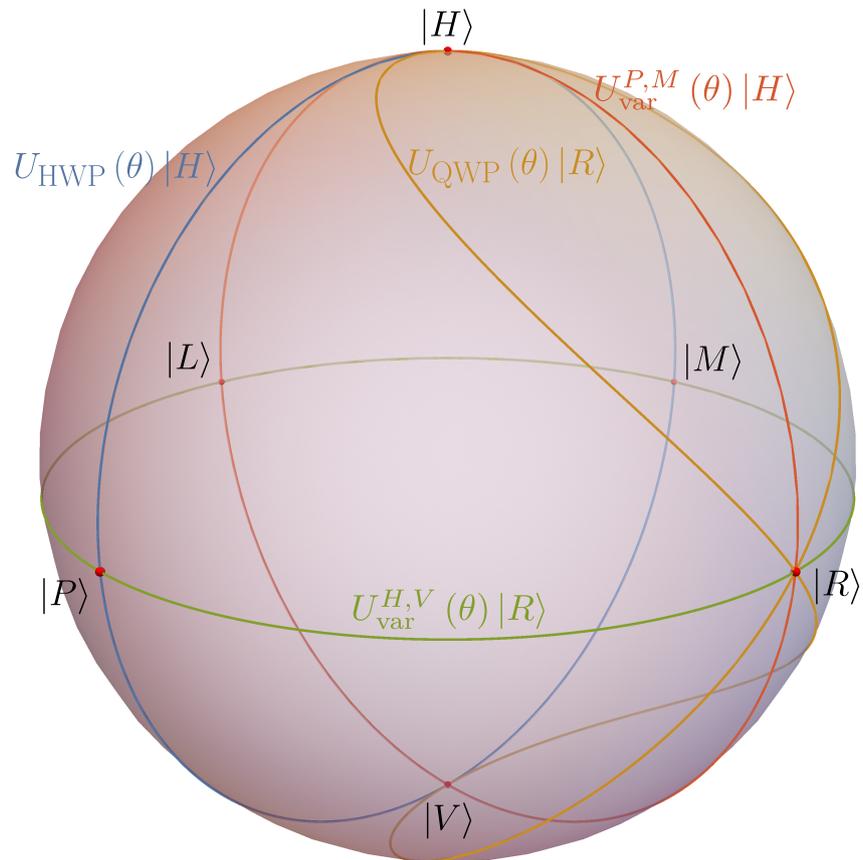


FIGURE 2.8: Transformations of a half-wave plate acting on the initial state  $|H\rangle$ , a quarter-wave plate acting on  $|R\rangle$ , and two differently aligned  $\text{YVO}_4$  crystals acting on  $|H\rangle$  and  $|R\rangle$ , respectively. By properly choosing combinations and angles of waveplates and birefringent crystals, all local unitary transformations are possible.

### 2.3.1.2 Variable Phase Retardance

Another usage of birefringent crystals is to set a *variable* relative phase between two orthogonal polarization components, e.g.,  $|H\rangle$  and  $|V\rangle$  polarization. When fast and slow axes coincide with the horizontal and vertical axes, respectively, a tilt of the crystal corresponds to effectively changing its thickness. This results in the transformation

$$U_{\text{var}}^{H,V}(\phi) \propto \begin{pmatrix} 1 & 0 \\ 0 & e^{i\phi} \end{pmatrix}, \quad (2.98)$$

where tilting varies the phase  $\phi$ . The change of the global phase is neglected. If instead the main axes of the crystal are rotated such that the relative phase between  $|P\rangle$  and  $|M\rangle$  can be set, the transformation matrix becomes (up to a global phase)

$$U_{\text{var}}^{P,M}(\phi) \propto \begin{pmatrix} 1 + e^{i\phi} & 1 - e^{i\phi} \\ 1 - e^{i\phi} & 1 + e^{i\phi} \end{pmatrix}. \quad (2.99)$$

Those transformations allow to obtain states lying on a great circle of the Bloch sphere, see Fig. 2.8. For the experiments presented in this thesis, a crystal made of Yttrium orthovanadate ( $\text{YVO}_4$ ) was used for introducing variable phases.

### 2.3.1.3 Beam Splitters

Another crucial element for setting up optical interferometric experiments are beam splitters. One distinguishes between polarizing and non-polarizing beam splitters as well as between balanced and unbalanced ones. The discussion here is limited to lossless beam splitters. For consequences of using (and tuning) the loss of beam splitters, see, e.g., [115]. Using the notation of creation and annihilation operators, one can describe a balanced, non-polarizing beam splitter with input modes  $a$  and  $b$  and output modes  $c$  and  $d$  by the transformations [49, 50, 116]

$$\hat{a}^\dagger \rightarrow \hat{c}^\dagger + i\hat{d}^\dagger, \quad (2.100a)$$

$$\hat{b}^\dagger \rightarrow i\hat{c}^\dagger + \hat{d}^\dagger. \quad (2.100b)$$

Thus, a photon in state  $\hat{a}^\dagger|0\rangle$  is transformed into a photon in a superposition of  $\hat{c}^\dagger|0\rangle$  and  $\hat{d}^\dagger|0\rangle$  with the corresponding phase. An incoming photon at port  $b$  behaves accordingly.

For a polarizing beam splitter, however, an additional index to the creation and annihilation operators has to be used in order to indicate the respective polarization. Then, one obtains for the transformations of the polarizing beam splitter (see Fig. 2.9)

$$\hat{a}_H^\dagger \rightarrow \hat{c}_H^\dagger, \quad (2.101a)$$

$$\hat{a}_V^\dagger \rightarrow i\hat{d}_V^\dagger, \quad (2.101b)$$

$$\hat{b}_H^\dagger \rightarrow \hat{d}_H^\dagger, \quad (2.101c)$$

$$\hat{b}_V^\dagger \rightarrow i\hat{c}_V^\dagger. \quad (2.101d)$$

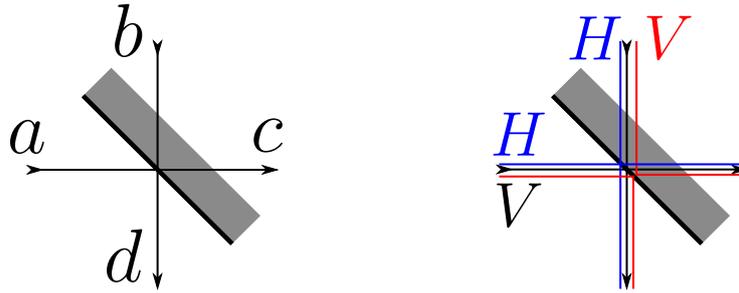


FIGURE 2.9: A non-polarizing (left) and a polarizing (right) beam splitter. Usually, one side of plate beam splitters is coated, such that light incoming from the other side first has to propagate through the material. Unintended reflections at the uncoated side can lead to disturbances of the spatial mode by a secondary mode, called “ghost”.

### 2.3.2 Non-linear Optics

In contrast to linear optics, the term *non-linear optics* is used to describe devices which can transfer energy between different frequencies. Due to the presence of the light field, the optical properties of the medium are changed, which can alter the properties of the light field [117]. Another type of non-linear process which does not depend on the pump power is the effect of *spontaneous parametric down conversion* [50], which will subsequently be discussed here.

#### 2.3.2.1 Spontaneous Parametric Down Conversion

In order to produce pairs of entangled photons, a process called spontaneous parametric down conversion is employed, where in an optically nonlinear crystal a pump photon can decay into two photons of correspondingly lower energy [87, 118–121]. While the incoming photon is usually called *pump photon*, the produced photons are denoted as *signal* and *idler* photon. This process has to conserve the initial energy  $\hbar\omega_p$  and initial momentum  $\hbar\vec{k}_p$ , therefore the *phase matching conditions* [117]

$$\omega_p = \omega_s + \omega_i, \quad (2.102)$$

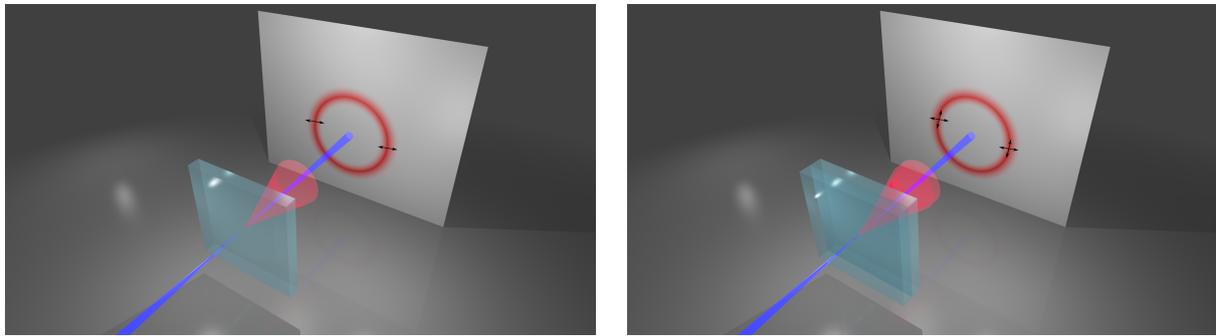
$$\vec{k}_p = \vec{k}_s + \vec{k}_i \quad (2.103)$$

imply constraints on the wavelengths and directions of the resulting photons. Down conversion cannot occur in vacuum, instead, it relies on the non-linearity of a suitable material. In an anisotropic, optically non-linear crystal, as needed for parametric down conversion, the induced polarization is described by [122]

$$P_i = \epsilon_0 \left( \sum_j \chi_{ij}^{(1)} E_j + \sum_{j,k} \chi_{ijk}^{(2)} E_j E_k + \dots \right), \quad (2.104)$$

where  $\epsilon_0$  denotes the vacuum permittivity and  $P_i$  the  $i$ -th component of the polarization vector  $\vec{P}$ .  $\chi^{(n)}$  is the  $n$ -th order of the susceptibility tensor. Higher orders of the susceptibility, such as  $\chi^{(3)}$  are used, e.g., for four wave mixing or the optical Kerr effect [117]. Please note that the process of down conversion is spontaneous. Thus, the produced signal and idler photons are created at random times.

Since there is a plethora of different materials available including variations of their usages regarding the types of phase matching, this overview will restrict itself to the types relevant for this thesis. Furthermore, only energy degenerate states are considered, i.e., only the case of  $\omega_s = \omega_i$  will be considered here. As all down conversion crystals used throughout this thesis are made of  $\beta$ -Barium borate (BBO), its operation will now exemplarily be described in Figs. 2.10 and 2.11 and in the following.



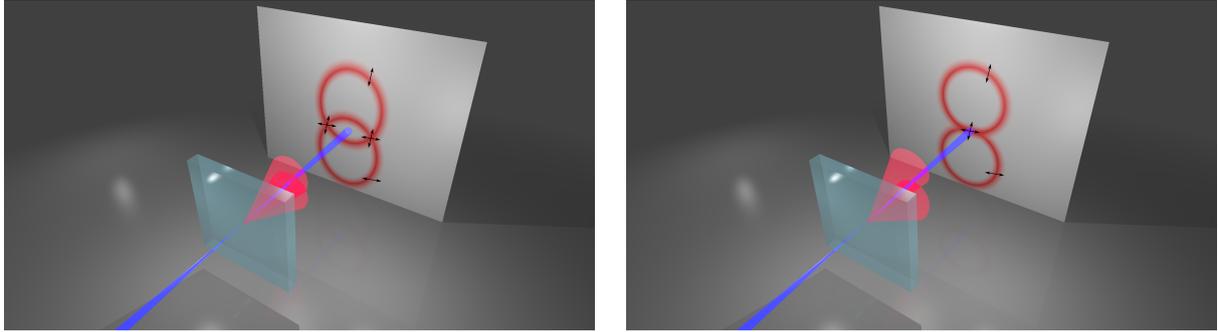
(A) Down conversion with type I phase matching conditions, where the pump polarization is orthogonal to the optical axis. Photons from a pump beam can create pairs of photons, which are identically polarized. Due to momentum conservation, both photons are emitted on a cone (a ring in projection). For coincidences, detectors are placed at diametrically opposed points.

(B) For creating polarization entanglement with type I-cut crystals, two crystals are combined such that their optical axes are orthogonal and, hence, the first one can create a pair of photons with one polarization, while the second crystal creates orthogonally polarized photons. The resulting quantum state can be, e.g.,  $1/\sqrt{2} (|HH\rangle + e^{i\phi}|VV\rangle)$ .

FIGURE 2.10: Down conversion with type I phase matching.

One distinguishes among other types between phase matching type I and type II. In a type I scenario, the pump photon is extraordinarily polarized with respect to the crystal axes and can create two ordinarily and thus equally polarized down converted photons [121]. Depending on the crystal, a typical pair production could thus look like  $|H\rangle \rightarrow |VV\rangle$ , see Fig. 2.10a. This state is a product state and thus not of large interest for the creation of entangled photon pairs. However, one can now utilize two crystals, which are rotated by  $90^\circ$  to each other, such that, say, the first and second crystal are capable of producing  $|VV\rangle$  and  $|HH\rangle$ , respectively. By pumping these sandwiched crystals with a coherent superposition state such as  $1/\sqrt{2} (|H\rangle + |V\rangle)$ , one can produce the polarization entangled two-photon state  $1/\sqrt{2} (|HH\rangle + e^{i\phi}|VV\rangle)$  with some relative phase  $\phi$ , see Fig. 2.10b.

For differently cut crystals, one can employ so-called type II down conversion. There,



(A) Down conversion with non-collinear type II phase matching. Extraordinarily polarized pump photons can create signal and idler photons, which are polarized extraordinarily and ordinarily, respectively, and are emitted along different emission cones. Detection, or coupling into single mode fibers, should occur along the intersection lines of the two cones (e.g., intersection of circles).

(B) Collinear type II phase matching conditions as in Fig. 2.11a, where the down conversion crystal is cut such that the emission cones are tangential. In order to observe spatially indistinguishable photons, the collection should take place collinearly with proper filtering of the pump light.

FIGURE 2.11: Down conversion with type II phase matching.

the extraordinarily polarized pump photon can spontaneously create a pair of photons, out of which one is ordinarily, one extraordinarily polarized [121]. For example, with properly oriented crystal, the process  $|H\rangle \rightarrow |HV\rangle$  is possible. The produced signal and idler photons are then emitted along the surfaces of cones as shown in Fig. 2.11a. In order to achieve spatial indistinguishability between both photons, the detection or coupling into fibers, respectively, should occur at the intersections of both crystals' cones. Ideally, one can then observe photons in a state of the form  $1/\sqrt{2}(|HV\rangle + |VH\rangle)$ . By changing the orientation of the crystal axes, one can alternatively use a collinear scheme, where the cone intersections become degenerate, see Fig. 2.11b.

Down version can also be used for more creating than two photons, either by multiple passes through the crystal, see, e.g., [81], or by higher orders of emission [123, 124]. Recently, down conversion has allowed for creation of up to eight [125] and even ten [126] photons. For optimizing the phase matching conditions, periodic pooling of the down version crystal can be exploited, see, e.g., [30, 31, 127]. An experimental demonstration of an integrated pair source is given by [128]. A recent review of spontaneous parametric down conversion can be found in [121].

# Chapter 3

## Quantum State Tomography

### Contents

---

<b>3.1</b>	<b>Introduction</b>	<b>48</b>
<b>3.2</b>	<b>Methods of Quantum State Tomography</b>	<b>48</b>
3.2.1	Formalism	48
3.2.2	Overcomplete Pauli Tomography	50
3.2.3	Informationally Complete Tomography	51
3.2.4	Symmetric Informationally Complete Tomography	52
3.2.5	Efficient Specialized Protocols	52
<b>3.3</b>	<b>Enforcing Physicality</b>	<b>55</b>
3.3.1	Understanding Unphysical Results	55
3.3.2	Constrained Optimization	56
3.3.3	Bias in Various Experiments	60
3.3.4	[P1] Systematic Errors in Current Qu. State Tomography Tools	61
<b>3.4</b>	<b>Multiqubit State Tomography with Finite Data</b>	<b>71</b>
3.4.1	Wigner Semicircle Distribution	71
3.4.2	Catalan Numbers	73
3.4.3	Illustration of Proof - Example of Fourth Moment	75
3.4.4	Degenerate Eigenvalues	79
3.4.5	Error Estimation	79
3.4.6	[P2] Multiqubit State Tomography with Finite Data	85
<b>3.5</b>	<b>Conclusion</b>	<b>102</b>

---

## 3.1 Introduction

Quantum state tomography (QST) together with quantum process tomography (QPT) is a powerful tool for certification and optimization of sources and implementations of quantum states and quantum information protocols, respectively [108]. State tomography allows to obtain the initially unknown or only vaguely known quantum state and deduce its density matrix. According to Quantum Mechanics, knowledge of the density matrix implies complete knowledge about all properties of the system and thus all other relevant quantities can be determined from it.

However, performing quantum state tomography can be a daunting task as the necessary measurement effort for usual tomography methods increases exponentially with the number of qubits of the considered system. Furthermore, the finite statistics in actual tomographic measurements leads with high probability to some deviations from the theoretical state which can result in some estimate for the density matrix which does not describe a physically meaningful quantum state. While usual tomography schemes are constructed to incorporate that a density matrix is supposed to be a normalized and Hermitian matrix, direct state estimates often violate the requirement of positive semidefiniteness, see Sec. 2.2.1.3.

This chapter briefly introduces different methods for performing quantum state tomography, which can thus also be readily applied for quantum process tomography. Afterwards, the problem of finite statistics is discussed alongside the usual means to cope with it, leading over to the publication [P1], “Systematic Errors in Current Quantum Tomography Tools”. After a subsequent, more detailed discussion of the effects of finite statistics on the experimentally obtained data and the resulting spectrum of density matrices, the publication [P2], “Multiqubit State Tomography with Finite Data” introduces a wholly new approach to infer a physically meaningful estimate of the density matrix.

## 3.2 Methods of Quantum State Tomography

Quantum state tomography aims at estimating an unknown quantum state. For the actual implementation and realization various means do exist, which define the mapping between the measured set of data and the estimated density matrix. In this section, a brief introduction to the formalism of quantum state estimation together with an overview of some commonly used tomography schemes is given. Besides full quantum state tomography with exponentially increasing measurement effort, known symmetries of the states can be employed to significantly facilitate the tomography process.

### 3.2.1 Formalism

An important criterion to distinguish different state tomography methods is the amount of performed measurements compared to the amount of parameters. In order to reveal the density matrix of an  $n$ -qubit state, one has to find a set of  $4^n - 1$  real parame-

ters, in the Pauli basis denoted as correlation tensor elements or, for short, correlations, see Sec. 2.2.1.3. Thus, at least  $4^n$  measurements have to be performed to obtain a tomographically complete set of data<sup>1</sup>. Typically, a (*tomographically* or *informationally*) *complete* tomography scheme thus is applied to use exactly  $4^n$  measurements. Any scheme with more measurements is denoted as (*tomographically* or *informationally*) *overcomplete*. Here, a short review of the reconstruction is given. For more details, see, e.g., the supplemental material of Ref. [P1]. A summary of different methods of state reconstruction for the case of a single qubit can also be found in [129].

Consider a given, but unknown quantum state  $\varrho$ , which is to be estimated. For this purpose, a set of measurement operators  $M = \{M_\nu\}$  is chosen, such that a probability  $P_\nu$  can be associated to each measurement operator with

$$P_\nu = \text{tr}(\varrho M_\nu) = \sum_{\mu_1, \dots, \mu_n} T_{\mu_1, \dots, \mu_n} \underbrace{\frac{1}{2^n} \text{tr}(\sigma_{\mu_1} \otimes \dots \otimes \sigma_{\mu_n} M_\nu)}_{B_{\nu; \mu_1, \dots, \mu_n}}. \quad (3.1)$$

The matrix  $B_{\nu; \mu_1, \dots, \mu_n}$  is determined by the relation of the measurement operators  $M$  to the Pauli basis and gives the mapping between the correlation tensor element  $T_{\mu_1, \dots, \mu_n}$  and the probability  $P_\nu$ . The density matrix can be written in terms of probabilities

$$\varrho = \frac{1}{2^n} \sum_{\mu_1, \dots, \mu_n} \underbrace{\sum_{\nu} (B_{\nu; \mu_1, \dots, \mu_n})^{-1} P_\nu}_{T_{\mu_1, \dots, \mu_n}} \sigma_{\mu_1} \otimes \dots \otimes \sigma_{\mu_n}, \quad (3.2)$$

where  $(B_{\nu; \mu_1, \dots, \mu_n})^{-1}$  is the matrix inverse of the reconstruction matrix  $B$ . If the tomography scheme is informationally complete, i.e., the set of measurement operators contains exactly  $4^n$  elements, the matrix  $B$  is quadratic and can directly be inverted. However, for overcomplete schemes, the system of equations is overdetermined. A canonical way to solve for the  $4^n$  unknown correlations is then given by using the pseudo inverse  $(B^\dagger B)^{-1} B^\dagger$ , which, for quadratic matrices, is equivalent to the matrix inverse.

Consequently, each tomography scheme with a fixed set of measurement operators  $M$  is determined by a fixed reconstruction matrix  $B$  (and its inverse). The density matrix can then be obtained using the reconstruction matrix and the set of experimentally observed frequencies  $f = \{f_\nu\}$ , which are drawn according to the set of probabilities  $\{P_\nu\}$ . The estimate  $\tilde{\varrho}$  for the state  $\varrho$  is given by

$$\tilde{\varrho} = \frac{1}{2^n} \sum_{\mu_1, \dots, \mu_n} \underbrace{\sum_{\nu} (B_{\nu; \mu_1, \dots, \mu_n})^{-1} f_\nu}_{\tilde{T}_{\mu_1, \dots, \mu_n}} \sigma_{\mu_1} \otimes \dots \otimes \sigma_{\mu_n} \quad (3.3)$$

<sup>1</sup>Albeit from the  $4^n$  parameters of a general decomposition of a Hermitian matrix into tensor products of Pauli matrices, one parameter can be fixed due to normalization,  $T_{00\dots 0} \equiv 1$ , experimentally, one still has to carry out  $4^n$  measurements to normalize the experimentally obtained values. This measurement can only be skipped if one has some means to normalize the obtained counts to relative frequencies, e.g., knowledge of the production rate of quantum states.

with  $\tilde{T}_{\mu_1, \dots, \mu_n}$  being the estimate for the correlation  $T_{\mu_1, \dots, \mu_n}$ .

Different tomography schemes can lead to better or worse conditioned reconstruction matrices, which can be beneficial or unfavorable in terms of errors of the estimated state [130]. If, however, the set of measurements is not tomographically complete or overcomplete, the reconstruction matrix  $B$  cannot be inverted and  $\rho$  cannot be estimated from  $f$  unless further assumptions regarding, e.g., the symmetry of the state are made.

### 3.2.2 Overcomplete Pauli Tomography

For optical experiments with qubits encoded in the polarization of the photons, the projective measurement onto polarization states can easily be implemented using polarizing beam splitters (PBS) as described in Sec. 2.3.1.3. The detection of a photon in the transmitted or reflected output of a PBS then corresponds to a successful projection onto the respective eigenstate of polarizing beam splitters, see left part of Fig. 3.1, and therefore allow to project onto  $|P\rangle$  and  $|M\rangle$ , onto  $|R\rangle$  and  $|L\rangle$ , and onto  $|H\rangle$  and  $|V\rangle$ . For the idealized case of perfect detectors, every incoming photon will be registered as an event. With a beam splitter acting on each of  $n$  qubits,  $2^n$  projective measurements are performed simultaneously, with one of them eventually being realized by detection of the  $n$ -photons in a particular combination of the outputs. To obtain the density matrix, it is necessary and sufficient to determine measurement results in a full set of complementary bases. Thus, to implement the respective measurement operator  $M$ , the polarization directions additionally have to be rotated to three different settings for each qubit [131, 132]. Usually, the bases are given by the  $3^n$  combinations of Pauli matrices, such that in the end  $2^n \cdot 3^n$  projective measurements onto the eigenstates of all combinations of Pauli matrices have been performed. Obviously, those measurements are informationally overcomplete, as a smaller set of measurements are theoretically sufficient to retrieve all necessary information.

In the single qubit case, one thus projects onto the six states as given in Eqs. (2.8) [133]. These states are depicted in Fig. 3.2, which compares different tomography methods.

The reconstruction matrix  $B$  and its pseudo inverse  $B^+ = (B^\dagger B)^{-1} B^\dagger$  for the single qubit case are given according to Sec. 3.2.1 by

$$B = \frac{1}{2} \begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & -1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & -1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & -1 \end{pmatrix}, \quad B^+ = \begin{pmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ 1 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & -1 \end{pmatrix}. \quad (3.4)$$

The first row of  $B^+$  indicates that all six projective measurements contribute equally to the normalization  $\langle \sigma_0 \rangle$ , while the second row corresponds to  $\langle \sigma_x \rangle = f_1 - f_2$ , where  $f_1$  and  $f_2$  are the relative frequencies of successful projections onto the eigenstates of  $\sigma_x$ . For  $n$  qubits, the reconstruction matrix  $B$  will be an element from  $\mathbb{R}^{6^n \times 4^n}$ . Consequently,  $B^+ \in \mathbb{R}^{4^n \times 6^n}$  connects the set of  $6^n$  measured frequencies with the  $4^n$  correlation values. This leads to

the relation

$$\varrho = \frac{1}{2^n} \sum_{s_1, \dots, s_n=1}^3 \sum_{r_1, \dots, r_n=0}^1 f_{\nu(s_1, \dots, s_n; r_1, \dots, r_n)} \bigotimes_{i=1}^n \left( \frac{1}{3} \sigma_0 + (-1)^{r_i} \sigma_{s_i} \right), \quad (3.5)$$

where  $s_j$  denotes the measurement setting on the  $j$ -th qubit and  $r_j$  the outcome (0 or 1) on the respective qubit. For agreement with the notation of the set of measured frequencies  $\{f_\nu\}$ , the frequencies  $f_{\nu(s_1, \dots, s_n; r_1, \dots, r_n)}$  in Eq. (3.5) have to be indexed ascendingly. A different, but equivalent explanation how to obtain a density matrix from the measured data can be found in [131].

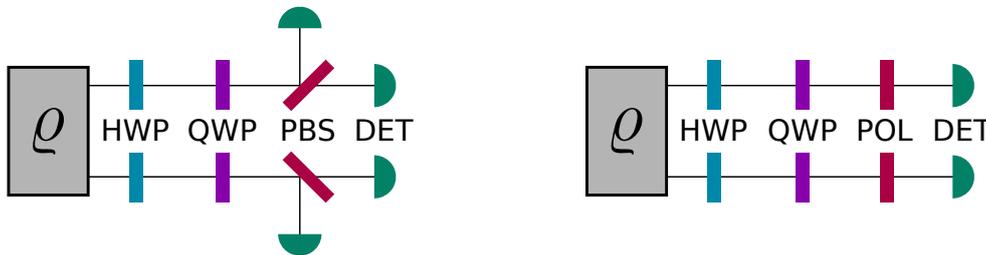


FIGURE 3.1: Typical implementation of a two qubit overcomplete Pauli tomography scheme (left) and a scheme used for informationally complete tomography (right) for qubits encoded in the polarization degree of freedom. A combination of half and quarter wave plates, HWP and QWP, respectively, see also Sec. 2.3.1.1, is used together with polarizing beam splitters (PBS) and polarizers (POL), respectively, to perform projective measurements onto various polarization states. By using polarizing beam splitters, the projection onto pairs of orthogonal states is done simultaneously, allowing a (in principle) lossless measurement. For complete tomography (right), typically, the minimal necessary amount of projective measurements is carried out.

### 3.2.3 Informationally Complete Tomography

Contrary to the overcomplete Pauli scheme, projective measurements onto exactly  $4^n$  mutually different states are done in the information complete tomography as proposed in [134]. An informationally complete set of easily accessible local states is chosen, as for example given by  $|H\rangle$ ,  $|V\rangle$ ,  $|P\rangle$ , and  $|R\rangle$ . By projecting onto the first two of these, one can obtain information about the total count rate of produced states, which is needed to normalize measured counts to relative frequencies. Together with the latter measurements, the full information about the quantum state can be retrieved. For  $n$  qubits, the observers agree to project onto all  $4^n$  combinations of those states. As the states are distributed asymmetrically, the reconstruction matrix is badly conditioned and measurement errors can propagate to the estimated state unfavorably [130].

Contrary to the above case of overcomplete tomography, the reconstruction matrix  $B$

in the informationally complete case is quadratic and reads

$$B = \frac{1}{2} \begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & -1 \end{pmatrix}, \quad B^+ = \begin{pmatrix} 0 & 0 & 1 & 1 \\ 2 & 0 & -1 & -1 \\ 0 & 2 & -1 & -1 \\ 0 & 0 & 1 & -1 \end{pmatrix}. \quad (3.6)$$

Obviously, only four projection measurements contribute to  $B$ . According to  $B^+$ , for the reconstruction of  $\langle\sigma_0\rangle$  as well as for  $\langle\sigma_z\rangle$ , only the last two projective measurements ( $|0\rangle\langle 0|$  and  $|1\rangle\langle 1|$  or  $|H\rangle\langle H|$  and  $|V\rangle\langle V|$  when considering photon polarization) are needed. For  $\langle\sigma_x\rangle$  ( $\langle\sigma_y\rangle$ ), however, the first (second) measurement together with the last two measurements contribute. Thus, some projectors are treated differently than others. The states onto which projective measurements are conducted are illustrated in Fig. 3.2.

### 3.2.4 Symmetric Informationally Complete Tomography

To avoid the large asymmetry of the orientation of measurements in the informationally complete scheme, one can distribute the  $4^n$  states more symmetrically to obtain results less prone to errors. In the single qubit case, the four states can be chosen to be [135, 136]

$$|H\rangle, \quad (3.7)$$

$$\cos(2\pi/3)|H\rangle + \sin(2\pi/3)|V\rangle, \quad (3.8)$$

$$\cos(2\pi/3)|H\rangle + e^{i2\pi/3}\sin(2\pi/3)|V\rangle, \quad (3.9)$$

$$\cos(2\pi/3)|H\rangle + e^{i4\pi/3}\sin(2\pi/3)|V\rangle \quad (3.10)$$

as given by the vertices of the blue tetrahedron shown in Fig. 3.2. Yet, for more qubits, the distribution of measurements can become a difficult task and result in measurements in entangled bases. One can make the compromise of using locally measurements according to the tetrahedron, however, accepting that the distribution is in general not symmetric for larger numbers of qubits.

### 3.2.5 Efficient Specialized Protocols

While full state tomography as discussed above can reveal a completely unknown quantum state, the measurement effort can be daunting. If, however, one has some prior knowledge about the quantum state or assumes some symmetries, more efficient protocols can be used. Here, a brief discussion of *permutationally invariant quantum state tomography* or, for short, *PI tomography* for states invariant under exchange of observers, the method of *compressed* (sometimes *compressive*) *sensing* for states with low rank as well as *matrix product state tomography* for states from short-ranged interactions is given.

Of course, combinations of those specialized methods are possible, as shown in [137], where permutationally invariant tomography and compressed sensing had been used together. Also, those methods may be connected with *adaptive tomography*, resulting in an

adaptive method working for example in the PI subspace or only with matrix product states.

### 3.2.5.1 Permutationally Invariant State Tomography

Many prominent quantum states are permutationally invariant. This includes, for example, the  $n$ -qubit Greenberger-Horne-Zeilinger state, see Sec. 2.2.4.3, which is proportional to  $|0\rangle_A \otimes |0\rangle_B \otimes \cdots \otimes |0\rangle_Z + |1\rangle_A \otimes |1\rangle_B \otimes \cdots \otimes |1\rangle_Z$ . Exchanging the particles of two observers, i.e., relabeling, e.g.,  $A \leftrightarrow B$  leaves the state unchanged. Since this property holds for all parties, the GHZ state is invariant under permutation of all parties. Dicke states share the same property.

Under the assumption that a state is permutationally invariant (PI), the measurement of, e.g.,  $\sigma_x \otimes \sigma_y$  gives the same result as  $\sigma_y \otimes \sigma_x$ . Using this symmetry (and usually distributing the remaining measurements in a more elaborate way) reduces the exponentially large Hilbert space to a polynomially large subspace. Consequently, the exponentially scaling measurement effort for full state tomography can be reduced to polynomial scaling [132, 137–139]. It is also important to stress that the overlap of the measured state with the permutationally symmetric subspace can be estimated in advance, justifying the application of this scheme.

### 3.2.5.2 Compressed Sensing

For quantum states of low rank  $r \ll d^2$ , i.e.,  $d^2 - r$  vanishing eigenvalues, where  $d$  denotes the number of dimensions, which for  $n$ -qubit systems is given by  $d = 2^n$ , the measurement effort can be largely reduced by a technique called *compressed sensing* [132, 137, 140–143]. It can be shown that instead of  $d^2$  measurements, only  $\mathcal{O}(rd \log^2 d)$  measurements are needed [141].

Out of all  $4^n$  tensor products of Pauli matrices including the identity matrix,  $\{\sigma_0 \equiv \mathbb{1}, \sigma_x, \sigma_y, \sigma_z\}$ ,  $m$  are randomly selected, which are called  $\{M_1, \dots, M_m\}$ , and used to perform measurements in the respective bases. From those  $m$  expectation values  $\text{tr}(\rho M_j)$ , the state  $\chi$  is reconstructed by the constrained minimization [141]

$$\chi = \arg \min_{\substack{\chi' \\ \text{tr}(\chi')=1 \\ \text{tr}(M_j \chi')=\text{tr}(M_j \rho)}} \sum_i |\lambda_i(\chi')|, \quad (3.11)$$

where  $\{\lambda_i(\chi')\}$  are the eigenvalues of  $\chi'$ . Eq. (3.11) is a convex problem [141], which simplifies finding a solution.

For implementing compressed sensing protocols, a state reconstruction algorithm with constrained optimization enforcing physical estimates as discussed in Sec. 3.3.2 is sufficient, see [144].

Experimental implementations of compressed sensing demonstrate its benefits for fast state reconstruction of rank-deficient states [137, 142, 143]. However, further research regarding its error performance seems valuable.

### 3.2.5.3 Matrix Product State Tomography

Quantum states of systems with rather short-range interactions can be parametrized in a very efficient way. Linear chains can be characterized by means of *matrix product states* (MPS), while different configurations require a modified approach as given by, for example, *projected entangled pair states* (PEPS) [145–147]. Matrix-product states can reduce the measurement effort significantly, depending on the range of considered correlations. Instead of performing measurements on all combinations of, e.g., Pauli operators, only neighboring subsets of particles are taken into account [148]. For example, instead of performing measurements in  $3^n$  bases, only  $(n-2) \times 3^3$  measurements have to be performed if the state to be detected is assumed to originate from an initial product state and only next-neighbor interactions. Then, it is sufficient to measure only subsets of neighboring 3 particles in all combinations of  $3^3$  Pauli matrices. Although in principle the formalism of matrix product states is capable to describe any state, it is most beneficial for states with short-range interactions [148].

### 3.2.5.4 Adaptive State Tomography

Finally, the group of adaptive quantum state and process tomography methods is to be mentioned. There, one tries to incorporate previously gained knowledge into the choice of subsequent measurements [149–157]. For example, if one fixes the total measurement time (or events) beforehand, one can perform the first half of the measurements in some previously chosen set of bases. The latter half of the measurements is then aligned according to the eigenbasis of the prior state estimate [152].

Alternatively, a small, *informationally incomplete* set of measurement bases can be chosen and used for intermediate state estimation based on some reconstruction as given by the maximum likelihood procedure, see Sec. 2.2.6.6 as well as Sec. 3.3.2. The next measurement setting is then chosen depending on the intermediate state estimate [154]. This procedure is per definition informationally incomplete, but seems to converge reasonably. A variation of this procedure could make use of the complementarity of correlations such that depending on previously performed measurements, the potential information gain for each subsequent measurement is assessed. Then, the following measurements are chosen based on the information potentially gained from them.

Closely related to the concept of adaptive tomography, the idea of self-guided tomography has emerged [158, 159], which avoids expensive post-processing and therefore reduce the requirements of data storage. Instead of collecting a large set of measurement results used for state reconstruction, self-guided tomography iteratively converges to the true state during runtime [158].

Another subclass of adaptive schemes makes use of techniques from machine learning [160] such as supervised or reinforcement learning. In the latter approach, an *agent* is provided some feedback depending on the overlap of the adjustable estimated state and the unknown state. By maximizing the rewards, the adjustable state is becoming an estimate for the unknown state [161, 162].

### 3.3 Enforcing Physicality

Although above methods allow to obtain an estimate for the prepared quantum state, in a practical scenario, the limitation of the measurement time and thus of the sample size leads to statistical deviations even for an otherwise ideal experimental setting. Therefore, the estimated matrix may in general not be in agreement with the requirements for a density matrix. Here, a result which does not describe a physical state ( $\varrho \not\geq 0$ ) is called *unphysical*. In the following sections, the reason for unphysical estimates will be illustrated together with an introduction into how this problem is often handled.

#### 3.3.1 Understanding Unphysical Results

TABLE 3.1: Example illustrating how quantum state tomography of the ideal state  $|0\rangle$  with finite statistics can lead to unphysical results. The colors illustrate the contributions of the correlations to the estimated density matrix and the actual density matrix, respectively.

	ideal state $\varrho^{\text{theo}}$		measured state $\varrho^{\text{exp}}$	
$\sigma_i$	probabilities	$\langle\sigma_i\rangle$	events	$\langle\sigma_i\rangle$
$\sigma_0$	(constant)	1	(constant)	1
$\sigma_1$	$p_1^+ = p_1^- = \frac{1}{2}$	0	$c_1^+ = 6, c_1^- = 4$	$\frac{6-4}{6+4} = 0.2$
$\sigma_2$	$p_2^+ = p_2^- = \frac{1}{2}$	0	$c_2^+ = 4, c_2^- = 6$	$\frac{4-6}{4+6} = -0.2$
$\sigma_3$	$p_3^+ = 1, p_3^- = 0$	1	$c_3^+ = 10, c_3^- = 0$	1
$\varrho$	$\frac{1}{2} \begin{pmatrix} 1+1 & 0 \\ 0 & 1-1 \end{pmatrix}$		$\frac{1}{2} \begin{pmatrix} 1+1 & 0.2 - (-0.2)i \\ 0.2 + (-0.2)i & 1-1 \end{pmatrix}$	
Eigenvalues	$\{0, 1\}$		$\approx \{-0.02, 1.02\}$	

In order to understand how finite statistics can lead to estimates for the density matrix which are not compatible with quantum theory, consider the following straightforward one-qubit example. The description follows the scheme of overcomplete Pauli tomography according to Sec. 3.2.2. The ideal state to be measured is  $\varrho^{\text{theo}} = |0\rangle\langle 0|$ , and thus eigenstate of  $\sigma_z$ . When performing full state tomography, the three correlation tensor elements, i.e., the Bloch vector elements, are to be determined, i.e.,  $T_1 = \langle\sigma_1\rangle$ ,  $T_2 = \langle\sigma_2\rangle$ , and  $T_3 = \langle\sigma_3\rangle$ . Assume that the measurement statistics are such that 10 events per measurement setting will be recorded. According to Sec. 2.2.1.3, a matrix describing a physical quantum state has to be normalized, Hermitian, and positive semi-definite. While the first two conditions are already ensured by construction using correlations as coefficients in a Pauli decomposition, finite statistics effects can lead to eigenvalues outside of the allowed regime  $[0, 1]$  as in the example shown in Tab. 3.1. The theoretically pure state is estimated by a matrix with eigenvalues approximately equal  $\{-0.02, 1.02\}$ , which does not describe a physical state.

This single qubit example of a tomography of  $|0\rangle$  clearly shows how small statistical fluctuations can lead to unphysical results. In fact, the eigenstates of the measurement

operators are stereotypical examples for this effect. Consider for example the situation of an odd number of measurement events instead of the above chosen set of 10 measurements per setting. For 11 measurements per setting,  $\langle\sigma_z\rangle$  is still found to be 1, as the probabilities for the corresponding projectors are given as before by  $p_3^+ = 1$ ,  $p_3^- = 0$ . However, it is now impossible to obtain, say,  $c_1^+ = c_1^-$ , leading to  $\langle\sigma_x\rangle \neq 0$ . Thus,  $\langle\sigma_x\rangle^2 + \langle\sigma_y\rangle^2 + \langle\sigma_z\rangle^2 > 1$ , which does not describe a physical state, as the Bloch vector is now outside the Bloch sphere. Fig. 3.3 illustrates a cross section through the Bloch sphere. For a fixed set of measurement settings, the probability to obtain an unphysical state after tomographic reconstruction thus depends on the initial state. For theoretically pure states with one qubit, it is largest if the initial state is aligned with one of the measurement axes, i.e., an eigenstate of one of the measurement operators [163]. The probability reduces to about 1/2 for maximal misalignment between measurement axes and initial state. A related discussion can be found in [164].

### 3.3.2 Constrained Optimization

As shown above, the linear inversion of tomographic data can (and for many states will with high probability) result in unphysical estimates. A common procedure to counteract this behavior is to enforce physicality of the estimated state as a constraint during state reconstruction. The general approach is to find the *physical* quantum state, which is in best agreement with the measured data. The meaning of *best agreement* is then still to be defined and depends on the chosen method.

The linear inversion method for overcomplete Pauli tomography with the reconstruction matrix  $B$ , see Eq. (3.4), provides a useful starting point for the discussion. As this method is *overcomplete*, due to statistical fluctuations the redundant information obtained in  $6^n$  measurements can be slightly contradictory when estimating the  $4^n$  real parameters (namely the correlations) of the density matrix. In order to obtain a result, those measurements can be assembled such that the reconstructed state is in best agreement with all different measurements. This canonical procedure, incorporated into the pseudo inverse of  $B$ ,  $B^+$ , is equivalent to the least squares fit of the density matrix, where the estimate should be a compromise between the redundant measurements. Yet, this approach does not contain additional constraints such as a constraint enforcing physical results. Instead, for constrained optimization, for the *loss* (or *cost*) function one either chooses a least squares term or, more commonly, the likelihood for a set of parameters given a measured set of data. According to the definition of the likelihood, see Sec. 2.2.6.5, and the method of maximum likelihood estimation, see Sec. 2.2.6.6, the set of  $k$  parameters (probabilities, correlations, etc.) is sought, which maximizes the likelihood to have produced a set of  $k$  measured values (frequencies, estimated correlations, etc.).

For a given set of experimentally observed events  $\{n_i\}$ , the negative log-likelihood, see Eq. (2.90), can be used as a loss function to find a set of probabilities  $\{p_i\}$ , which might have produced the observed events. The negative log-likelihood reads [165]

$$\tilde{\mathcal{L}}(\{p_i\}|\{n_i\}) \equiv - \sum_i n_i \log p_i. \quad (3.12)$$

The set of probabilities is then given by

$$\{p_i\} = \arg \min_{\{\tilde{p}_i\}} \check{\mathcal{L}}(\{\tilde{p}_i\}|\{n_i\}). \quad (3.13)$$

By changing the definition of the log-likelihood function, one obtains

$$\varrho_{\text{MLE}}^{\text{uncon}} = \arg \min_{\tilde{\varrho}} \check{\mathcal{L}}(\tilde{\varrho}|\{n_i\}) \quad (3.14)$$

for the unconstrained maximum likelihood estimate. In order to incorporate the physicality constraint, the optimization has to be restricted to the set of positive semidefinite (density) matrices,  $\tilde{\varrho} \in \{\varrho|\varrho \geq 0\}$ . Hence, the (constrained) maximum likelihood estimate  $\varrho_{\text{MLE}}$  is found by means of

$$\varrho_{\text{MLE}} = \arg \min_{\tilde{\varrho} \geq 0} \check{\mathcal{L}}(\tilde{\varrho}|\{n_i\}). \quad (3.15)$$

As an alternative to the negative log-likelihood as a loss function, one can use the least squares between the measured frequencies and the expected probabilities,

$$\varrho_{\text{LS}} = \arg \min_{\tilde{\varrho} \geq 0} \sum_i w_i (f_i - p_i)^2 \quad (3.16)$$

with adequate weights such as  $w_i = 1$  or  $w_i = p_i$  [134]. Those loss functions can further be modified to, e.g., accommodate for *boundary effects* such that to a specific event which was never observed should be attributed a small, but non-vanishing probability as by the so-called *hedged maximum likelihood estimation* [166].

Finding the optimum of the constrained optimization in Eq. (3.15) is generally a numerically demanding task [108, 134]. While first approaches could not verify the successful optimization, a series of powerful reconstruction methods emerged [108, 134, 165], see also [P1], partly making use of a convex formulation, which allows to guarantee convergence [132, 154]. Eventually, this lead recently to efficient state estimation tools [167, 168].

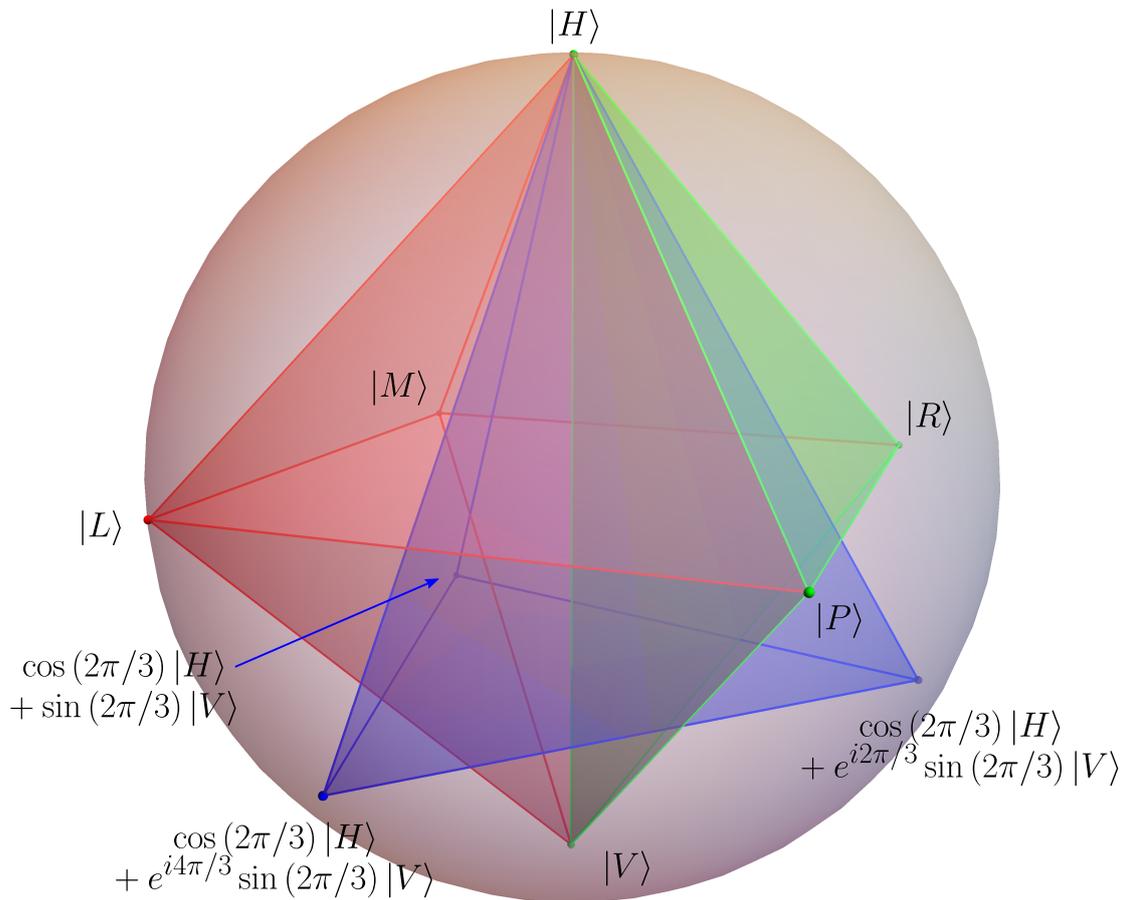


FIGURE 3.2: Different commonly used tomography methods for one qubit states. The over-complete Pauli scheme is based upon projection onto the states  $|H\rangle$ ,  $|V\rangle$ ,  $|P\rangle$ ,  $|M\rangle$ ,  $|R\rangle$ , and  $|L\rangle$ , as shown by the edges of the red octahedron. The informationally complete (green) scheme makes use of the states  $|H\rangle$ ,  $|V\rangle$ ,  $|P\rangle$ , and  $|R\rangle$ , which allow to reconstruct the density matrix. However, those states are distributed in a suboptimal way. In contrast, the symmetric, tomographically complete scheme, shown by the blue tetrahedron, distributes the 4 measurements in a more uniform way on the Bloch sphere such that the measurements have equal mutual overlap.

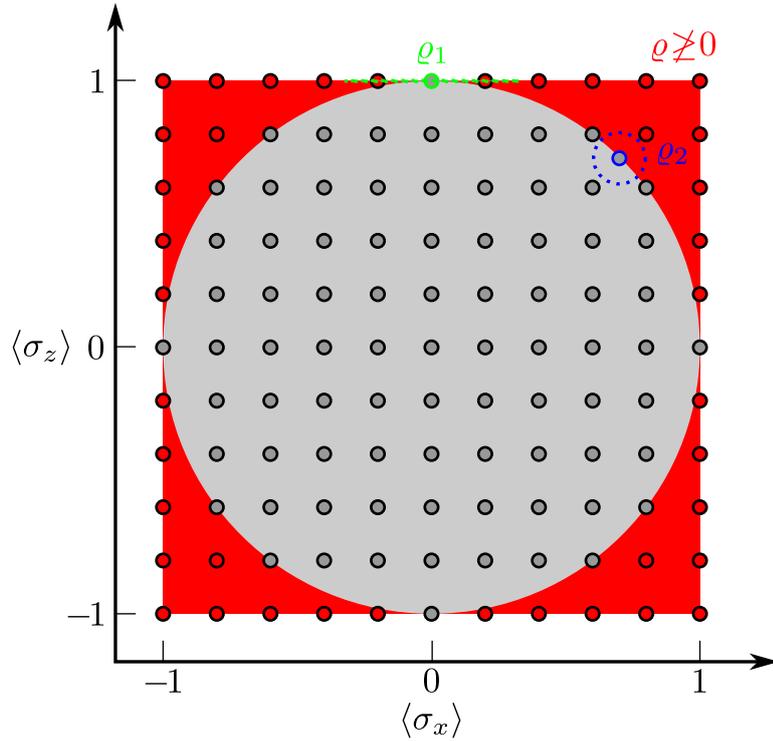


FIGURE 3.3: All *rebits* (real qubits) lie within the cross section with  $\langle \sigma_y \rangle = 0$  on the Bloch sphere, cf. Fig. 2.1. The quantum state reconstruction based on the linear inversion as described in Eq. (3.5) leads for only  $N = 10$  counts per basis setting to a finite set of  $11^3$  different statistically obtainable outcomes and thus can resolve only  $11^3$  different quantum states. For the example of  $\rho_1 = |0\rangle\langle 0|$  (green dot) as discussed in Tab. 3.1, the outcomes will have  $\langle \sigma_z \rangle = 1$ . The green dotted line indicates the uncertainty of the possible results, indicating that the outcome will be unphysical with high probability. The rotated state  $\rho_2$  (blue dot), however, has approximately 50% probability to result in a physical state as the uncertainty circle (blue dotted) has similar overlap with the physical region ( $\rho \geq 0$ ) as with the unphysical one ( $\rho \not\geq 0$ ).

### 3.3.3 Bias in Various Experiments

Although constraining the estimation to the subspace of physical results inherently avoids unphysical density matrices, it can lead to severe side effects. In [P1], which is reprinted in the following section, the systematic errors in state tomography due to constrained optimization are studied. When considering tomography of an unknown state with finite statistics, the reconstructed state may show systematically wrong values for the state fidelity or entanglement monotones. This effect, often called *estimation bias*, can hence lead to spurious entanglement detection of a state initially fully factorizable [P1]. The bias strongly depends on the initial quantum state, the number of qubits, the measurement settings, and the measurement statistics per setting. Although for increasing sample size, i.e., longer measurement time, the estimator is consistent and hence converges to the true result in the limit of infinite sample size, one has to carefully evaluate the measurement data for typical samples sizes in photonic experiments. For systems with a single qubit, similar results to [P1] had been reported in [150]. The systematics discussed in [P1] have also been seen among others in [169–171].

In [P1], state-of-the-art tomography tools are discussed. For finite statistics with sample size typical for multiqubit quantum state tomography, the systematic deviations between some theoretical states and results of simulated state tomography are discussed. The influences of various properties of the theoretical state, such as the number of qubits, its purity and the entanglement class, are investigated. Finally, a method is presented how meaningful one-sided error bars can be obtained for various quantities of interest.

This publication originated from a close collaboration with Tobias Moroder, Matthias Kleinmann and Otfried Gühne. It is based both on theoretical work, which was led by Otfried Gühne, and on numerical studies, which were mainly conducted by us. For the numerical studies, we developed an efficient computer program for constrained optimization of tomographic results as described above. Together with Christian Schwemmer, and also Daniel Richart, I worked on this program, significantly increased its efficiency and simulated the required sets of tomographic data. I implemented several of the data evaluation procedures and, together with Christian Schwemmer, evaluated the data. For five and six qubits, I simulated and evaluated the data sets on a computation cluster. All authors contributed to the preparation and editing of the manuscript.

## Systematic Errors in Current Quantum State Tomography Tools

Christian Schwemmer,<sup>1</sup> Lukas Knips,<sup>1</sup> Daniel Richart,<sup>1</sup> Harald Weinfurter,<sup>1</sup> Tobias Moroder,<sup>2</sup>  
Matthias Kleinmann,<sup>2,3,4</sup> and Otfried Gühne<sup>2</sup><sup>1</sup>Max-Planck-Institut für Quantenoptik, Hans-Kopfermann-Straße 1, D-85748 Garching, Germany and Department für Physik,  
Ludwig-Maximilians-Universität, D-80797 München, Germany<sup>2</sup>Naturwissenschaftlich-Technische Fakultät, Universität Siegen, Walter-Flex-Straße 3, D-57068 Siegen, Germany<sup>3</sup>Departamento de Matemática, Universidade Federal de Minas Gerais, Caixa Postal 702, Belo Horizonte,  
Minas Gerais 31270-901, Brazil<sup>4</sup>Department of Theoretical Physics, University of the Basque Country UPV/EHU, P.O. Box 644, E-48080 Bilbao, Spain

(Received 31 October 2013; published 26 February 2015)

Common tools for obtaining physical density matrices in experimental quantum state tomography are shown here to cause systematic errors. For example, using maximum likelihood or least squares optimization to obtain physical estimates for the quantum state, we observe a systematic underestimation of the fidelity and an overestimation of entanglement. Such strongly biased estimates can be avoided using linear evaluation of the data or by linearizing measurement operators yielding reliable and computational simple error bounds.

DOI: 10.1103/PhysRevLett.114.080403

PACS numbers: 03.65.Wj, 03.65.Ud, 06.20.Dk

**Introduction.**—Quantum state tomography (QST) [1] enables us to fully determine the state of a quantum system and, thereby, to deduce all its properties. As such, QST and the closely related quantum process tomography (QPT) are widely used to characterize and to evaluate numerous experimentally implemented qubit states or their dynamics, e.g., in ion trap experiments [2,3], photonic systems [4,5], superconducting circuits [6], or nuclear magnetic resonance systems [7,8]. The increasing complexity of today’s multiqubit or multiqutrit quantum systems brought new challenges but, also, progress. Now, highly efficient methods allow an even scalable analysis for important subclasses of states [9,10]. The calculation of errors of QST was significantly improved, although the errors remain numerically expensive to evaluate for larger systems [11]. Moreover, QST and QPT were adopted to detect systematic errors in the alignment of an experiment itself [12].

A central step in QST is to establish the state from the acquired experimental data. A direct, linear evaluation of the data returns, almost for sure, an unphysical density matrix with negative eigenvalues [13]. Thus, several schemes have been developed to obtain a physical state which resembles the observed data as closely as possible [4,14,15]. From classical statistics, it is known that a constraint, such as the physicality of a state, can lead to systematic deviations, called bias, in parameter estimation for finite statistics [16,17]. However, in quantum tomography experiments, this effect has hardly ever been considered.

In this Letter, we test whether the naive expectation is met that QST delivers meaningful estimates for physical quantities. We test this for the two most commonly used reconstruction schemes—maximum likelihood (ML) [15]

and least squares (LS) [4]—using Monte Carlo simulations. In detail, we investigate whether or not a possibly occurring bias of these reconstruction schemes is relevant at all on the example of some of the most prominent multiqubit quantum states. We find that, due to the constraint of physicality, both ML and LS return states which deviate systematically from the true state. Foremost, for small sample sizes, as they are typical in multiqubit experiments, it leads to significantly differing estimates for physical quantities as illustrated for the fidelity with respect to the Greenberger-Horne-Zeilinger (GHZ) state in Fig. 1 [18]. These deviations depend on the experimental and statistical noise but are typically larger than commonly deduced errors [19].

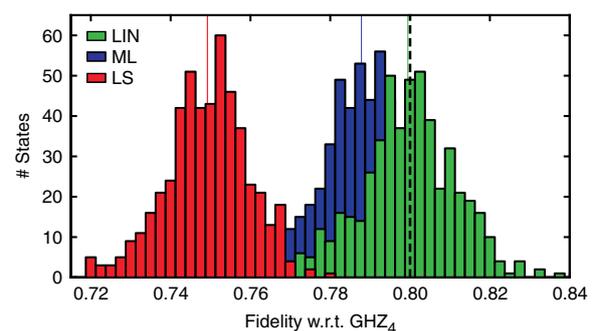


FIG. 1 (color online). Histogram of the fidelity estimates of 500 independent simulations of QST of a noisy four-party GHZ state for three different reconstruction schemes. The values obtained via ML (blue) or LS (red) fluctuate around a value that is lower than the initial fidelity of 80% (dashed line). For comparison, we also show the result using LIN (green), which does not have this systematic error called bias.

We demonstrate that, for linear quantities, one can provide meaningful confidence regions directly from the raw data and that it is also possible to compute meaningful lower (upper) bounds on convex (concave) functions like the bipartite negativity.

*Standard state tomography tools.*—The aim of QST is to find an estimate together with some confidence region for the initially unknown state  $\rho_0$  of a system via appropriate measurements on multiple preparations of this state. For an  $n$ -qubit system, the so-called Pauli tomography scheme consists of measuring in the eigenbases of all  $3^n$  possible combinations of local Pauli operators, each yielding  $2^n$  possible results [4]. In more general terms, in a tomography protocol, one repeats, for each measurement setting  $s$ , the experiment a certain number of times  $N_s$  and obtains  $c_r^s$  times the result  $r$ . These numbers then yield the frequencies  $f_r^s = c_r^s/N_s$ . The probability to observe the outcome  $r$  for setting  $s$  is given by  $P_{\rho_0}^s(r) = \text{tr}(\rho_0 M_r^s)$ . Here,  $M_r^s$  labels the measurement operator corresponding to the result  $r$  when measuring setting  $s$ . The probabilities  $P_{\rho_0}^s(r)$  will uniquely identify the unknown state  $\rho_0$ , if the set of operators  $M_r^s$  spans the space of traceless Hermitian operators.

Provided the data  $f$ , i.e., the experimentally determined frequencies  $f_r^s$ , one requires a method to determine the estimate  $\hat{\rho} \equiv \hat{\rho}(f)$  of the unknown state  $\rho_0$ . Simply inverting the relations for  $P_{\rho_0}^s(r)$  we obtain

$$\hat{\rho}_{\text{LIN}} = \sum_{r,s} A_r^s f_r^s, \quad (1)$$

where  $A_r^s$  are determined from the measurement operators  $M_r^s$  [8,20]. Note that there is a canonical construction of  $A_r^s$  even for the case of an overcomplete set of  $M_r^s$ , see SM 1 in the Supplemental Material (SM) [21]. This reconstruction of  $\hat{\rho}_{\text{LIN}}$  is computationally simple and has become known as linear inversion (LIN) [23]. In principle, Gaussian error propagation could also be used here to determine confidence regions.

Yet, due to unavoidable statistical fluctuations, the estimate  $\hat{\rho}_{\text{LIN}}$  is not a physical density operator for typical experimental situations; i.e., generally some eigenvalues are negative. Apart from causing issues related to a physical interpretation of such a “state”, negative eigenvalues impedes the evaluation of interesting functions like the von Neumann entropy, the quantum Fisher information, or an entanglement measure like the negativity, as these functions are defined, or meaningful, only for valid, i.e., positive semidefinite, quantum states.

For this reason, different methods have been introduced that mostly follow the paradigm that the reconstructed state  $\hat{\rho} = \arg \max_{\rho \geq 0} T(\rho|f)$  maximizes a target function  $T(\rho|f)$  within the set of valid density operators. This target function, thereby, measures how well a density operator  $\rho$  agrees with the observed data  $f$ . Two common choices are ML [15] where  $T_{\text{ML}} = \sum_{r,s} f_r^s \log[P_{\rho}^s(r)]$ , and LS [4] where  $T_{\text{LS}} = -\sum_{r,s} [f_r^s - P_{\rho}^s(r)]^2 / P_{\rho}^s(r)$ . We denote the

respective optima by  $\hat{\rho}_{\text{ML}}$  and  $\hat{\rho}_{\text{LS}}$ . From these estimates, one then easily computes any physical quantity of the observed state, e.g., the fidelities  $\hat{F}_{\text{ML}} = \langle \psi | \hat{\rho}_{\text{ML}} | \psi \rangle$  and  $\hat{F}_{\text{LS}} = \langle \psi | \hat{\rho}_{\text{LS}} | \psi \rangle$  with respect to the target state  $|\psi\rangle$ .

*Numerical simulations.*—To enable detailed analysis of the particular features of the respective state reconstruction algorithm and to exclude influence of systematic experimental errors, we perform Monte Carlo simulations. For a chosen state  $\rho_0$ , the following procedure is used: (i) Compute the single event probabilities  $P_{\rho_0}^s(r)$ , (ii) toss coins to get frequencies distributed according to the multinomial distribution determined by  $P_{\rho_0}^s(r)$  and  $N_s$ , (iii) reconstruct the state with either reconstruction method and compute the functions of interest, (iv) carry out steps (ii) and (iii) 500 times. Note that the optimality of the maximizations for ML and LS in step (iii) is certified by convex optimization [10,24].

First, we consider the four-qubit GHZ state  $|\text{GHZ}_4\rangle = (|0000\rangle + |1111\rangle)/\sqrt{2}$  mixed with white noise, i.e.,  $\rho_0 = p|\text{GHZ}_4\rangle\langle\text{GHZ}_4| + (1-p)\mathbb{1}/16$  where  $p$  is chosen such that the fidelity is  $\langle\text{GHZ}_4|\rho_0|\text{GHZ}_4\rangle = 0.8$ . This state is used to simulate the Pauli tomography scheme. Figure 1 shows a typical histogram of the resulting fidelities for  $N_s = 100$  measurement repetitions, which is a typical value used for various multiqubit experiments. The fidelities obtained via LIN reconstruction fluctuate around the initial value ( $\bar{F}_{\text{LIN}} = 0.799 \pm 0.012$ ). (The values given there are the mean and the standard deviation obtained from the 500 reconstructed states). In stark contrast, both ML ( $\bar{F}_{\text{ML}} = 0.788 \pm 0.010$ ) and even more LS ( $\bar{F}_{\text{LS}} = 0.749 \pm 0.010$ ) systematically underestimate the fidelity, i.e., are strongly biased. Evidently, the fidelities of the reconstructed states differ by more than 1 standard deviation for ML and even more than 5 standard deviations for LS. The question of how these systematic errors depend on the parameters of the simulation arises. Let us start by investigating the dependence on the number of repetitions  $N_s$ . Figure 2(a) shows the mean and the standard deviations of histograms like the one shown in Fig. 1 for different  $N_s$ . As expected, the systematic errors are more profound for low numbers of repetitions  $N_s$  per setting  $s$  and decrease with increasing  $N_s$ . Yet, even for  $N_s = 500$ , a number hardly used in multiqubit experiments,  $\bar{F}_{\text{LS}}$  still deviates by 1 standard deviation from the value for the initial state. The effect is also, by no means, special for the GHZ state but was equally observed for other prominent four-party states, here, also, chosen with a fidelity of 80%, see Figs. 2(b)–2(d) and the SM [21].

The systematic deviations also vary with the number of qubits or the purity of the initial state. Figure 3(a) shows the respective dependencies of the fidelity for  $n$ -qubit states  $\rho_0 = p|\text{GHZ}_n\rangle\langle\text{GHZ}_n| + (1-p)\mathbb{1}/2^n$  (for  $N_s = 100$ ). Here, a significant increase of the bias with the number of qubits is observed especially for LS. Also, when varying the purity or fidelity with the GHZ state, respectively,

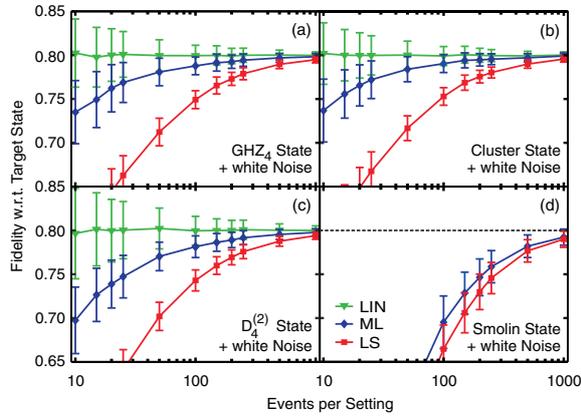


FIG. 2 (color online). The performance of ML, LS, and LIN methods depending on the number of events  $N_s$  per setting and for four different noisy initial states  $\rho_0$ . Note that the fidelity can only be calculated linearly if the reference state is pure which is not the case for the Smolin state [25]. Therefore, only the curves for ML and LS are plotted for the Smolin state.

we observe a large deviation for ML and LS estimators [Fig. 3(b)]. If the initial fidelity is very low, the effect is negligible, but large fidelity values suffer from stronger deviations, especially for LS.

The reliability of the estimates  $\hat{\rho}$  or of physical quantities deduced thereof are quantified by the size of confidence regions which commonly are deduced by bootstrapping methods [19]. Starting either from the estimate  $\hat{\rho}_{\text{EST}}$  ( $\text{EST} \in \{\text{ML}, \text{LS}\}$ ) or the observed data set  $f$ , this error is typically accessed by Monte Carlo sampling: One repeatedly simulates data  $f^{(i)}$  according to the state  $\rho_{\text{EST}}$  or  $f$  together with a reasonable noise model for the respective experiment and reconstructs the state  $\hat{\rho}(f^{(i)})$ . From the resulting empirical distribution, one then reports the standard deviation (or a region including, say, 68% of the simulated states) for the matrix elements or for quantities of interest [19], see also SM 3. However, the problem with such error regions is that they are typically too small since they reflect only statistical fluctuations of

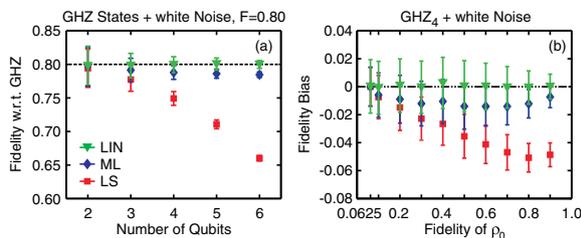


FIG. 3 (color online). The behavior of ML, LS, and LIN depending on the number of qubits  $n$  (left) and the fidelity of  $\rho_0$  (right).

the biased estimate, which can easily be smaller than the systematic error [26].

In summary, we observe systematic errors, which depend on the state reconstruction method and the strength of the statistical fluctuations of the count rates. Any manual correction of the bias and the statistical fluctuations [17] seems out of reach, since the effect depends on the unknown initial state which cannot be calculated from the observed data [16]. Let us emphasize that in most cases the initial value differs by more than the statistical error determined via bootstrapping (cf. SM 3 [21]).

*Biased and unbiased estimators.*—The systematic offset discussed above is well known in the theory of point estimates [26]. Expressed for QST, an estimator  $\hat{\rho}$  is called unbiased if its fluctuations are centered around the true mean, such that, for its expectation value,

$$\mathbb{E}_{\rho_0}(\hat{\rho}) \equiv \sum_f P_{\rho_0}(f) \hat{\rho}(f) = \rho_0 \quad (2)$$

holds for all possible states  $\rho_0$  with  $P_{\rho_0}(f)$  the probability to observe the data  $f$ . An estimator that violates Eq. (2) is called biased. Similar definitions hold, for instance, for fidelity estimators,  $\mathbb{E}_{\rho_0}(\hat{F}) = \langle \psi | \rho_0 | \psi \rangle \equiv F_0$ . This terminology is motivated by the form of the mean squared error, which decomposes, for example, for the fidelity into

$$\mathbb{E}_{\rho_0}[(\hat{F} - F_0)^2] = \mathbb{V}_{\rho_0}(\hat{F}) + [\mathbb{E}_{\rho_0}(\hat{F}) - F_0]^2, \quad (3)$$

where  $\mathbb{V}(\hat{F}) \equiv \mathbb{E}(\hat{F}^2) - \mathbb{E}(\hat{F})^2$  denotes the variance. Equation (3) consists of two conceptually different parts. The first is a statistical term quantifying the fluctuations of the estimator  $\hat{F}$  itself. The second, purely systematic term, is called bias and vanishes for unbiased estimators [27]. Note that, since the expectation values of the frequencies are the probabilities,  $\mathbb{E}_{\rho_0}(f_r^s) = P_{\rho_0}^s(r)$ , and because  $\hat{\rho}_{\text{LIN}}$  as given by Eq. (1) is linear in  $f_r^s$ , the determination of a quantum state using LIN is unbiased. However, as shown below, for QST, the bias is inherent to estimators constrained to giving only physical answers.

*Proposition.*—A reconstruction scheme for QST that always yields valid density operators is biased.

*Proof.*—For a tomography experiment on the state  $|\psi_i\rangle$  with finite measurement time, there is a set of possible data  $\mathcal{S}_i = \{f_i | P_{|\psi_i\rangle}(f_i) > 0\}$ , with  $P_{|\psi_i\rangle}(f_i)$  the probability to obtain data  $f_i$  when observing state  $|\psi_i\rangle$ .

Consider two pure nonorthogonal states  $|\psi_1\rangle$  and  $|\psi_2\rangle$  ( $\langle \psi_1 | \psi_2 \rangle \neq 0$ ). For these two states, there exists a nonempty set of data  $\mathcal{S}_{12} = \{f' | P_{|\psi_1\rangle}(f') \cdot P_{|\psi_2\rangle}(f') > 0\} = \mathcal{S}_1 \cap \mathcal{S}_2$ , which can occur for both states.

Now, let us assume that a reconstruction scheme  $\hat{\rho}$  provides a valid quantum state  $\hat{\rho}(f)$  for all possible outcomes  $f$  and that Eq. (2) is satisfied for  $|\psi_1\rangle$ , i.e.,  $\sum_{\mathcal{S}_1} P_{|\psi_1\rangle}(f_1) \hat{\rho}(f_1) = |\psi_1\rangle \langle \psi_1|$ . This incoherent sum over all  $\hat{\rho}(f_1)$  can be equal to the pure state  $|\psi_1\rangle \langle \psi_1|$  only for the

(already pathological) case that  $\hat{\rho}(f_1) = |\psi_1\rangle\langle\psi_1|$  for all  $f_1 \in \mathcal{S}_1$ . This means that the outcome of the reconstruction is fixed for all  $f_1$  including all data  $f' \in \mathcal{S}_{12}$ . As these data also occur for state  $|\psi_2\rangle$ , there exist  $f_2 \in \mathcal{S}_{12}$  with  $\hat{\rho}(f_2) = |\psi_1\rangle\langle\psi_1| \neq |\psi_2\rangle\langle\psi_2|$ . Thus, in Eq. (2), the sum over all reconstructed states now is an incoherent mixture of at least two pure states, and the condition  $\sum_{\mathcal{S}_2} P_{|\psi_2\rangle}(f_2) \hat{\rho}(f_2) = |\psi_2\rangle\langle\psi_2|$  is violated for  $|\psi_2\rangle$ . Hence,  $\hat{\rho}$  does not obey Eq. (2) for  $|\psi_2\rangle$  and is, therefore, biased [28].

This leaves us with the tradeoff: Should one necessarily use an algorithm like ML or LS to obtain a valid quantum state but suffer from a bias, or should one use LIN which is unbiased but typically delivers an unphysical result? In the following, we propose a scheme using linearized operators to provide a valid, lower or upper bound and an easily computable confidence region for many quantities of interest.

*Parameter estimation by linear evaluation.*—Many relevant functions are either convex, like most entanglement measures or the quantum Fisher information, or concave, like the von Neumann entropy. Thus, these operators can be linearized around some properly chosen state in order to obtain a reliable lower (upper) bound. Note that, typically, a lower bound on an entanglement measure is often suited for evaluating experimental states, whereas an upper bound does not give much additional information.

Recall that a differentiable function  $g(x)$  is convex if  $g(x) \geq g(x') + \nabla g(x')^T(x - x')$  holds for all  $x, x'$ . In our case, we are interested in a function  $g(x) = g[\rho(x)]$  where  $x$  is a variable to parametrize a quantum state  $\rho$  in a linear way. From convexity, it follows that it is possible to find an operator  $L$ , such that

$$\text{tr}(\rho_0 L) \leq g(\rho_0) \quad (4)$$

holds for all  $\rho_0$  (similarly an upper bound is obtained for concave functions). This operator  $L$  can be determined from the derivatives of  $g(x)$  with respect to  $x$  at a suitable point  $x'$ , from the Legendre transformation [29], or directly inferred from the definition of the function  $g(x)$  [30]. A detailed discussion is given in SM 5 [21].

For this bound, a confidence region, i.e., the error region for the frequentistic approach, can be calculated. For example, a one-sided confidence region of level  $\gamma$  can be described by a function  $\hat{C}$  on the data  $f$  such that  $\text{Prob}_{\rho_0}[\hat{C} \leq g(\rho_0)] \geq \gamma$  holds for all  $\rho_0$  [26]. According to Hoeffding's tail inequality [31] and a given decomposition of  $L = \sum_r l_r^s M_r^s$  into the measurement operators  $M_r^s$ , a confidence region, then, is

$$\hat{C} = \text{tr}(\hat{\rho}_{\text{LIN}} L) - \sqrt{\frac{h^2 |\log(1 - \gamma)|}{2N_s}}, \quad (5)$$

where  $h^2$  is given by  $h^2 = \sum_s (l_{\max}^s - l_{\min}^s)^2$ , and  $l_{\max/\min}^s$  denotes the respective extrema of  $l_r^s$  over  $r$  for each

setting  $s$ . Although not optimal, such error regions are easy to evaluate and valid without extra assumptions. Since we directly compute a confidence region on  $g(x)$ , we obtain, generally, a tighter result than what would be obtained from a “smallest” confidence region on density operators which tend to drastically overestimate the error (see SM 4 [21] for an example).

In the following, we show how to use a linearized operator on the example of the bipartite negativity [30]. (For the quantum Fisher information [32] and additional discussion, see SM 5 [21].) A lower bound on the negativity  $N(\rho_{AB})$  of a bipartite state  $\rho_{AB}$  is given by  $N(\rho_{AB}) \geq \text{tr}(\rho_{AB} L)$  for any  $L$  satisfying  $\mathbb{1} \geq L^{T_A} \geq 0$ , where the superscript  $T_A$  denotes partial transposition [33] with respect to party  $A$ . This bound is tight if  $L$  is the projector on the negative eigenspace of  $\rho_{AB}^{T_A}$ . Using this linear expression, one can directly compute the lower bound on the negativity and, by using Eq. (5), the one-sided confidence region. Any choice of  $L$  is, in principle, valid, but, for a good performance,  $L$  should be chosen according to the experimental situation. We assume, however, no prior knowledge and rather estimate  $L$  independent of the tomographic data by the projection on the negative eigenspace of  $\hat{\rho}_{\text{ML}}^{T_A}$  deduced from an additional tomography, again with  $N_s = 100$  counts per setting. One can, of course, also start with an educated guess of  $L$  motivated by the target state one wants to prepare.

Figure 4 shows the distributions of the negativity between qubits  $A = \{1, 2\}$  and  $B = \{3, 4\}$  for the four-qubit GHZ state and for the separable four-qubit state  $|\psi_{\text{sep}}\rangle \propto (|0\rangle + |+\rangle)^{\otimes 4}$ , with  $|+\rangle \propto (|0\rangle + |1\rangle)$ , each mixed with white noise such that the fidelity with the respective pure state is 80%. In both cases, we observe that ML and LS overestimate the amount of entanglement. Even if no entanglement is present, ML and LS clearly indicate entanglement. In contrast, the lower bound of the negativity does not indicate spurious entanglement. As negativity

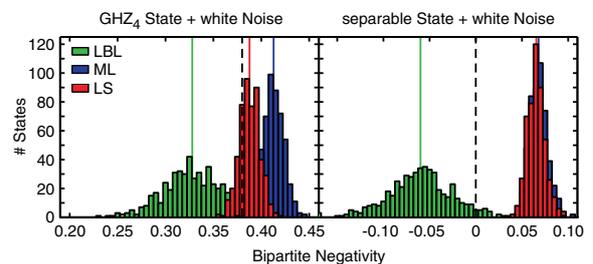


FIG. 4 (color online). Lower bound obtained by linearizing bipartite negativity (LBL) for the GHZ (left) and a four-qubit product state (right) both mixed with white noise resulting in 80% fidelity. The ML and LS reconstruction leads to a systematic overestimation of the negativity, while the lower bound yields a valid estimate.

gives lower bounds of other entanglement measures, those would overestimate entanglement of a state, too [34].

*Conclusion.*—Any state reconstruction algorithm enforcing physicality of the result suffers from systematic deviations. We have shown that, for the commonly used methods, this bias is significant for data sizes typical in current experiments. Quantities that are computed from such a point estimate can easily inherit this bias and lead to erroneous conclusions, as shown here on the examples of the fidelity, the negativity, and the Fisher information. Equivalent statements can be inferred for process tomography.

Recently, methods have been used to obtain confidence regions via the likelihood function. However, these are notoriously difficult to compute. The linearization method developed here yields a well defined confidence region for interesting quantities. This quantity is easily calculable, yet pessimistic. The quest is, thus, open for finding tighter, but still computationally accessible, confidence regions.

We would like to thank P. Drummond, B. Englert, D. Gross, H. K. Ng, M. Raymer, and J. Shang for stimulating discussions. This work has been supported by the EU (Marie Curie CIG 293993/ENFOQI, ERC Starting Grant GEDENTQOPT, FET Project QWAD, ERC QOLAPS), the BMBF (Chist-Era Project QUASAR), the Brazilian agency CAPES, through the program Science without Borders, and the Elite Network of Bavaria (PhD programs QCCC and ExQM).

- 
- [1] *Quantum State Estimation*, edited by M. G. A. Paris and J. Reháček, (Springer, Berlin, 2004).
- [2] H. Häffner, W. Hänsel, C. F. Roos, J. Benhelm, D. Chek-al-kar, M. Chwalla, T. Körber, U. D. Rapol, M. Riebe, P. O. Schmidt *et al.*, *Nature (London)* **438**, 643 (2005).
- [3] J. P. Home, D. Hanneke, J. D. Jost, J. M. Amini, D. Leibfried, and D. J. Wineland, *Science* **325**, 1227 (2009); J. P. Gaebler, A. M. Meier, T. R. Tan, R. Bowler, Y. Lin, D. Hanneke, J. D. Jost, J. P. Home, E. Knill, D. Leibfried, and D. J. Wineland, *Phys. Rev. Lett.* **108**, 260503 (2012).
- [4] D. F. V. James, P. G. Kwiat, W. J. Munro, and A. G. White, *Phys. Rev. A* **64**, 052312 (2001).
- [5] K. J. Resch, P. Walther, and A. Zeilinger, *Phys. Rev. Lett.*, **94**, 070402 (2005); N. Kiesel, C. Schmid, G. Tóth, E. Solano, and H. Weinfurter, *Phys. Rev. Lett.* **98**, 063604 (2007); M. W. Mitchell, C. W. Ellenor, S. Schneider, and A. M. Steinberg, *Phys. Rev. Lett.* **91**, 120402 (2003).
- [6] L. DiCarlo, J. M. Chow, J. M. Gambetta, L. S. Bishop, B. R. Johnson, D. I. Schuster, A. Majer, J. Blais, L. Frunzio, S. M. Girvin, and R. J. Schoelkopf, *Nature (London)* **460**, 240 (2009); M. Neeley, R. C. Bialczak, M. Lenander, E. Lucero, M. Mariantoni, A. D. O'Connell, D. Sank, H. Wang, M. Weides, J. Wenner *et al.*, *Nature (London)* **467**, 570 (2010); A. Fedorov, L. Steffen, M. Baur, M. da Silva, and A. Wallraff, *Nature (London)* **481**, 170 (2012); J. M. Chow, J. M. Gambetta, A. D. Córcoles, S. T. Merkel, J. A. Smolin, C. Rigetti, S. Poletto, G. A. Keefe, M. B. Rothwell,

- J. R. Rozen, M. B. Ketchen, and M. Steffen, *Phys. Rev. Lett.* **109**, 060501 (2012).
- [7] O. Mangold, A. Heidebrecht, and M. Mehring, *Phys. Rev. A* **70**, 042307 (2004); M. A. Nielsen, E. Knill, and R. Laflamme, *Nature (London)* **396**, 52 (1998).
- [8] I. L. Chuang and M. A. Nielsen, *J. Mod. Opt.* **44**, 2455 (1997).
- [9] D. Gross, Y.-K. Liu, S. T. Flammia, S. Becker, and J. Eisert, *Phys. Rev. Lett.* **105**, 150401 (2010); G. Tóth, W. Wicczorek, D. Gross, R. Krischek, C. Schwemmer, and H. Weinfurter, *Phys. Rev. Lett.* **105**, 250403 (2010); M. Cramer, M. B. Plenio, S. T. Flammia, D. Gross, S. D. Bartlett, R. Somma, O. Landon-Cardinal, Y.-K. Liu, and D. Poulin, *Nat. Commun.* **1**, 149 (2010); O. Landon-Cardinal and D. Poulin, *New J. Phys.* **14**, 085004 (2012).
- [10] T. Moroder, P. Hyllus, G. Tóth, C. Schwemmer, A. Niggebaum, S. Gaile, O. Gühne, and H. Weinfurter, *New J. Phys.* **14**, 105001 (2012).
- [11] M. Christandl and R. Renner, *Phys. Rev. Lett.* **109**, 120403 (2012); T. Sugiyama, P. S. Turner, and M. Muraio, *Phys. Rev. Lett.* **111**, 160406 (2013); R. Blume-Kohout, arXiv:1202.5270; J. Shang, H. K. Ng, A. Sehwat, X. Li, and B.-G. Englert, *New J. Phys.* **15**, 123026 (2013).
- [12] T. Moroder, M. Kleinmann, P. Schindler, T. Monz, O. Gühne, and R. Blatt, *Phys. Rev. Lett.* **110**, 180401 (2013); D. Rosset, R. Ferretti-Schöbitz, J.-D. Bancal, N. Gisin, and Y.-C. Liang, *Phys. Rev. A* **86**, 062325 (2012); C. Stark, *Phys. Rev. A* **89**, 052109 (2014); N. K. Langford, *New J. Phys.* **15**, 035003 (2013); S. J. van Enk and R. Blume-Kohout, *New J. Phys.* **15**, 025024 (2013); S. T. Merkel, J. M. Gambetta, J. A. Smolin, S. Poletto, A. D. Córcoles, B. R. Johnson, C. A. Ryan, and M. Steffen, *Phys. Rev. A* **87**, 062119 (2013).
- [13] D. Smithey, M. Beck, M. G. Raymer, and A. Faridani, *Phys. Rev. Lett.* **70**, 1244 (1993); G. M. D'Ariano, C. Macchiavello, and N. Sterpi, *Quantum Semiclass. Opt.* **9**, 929 (1997); U. Leonhardt, M. Munroe, T. Kiss, T. Richter, and M. Raymer, *Opt. Commun.* **127**, 144 (1996).
- [14] R. Blume-Kohout, *New J. Phys.* **12**, 043034 (2010); *Phys. Rev. Lett.* **105**, 200504 (2010).
- [15] Z. Hradil, *Phys. Rev. A* **55**, R1561 (1997).
- [16] T. Sugiyama, P. S. Turner, and M. Muraio, *New J. Phys.* **14**, 085005 (2012).
- [17] C. W. Helstrom, *Phys. Lett.*, **25A**, 101 (1967); S. L. Braunstein and C. M. Caves, *Phys. Rev. Lett.* **72**, 3439 (1994).
- [18] We stress that ML methods can, in principle, be used for directly estimating the fidelity without reconstructing the quantum state, but most experiments using ML first reconstruct the state and then calculate the fidelity. This procedure is the subject of our criticism.
- [19] B. Efron and R. J. Tibshirani, *An Introduction to the Bootstrap* (Chapman & Hall, London, 1994).
- [20] J. F. Poyatos, J. I. Cirac, and P. Zoller, *Phys. Rev. Lett.* **78**, 390 (1997); N. Kiesel, Ph.D. thesis, Ludwig-Maximilians-Universität München, 2007; T. Kiesel, *Phys. Rev. A* **85**, 052114 (2012).
- [21] See Supplemental Material at <http://link.aps.org/supplemental/10.1103/PhysRevLett.114.080403> for details on LIN, additional examples of states with a strong bias, examples for bootstrapping, the difference between multi-parameter and scalar confidence regions, and further details

- on the proposed linearization method for reliable confidence regions, which includes Ref. [22].
- [22] R. A. Horn and C. R. Johnson, *Topics in Matrix Analysis* (Cambridge University Press, Cambridge, England, 1991).
- [23] Clearly, LIN is only reasonable if tomographically complete information about the state is available.
- [24] S. Boyd and S. Vandenberghe, *Convex Optimization* (Cambridge University Press, Cambridge, England, 2004).
- [25] J. A. Smolin, *Phys. Rev. A* **63**, 032306 (2001).
- [26] A. F. Mood, *Introduction to the Theory of Statistics* (McGraw-Hill Inc., New York, 1974).
- [27] It is well known that estimators can be biased for a finite number of samples or for particular additional conditions and constraints. An example is the estimate of the variance  $\sigma^2$  of a Gaussian distribution with known mean  $\mu$ . The ML estimate from  $N$  samples  $x_i$  gives  $\hat{\sigma}_{\text{ML}}^2(x_i) = 1/N \sum_i (x_i - \mu)^2$ , while an unbiased estimator is  $\hat{\sigma}^2(x_i) = (1/N - 1) \sum_i (x_i - \mu)^2$ .
- [28] This proof, given here for pure states, can be generalized for mixed states  $\rho_1$  and  $\rho_2$  with the ranges of the two states not sharing any vector.
- [29] O. Gühne, M. Reimpell, and R. F. Werner, *Phys. Rev. Lett.* **98**, 110502 (2007); J. Eisert, F. G. S. L. Brandao, and K. M. R. Audenaert, *New J. Phys.* **9**, 46 (2007).
- [30] G. Vidal and R. F. Werner, *Phys. Rev. A* **65**, 032314 (2002).
- [31] M. Tomamichel, C. C. W. Lim, N. Gisin, and R. Renner, *Nat. Commun.* **3**, 634 (2012); S. T. Flammia and Y.-K. Liu, *Phys. Rev. Lett.* **106**, 230501 (2011); M. P. da Silva, O. Landon-Cardinal, and D. Poulin, *Phys. Rev. Lett.* **107**, 210404 (2011).
- [32] D. Petz and C. Ghinea, QP-PQ: Quantum Probab. White Noise Anal. **27**, 261 (2011).
- [33] A. Peres, *Phys. Rev. Lett.* **77**, 1413 (1996).
- [34] K. Chen, S. Albeverio, and S.-M. Fei, *Phys. Rev. Lett.* **95**, 040504 (2005); **95**, 210501 (2005).

## SUPPLEMENTAL MATERIAL: SYSTEMATIC ERRORS IN CURRENT QUANTUM STATE TOMOGRAPHY TOOLS

Christian Schwemmer,<sup>1</sup> Lukas Knips,<sup>1</sup> Daniel Richart,<sup>1</sup> and Harald Weinfurter,<sup>1</sup> Tobias Moroder,<sup>2</sup> Matthias Kleinmann,<sup>2,3,4</sup> and Otfried Gühne,<sup>2</sup>

<sup>1</sup>Max-Planck-Institut für Quantenoptik, Hans-Kopfermann-Straße 1, D-85748 Garching, Germany and Department für Physik, Ludwig-Maximilians-Universität, D-80797 München, Germany

<sup>2</sup>Naturwissenschaftlich-Technische Fakultät, Universität Siegen, Walter-Flex-Straße 3, D-57068 Siegen, Germany

<sup>3</sup>Departamento de Matemática, Universidade Federal de Minas Gerais, Caixa Postal 702, Belo Horizonte, Minas Gerais 31270-901, Brazil

<sup>4</sup>Department of Theoretical Physics, University of the Basque Country UPV/EHU, P.O. Box 644, E-48080 Bilbao, Spain

### SM 1: QUANTUM STATE RECONSTRUCTION USING LINEAR INVERSION

In [4] it is explained how to obtain the estimate  $\hat{\rho}_{\text{LIN}}$  for an  $n$ -qubit state from the observed frequencies of a complete set of projection measurements, i.e.  $4^n$  results. Yet, the scheme described there is more general and can be used for any (over)complete set of projection measurements.

In the standard Pauli basis  $\{\sigma_0, \sigma_x, \sigma_y, \sigma_z\}$  the density matrix of the state  $\rho$  is given by

$$\rho = \frac{1}{2^n} \sum_{\mu} T_{\mu} \Gamma_{\mu} \quad (\text{S1})$$

where  $\mu = 1 \dots 4^n$  enumerates all possible  $n$ -fold tensor products of Pauli matrices  $\Gamma_1 = \sigma_0 \otimes \sigma_0 \otimes \dots \otimes \sigma_0$ ,  $\Gamma_2 = \sigma_0 \otimes \sigma_0 \otimes \dots \otimes \sigma_x$ , etc. and with correlations  $T_{\mu} = \text{tr}(\rho \Gamma_{\mu})$ . To simplify our notation we will use the following mapping for a setting  $s$  with a respective outcome  $r$ :  $(r, s) \rightarrow \nu = 2^{n(s-1)} + r - 1$ , hence for the projectors,  $M_r^s \rightarrow M_{\nu}$ , and for the  $A_r^s \rightarrow A_{\nu}$ , etc. Then the probabilities to observe a result  $r$  for setting  $s$ , or  $\nu$  respectively, are given by

$$P_{\nu} = \text{tr}(\rho M_{\nu}) = \frac{1}{2^n} \sum_{\mu} \text{tr}(M_{\nu} \Gamma_{\mu}) T_{\mu}. \quad (\text{S2})$$

Introducing the matrix  $\hat{B}$  with elements

$$B_{\nu, \mu} = \frac{1}{2^n} \text{tr}(M_{\nu} \Gamma_{\mu}) \quad (\text{S3})$$

Eq. (S2) simplifies to

$$\vec{P} = \hat{B} \vec{T}. \quad (\text{S4})$$

Inverting Eq. (S4), the correlations can be obtained from the probabilities  $P_{\nu}$ , i.e.,  $T_{\mu} = \sum_{\nu} (\hat{B}^{-1})_{\mu, \nu} P_{\nu}$ . Note that this is possible for any set of measurement operators. In case of a tomographically overcomplete set, the inverse  $\hat{B}^{-1}$  has to be replaced by the pseudo inverse  $\hat{B}^{-1} \rightarrow$

$B^+ = (B^{\dagger} B)^{-1} B^{\dagger}$ . Reinserting  $T_{\mu}$  one obtains

$$\rho = \frac{1}{2^n} \sum_{\nu, \mu} (\hat{B}^{-1})_{\mu, \nu} \Gamma_{\mu} P_{\nu}. \quad (\text{S5})$$

For finite data sets, the  $P_{\nu}$  are replaced by the frequencies  $f_{\nu}$  and with

$$A_{\nu} = \frac{1}{2^n} \sum_{\mu} (\hat{B}^{-1})_{\mu, \nu} \Gamma_{\mu} \quad (\text{S6})$$

Eq. (1) is obtained.

### SM 2: BIAS FOR OTHER PROMINENT STATES

The occurrence of a bias for fidelity estimation based on ML and LS state reconstruction is by no means a special feature of the GHZ state. In Fig. S1 we show some further examples of the corresponding dependencies of the bias on the number of measurements per setting  $N_s$  for the  $W$  and the fully separable state  $|\psi\rangle \propto (|0\rangle + |+\rangle)^{\otimes 4}$ . For all these pure states we assume that they are mixed with white noise for an overall initial fidelity of 80%, so that the states are not at the border of the state space.

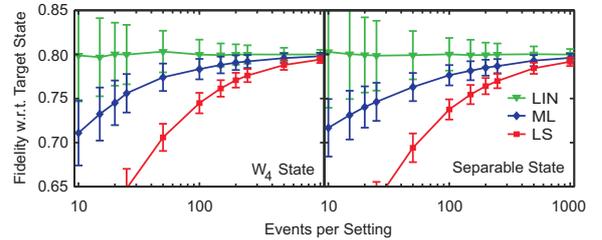


FIG. S1. The behavior of ML, LS and LIN depending on the number of events  $N_s$  per setting for different noisy initial states  $\rho_0$ .

Furthermore we observed that the fidelity values as inferred via LS are systematically lower than those obtained using ML, see Fig. S2.

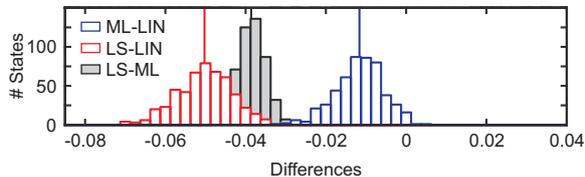


FIG. S2. Here we show the differences of the respective fidelity estimates evaluated for each single simulated tomography experiment as shown in Fig. 1 of the main text. It shows that the respective ML or LS estimate, with one rare exception, is always lower than the LIN estimate. Comparing ML and LS (gray) shows that not only on average but also for every single data set LS delivers a smaller fidelity value than ML.

### SM 3: BOOTSTRAPPING

As already mentioned in the main text, in many publications where QST is performed the standard error bar is calculated by bootstrapping based on Monte Carlo methods. One can here distinguish between parametric bootstrapping, where  $f^{(i)}$  are sampled according to  $\hat{P}^s(r) = \text{tr}(\hat{\rho}(f_{\text{obs}})M_r^s)$ , and non-parametric bootstrapping, where  $\hat{P}^s(r) = f_{\text{obs}}$  is used instead.

We consider again the four-qubit GHZ state of 80% fidelity and  $N_s = 100$ . Interpreting the simulations of Fig. 1 as Monte-Carlo simulations from the parametric bootstrap with  $\hat{P}^s(r) = \text{tr}(\rho_0 M_r^s)$  we have already seen that ML and LS yield fidelity estimates below the actual value. If one uses now one of these data sets  $f_{\text{obs}}$  as a seed to generate new samples  $f^{(i)}$  the fidelity decreases further. As shown in Fig. S3 this happens in particular for parametric bootstrapping ( $0.777 \pm 0.011$  for ML and  $0.700 \pm 0.012$  for LS) while non-parametric bootstrapping ( $0.780 \pm 0.011$  for ML and  $0.714 \pm 0.012$  for LS) weakens this effect. However, in this context, one is interested in fact in the standard deviation of the simulated distribution. In our simulations it is somewhat smaller than the distribution of linearly evaluated fidelities. This means, the biasedness of ML and LS methods leads to a false estimate of the error, too.

### SM 4: CONFIDENCE REGIONS FOR STATES VS. SCALAR QUANTITIES

Let us now comment on confidence regions (CR) for density operators and CR on parameter functions  $Q$ . Having a (tractable) method to compute CR for states  $\hat{C}_\rho(f)$  [11], one could think that this region of states also provides good CR for the parameter functions  $Q$ , if one manages to evaluate the minimal and maximal values of  $Q(\rho)$  for all  $\rho \in \hat{C}_\rho(f)$ . However, such CR are typically much worse than CR evaluated for  $Q$  directly, the reason being the large freedom in how to build up a CR. Let us

give the following illustrative example, see also Fig. S4:

Let us consider the task to obtain a CR for the two mean values  $\vec{\mu} = (\mu_1, \mu_2)$  of two independent Gaussian experiments, where the first  $N$  samples  $x_i$  are drawn from  $\mathcal{N}(\mu_1, \sigma^2)$  while the remaining  $N$  instances  $y_i$  originate from  $\mathcal{N}(\mu_2, \sigma^2)$ , both with the same known variances. If one is interested in an 68% CR for both mean values  $\vec{\mu}$  then both possible recipes

$$\hat{C}^{(1)} = \{\vec{\mu} : \|\vec{\mu} - (\bar{x}, \bar{y})\| \leq 1.52\sigma/\sqrt{N}\}, \quad (\text{S7})$$

$$\hat{C}^{(2)} = [\bar{x} - \sigma/\sqrt{N}, \bar{x} + \sigma/\sqrt{N}] \times (-\infty, \infty) \quad (\text{S8})$$

with  $\bar{x} = \frac{1}{N} \sum_i x_i$  and similar for  $\bar{y}$  are valid 68% CR. However, while  $\hat{C}^{(1)}$  yields the smallest area for the CR, it gives a much larger confidence region for  $Q(\vec{\mu}) = \mu_1$  than if we would directly use  $\hat{C}^{(2)}$ , which in fact is the smallest one for  $\mu_1$ . Note that this effect increases roughly with  $\sqrt{\text{dim}}$  if one adds further parameters in the considered Gaussian example. Therefore we see that “errors” associated with CR on the density operator are not the best choice if one is interested only on a few key properties of the state.

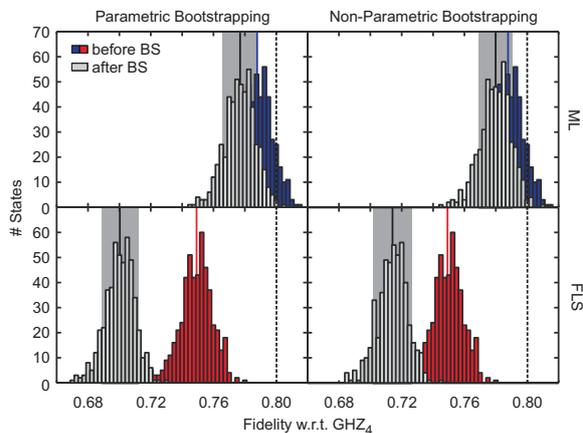


FIG. S3. Error bar computation for the fidelity of the four-qubit GHZ state via Monte-Carlo simulation using either parametric or non-parametric bootstrapping with the data from Fig. 1. For each of these 500 observations  $f_{\text{obs}}$ , 100 new data sets  $f^{(i)}$  were generated and reconstructed in order to deduce the mean and standard deviation as an error bar for the fidelity. The histograms denoted by “after BS” show the distributions of these means together with an averaged error bar given by the gray shaded areas. The initial values for the fidelities are described by the dashed lines.

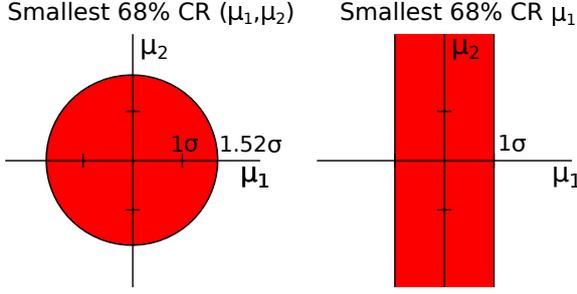


FIG. S4. Which confidence region is the smallest? If one is interested in both mean values  $\vec{\mu} = (\mu_1, \mu_2)$  then clearly the left one represents the smallest one, but if  $Q(\vec{\mu}) = \mu_1$  is chosen, then the CR obtained by projection onto the  $\mu_1$ -axis is much larger for the disc-shaped area.

### SM 5: BOUNDS ON CONVEX/CONCAVE FUNCTIONS

As mentioned in the main text, one can directly bound convex (or concave) functions  $g(x)$  by linear ones using an operator  $L$

$$\text{tr}(\varrho_0 L) \leq g(\varrho_0). \quad (\text{S9})$$

Here, we want to explain in detail how the operator  $L$  can be determined from the derivatives of  $g[\varrho(x)]$ . Therefore, we parametrize the density operator  $\varrho(x) = \mathbb{1}/\text{dim} + \sum_i x_i S_i$  via an orthonormal basis  $S_i$  of Hermitian traceless operators. A possible choice for the  $S_i$  are all normalized traceless tensor products of the Pauli matrices and the identity. Since we employ an affine parametrization, the function  $g(x) = g[\varrho(x)]$  is convex. Direct calculation shows that choosing the operator  $L[\varrho(x')] = l_0 \mathbb{1} + \sum_i l_i S_i$  as

$$l_0 = g[\varrho_{\text{guess}}(x')] - \sum_i x'_i \frac{\partial}{\partial x_i} g[\varrho_{\text{guess}}(x')] \quad (\text{S10})$$

$$l_i = \frac{\partial}{\partial x_i} g[\varrho_{\text{guess}}(x')] \quad (\text{S11})$$

gives due to the convexity condition  $g(x) \geq g(x') + \nabla g(x')^T (x - x')$  a lower bound as in Eq. (S9). Here,  $L[\varrho(x')]$  is computed on a “guess”  $x'$ , i.e.,  $\varrho_{\text{guess}}(x')$  of the true state  $\varrho_0$ . Recall that while the guess  $\varrho_{\text{guess}}$  must be a valid state the lower bound  $\text{tr}(\varrho_0 L)$  is well-defined also for nonphysical density operators.

As an example how to apply this linearization, let us consider the quantum Fisher information  $f(x) = F(\varrho, H)$ , which measures the suitability of a state  $\varrho$  to determine the parameter  $\theta$  in an evolution  $U(\theta, H) =$

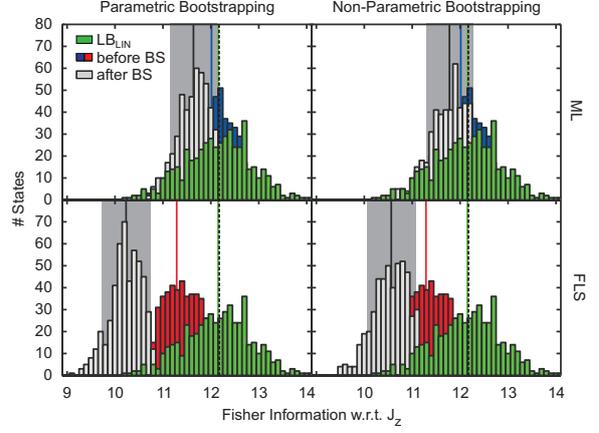


FIG. S5. Full analysis of a Pauli QST scheme with  $N_s = 100$  on four qubits in order to deduce the quantum Fisher information with respect to  $H = J_z$ . As the true underlying state we assume again a noisy four-party GHZ state. We observe that the quantum Fisher information is underestimated from both ML and LS, while the lower bound deduced from LIN is fine.

$e^{-i\theta H}$ . More explicitly the formulae are given by

$$f(x) = 2 \sum_{jk} \frac{(\lambda_j - \lambda_k)^2}{\lambda_j + \lambda_k} H_{jk} H_{kj}, \quad (\text{S12})$$

$$\frac{\partial}{\partial x_i} f(x) = 4 \sum_{jkl} \frac{\lambda_j \lambda_k + \lambda_j \lambda_l + \lambda_k \lambda_l - 3\lambda_j^2}{(\lambda_j + \lambda_k)(\lambda_j + \lambda_l)} H_{jk} S_{i,kl} H_{lj} \quad (\text{S13})$$

where  $\{\lambda_i, |\psi_i\rangle\}$  denotes the eigenspectrum of  $\varrho(x)$ ,  $H_{jk} = \langle \psi_j | H | \psi_k \rangle$  and  $S_{i,kl} = \langle \psi_k | S_i | \psi_l \rangle$ . In order to compute the derivative of the Fisher information one can employ the alternative form, as given for instance in Ref. [32],

$$F(\varrho, H) = \text{tr}[(H\varrho^2 + \varrho^2 H - 2\varrho H\varrho) J_\varrho^{-1}(H)], \quad (\text{S14})$$

$$J_\varrho^{-1}(H) = \int_0^\infty dt e^{-t/2\varrho} H e^{-t/2\varrho}. \quad (\text{S15})$$

such that the derivative can be computed via the help of matrix derivatives [22].

Now let us imagine that we want to determine the quantum Fisher information of a four-qubit state with respect to  $H = J_z$ , while our true underlying state  $\varrho_0$  is once more the noisy GHZ state of 80% fidelity. Figure S5 shows the full simulation of a Pauli tomography experiment with  $N_s = 100$  together with the standard error analysis using parametric or non-parametric bootstrapping. As with the other examples, we observe a systematic discrepancy between the results of standard QST tools and the true value. In this case, though the quantum Fisher information is typically larger for stronger

entangled states, ML or LS underestimate the true capabilities of the state. However, if we use the described method for LIN (with an in this case optimized operator  $L$ ) the lower bound via LIN is fine.

$$g(x) = -\text{tr}(\varrho \log \varrho)$$

$$\frac{\partial}{\partial x_i} g(x) = -\text{tr}[S_i(\log \varrho(x))]. \quad (\text{S17})$$

For completeness, we also give the respective derivatives for further convex functions of interest like the purity  $g(x) = \text{tr}(\varrho^2)$

$$\frac{\partial}{\partial x_i} g(x) = 2\text{tr}[S_i \varrho(x)] \quad (\text{S16})$$

and correspondingly for the von Neumann entropy

- 
- [4] D. F. V. James, P. G. Kwiat, W. J. Munro, and A. G. White, *Phys. Rev. A* **64**, 052312 (2001).
  - [11] M. Christandl and R. Renner, *Phys. Rev. Lett.* **109**, 120403 (2012); T. Sugiyama, P. S. Turner, and M. Muraio, *Phys. Rev. Lett.* **111**, 160406 (2013); R. Blume-Kohout, arXiv:1202.5270; J. Shang, H. K. Ng, A. Sehwat, X. Li, and B.-G. Englert, *New J. Phys.* **15**, 123026 (2013).
  - [22] R. A. Horn and C. R. Johnson, *Topics in matrix analysis* (Cambridge University Press, 1991).
  - [32] D. Petz and C. Ghinea, *QP-PQ: Quantum Probab. White Noise Anal.* **27**, 261 (2011).

## 3.4 Multiqubit State Tomography with Finite Data

In the previous sections, problems of constrained optimization methods such as maximum likelihood estimation have been discussed. In this section, the reason for the appearance of negative eigenvalues in the first place will be investigated. Here, the consequences of finite statistics will be studied in greater detail, leading to another quantum state estimation method, which is based on understanding the effects of finite statistics.

Also, the effects of finite statistics in quantum state tomography may also be used to certify a given set of experimental data. Manipulation of experimental data, either intentionally or accidentally, can be identified, which may allow to verify the integrity of measurements.

### 3.4.1 Wigner Semicircle Distribution

For increasing statistics, the relative frequencies obtained in quantum state estimation will tend towards the actual (unknown) probabilities. In the hypothetical scenario of infinite sample size, the frequencies actually become equal to the probabilities. Thus, the reconstructed state equals the unknown state.

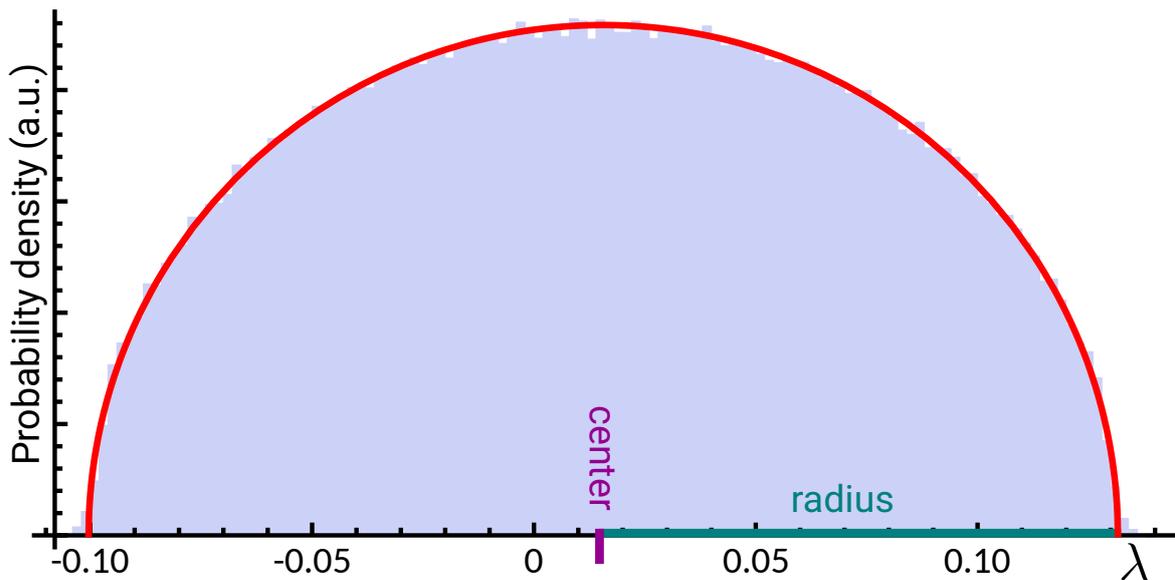


FIGURE 3.4: The Wigner semicircle distribution describes the distribution of eigenvalues in a tomography experiment with finite statistics and an infinite number of qubits. Already for a modest number of qubits, the description, which is based on two parameters only, the radius and the center, provides a suitable approximation. Here, the distribution of 10 000 simulated tomographies of the maximally mixed state with six qubits and 100 measurements per basis setting is shown.

Conversely, for finite statistics, the deviations between frequencies and probabilities

will lead to statistical errors. Consider first the maximally mixed state with  $n$  qubits. As its density matrix is proportional to the identity matrix with  $2^n \times 2^n$  elements, it has  $2^n$  degenerate eigenvalues. The finite statistics now leads to a lifting of the degeneracy. When using the overcomplete Pauli tomography and for a sufficiently large number of qubits, the eigenvalues will be distributed according to the Wigner semicircle distribution, see Fig. 3.4. The emergence of this distribution will be explained later and proven in the publication [P2], which is reprinted in Sec. 3.4.6.

Although the Wigner semicircle distribution describes the distribution of eigenvalues only in the limit of large density matrices, it is a sufficient approximation already for a modest number of qubits. A characteristic feature of the semicircle distribution, described by its radius  $R$  and the center  $c$ , is its compact support with the clear cutoffs at  $c \pm R$ . Given a distribution of eigenvalues for a finite number of qubits, the best fitting semicircle can be found. In Tab. 3.2, the radius and the center of the approximative Wigner semicircle are estimated for simulations of the maximally mixed state and overcomplete Pauli tomography with  $N = 100$  samples per basis setting for various numbers of qubits  $n$ . In order to assess the behavior of the eigenvalue distribution of density matrices with finite  $n$  close to the point, at which the theoretical model in the limit  $n \rightarrow \infty$  predicts a cutoff, numerical simulations have been conducted. A linear approximation in the vicinity of the expected cutoff at  $c - R$  and  $c + R$  provided the estimates for the slopes  $s_1$  and  $s_2$  of distributions of the numerically simulated density matrices, respectively. Hence, a large positive value for  $s_1$ , i.e., for the slope at the left hand side at which the distribution is rising, and a large negative value for  $s_2$ , i.e., for the slope at the right hand side at which the distribution is falling, indicate steep tails. As the evaluation is based on a finite number of simulated samples and relies on a difference quotient with finite interval length, the values  $s_1$  and  $s_2$  are only rough estimates. Due to the symmetry of the Wigner semicircle, in the following, only the averaged slope  $(|s_1| + |s_2|)/2$  is considered.

TABLE 3.2: Radius  $R$  and center  $c$  of the approximative Wigner semicircle are given for the  $n$  qubit maximally mixed state. Around the cutoffs  $c \pm R$ , the slopes  $s_1$  and  $s_2$  of the distributions are estimated by means of linear fits. The mean of the moduli of the slopes,  $\frac{|s_1| + |s_2|}{2}$ , indicates the increasing steepness at the tails of the distribution. The last column shows the relative amount of samples outside the interval  $\mathcal{I} = [c - R, c + R]$ . Fig. 3.5 shows  $\frac{|s_1| + |s_2|}{2}$  versus the number of qubits  $n$ .

n	samples	$R$	$c$	$s_1$	$s_2$	$\frac{ s_1  +  s_2 }{2}$	$ \{\lambda_i : \lambda_i \notin \mathcal{I}\} $
1	$200 \times 10^4$	0.18	0.50	8.5	-9.2	8.9	0.47%
2	$400 \times 10^4$	0.17	0.25	20.9	-24.2	22.5	0.99%
3	$800 \times 10^4$	0.15	0.13	40.2	-45.1	42.6	0.69%
4	$1390 \times 10^4$	0.14	0.06	60.8	-76.3	68.5	0.46%
5	$4 \times 10^4$	0.13	0.03	67.2	-144.5	105.9	0.26%
6	$64 \times 10^4$	0.12	0.02	152.2	-200.1	176.1	0.19%

As can be seen in Tab. 3.2, the slope at the anticipated cutoffs of a corresponding Wigner semicircle distribution strongly increases with increasing number of qubits  $n$ . Fig. 3.5

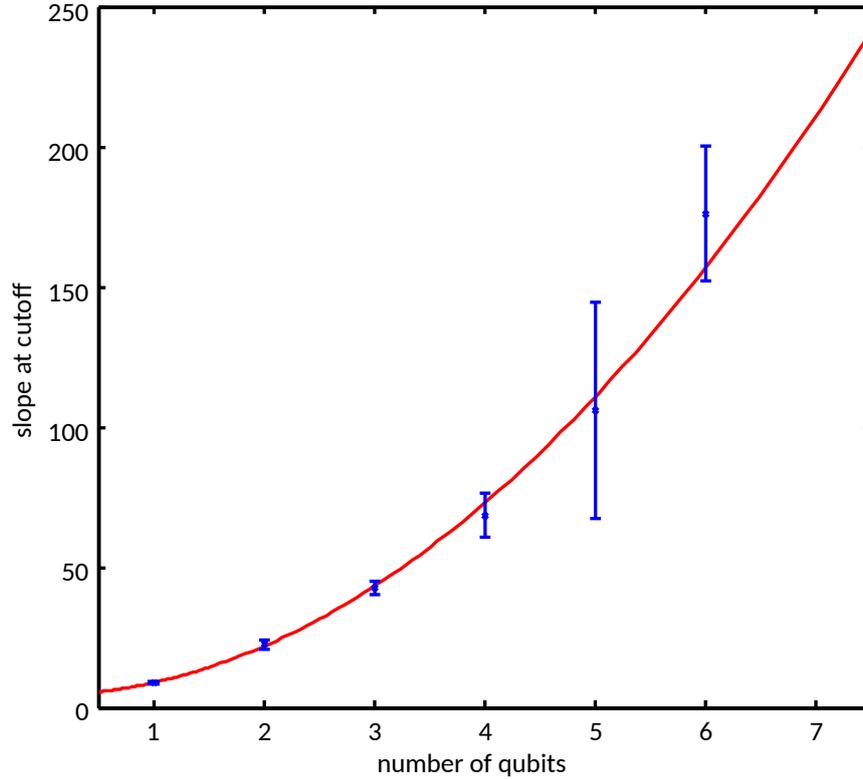


FIGURE 3.5: Slope of simulated eigenvalue distribution at  $c \pm R$ . For various numbers of qubits, the averaged moduli of the slopes are shown with blue markers, see Tab. 3.2. The fit (red line) corresponds to the fitmodel  $a(x^b + 1)$  with  $a \approx 4.4$ ,  $b \approx 2.0$ . Obviously, for increasing number of qubits, the tails of the distribution become less pronounced and turn into a steep cutoff.

shows this relation, indicating that the steepness of the simulated eigenvalue distributions diverges, as one expects for a semicircle distribution. Thus, already for a modest number of qubits, the spectral distribution features strongly damped tails. For more details, see also the discussion on the smallest eigenvalue in appendix C of [P2].

### 3.4.2 Catalan Numbers

The so called *Catalan numbers* play a significant role in proving that the distribution of eigenvalues leads to the Wigner semicircle distribution when performing quantum state tomography with finite statistics. An ideal semicircle distribution centered around 0 with radius  $R$  can be described by

$$f(x) = \frac{2}{\pi R^2} \begin{cases} \sqrt{R^2 - x^2} & x^2 \leq R^2, \\ 0 & x^2 > R^2. \end{cases} \quad (3.17)$$

A general description for the sequence of all moments  $m_k^{\text{sc}}$  of this distribution can be obtained using the Catalan numbers  $\mathcal{C}_n$  as

$$m_k^{\text{sc}} \equiv \int_{-R}^R dx f(x) x^k = \begin{cases} 1 & k = 0, \\ 0 & k = 1, \\ \left(\frac{R}{2}\right)^2 & k = 2, \\ 0 & k = 3, \\ 2 \left(\frac{R}{2}\right)^4 & k = 4, \\ 0 & k = 2n + 1, \\ \mathcal{C}_n \left(\frac{R}{2}\right)^k & k = 2n. \end{cases} \quad (3.18)$$

This important fact will be used in the following section to prove the appearance of the semicircle distribution in quantum state tomography.

The Catalan numbers are a series of numbers which can be used to count the number of so-called *non-crossing partitions* [172].

When four persons (e.g., Alice, Bob, Charlie, and Dave) meet and stand together, there are two out of three total possibilities for them to shake hands without the arms of one pair crossing another pair's arms (in the arrangement ABCD possible combinations are AB/CD and AD/BC, while AC/BD contains crossing arms). For six persons, there are in total nine possible ways how those persons can shake hands, i.e., how those persons can be partitioned, see Fig. 3.6. Out of those, five partitions are so called *non-crossing* as shown in green, while the remaining four possibilities contain crossing arms or lines, respectively.

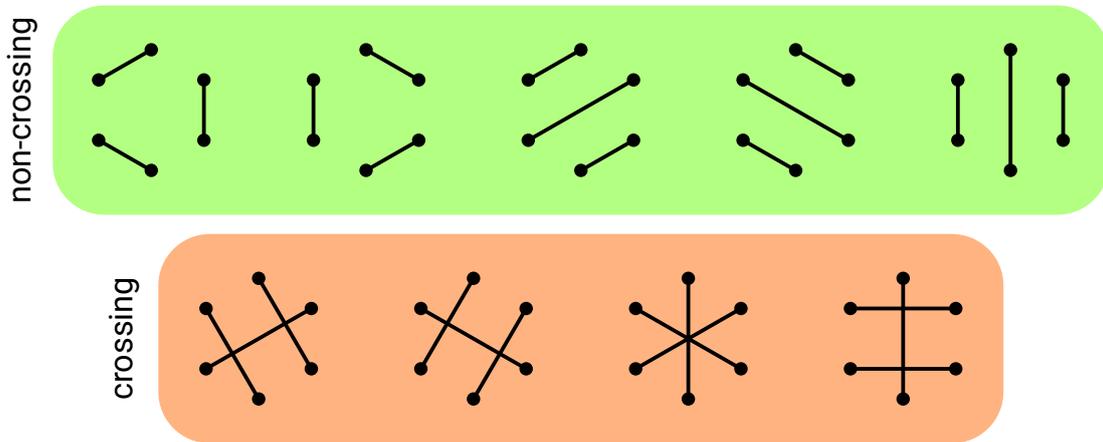


FIGURE 3.6: Non-crossing and crossing partitions for six elements. The upper (green shaded) row shows all possibilities how to connect pairwise two vertices such that those connections do not cross. The lower (orange shaded) row lists the four crossing partitions. For example, consider the vertices to be persons shaking hands along the edges, as described in the text.

The number of non-crossing partitions also corresponds to counting the number of valid possibilities how to arrange  $n$  pairs of parentheses. Three pairs can be arranged in five

possible ways,  $((()))$ ,  $(()())$ ,  $()(())$ ,  $()(())$ ,  $(())()$ , while  $)()()$ ,  $)()()$ ,  $)()()$ ,  $)()()$ ,  $)()()$  are invalid combinations. Another equivalent example is the number of combinations of going in total 3 steps upwards and 3 steps downwards under the constraint of staying at least at the initial height (cf. climbing a ladder starting at the floor).

The Catalan numbers are given by

$$\mathcal{C}_n = \frac{1}{n+1} \binom{2n}{n} = \binom{2n}{n} - \binom{2n}{n+1}. \quad (3.19)$$

The first Catalan numbers starting with  $\mathcal{C}_0 = 1$  are then  $\{1, 1, 2, 5, 14, 42, \dots\}$ . In the above examples, the number of non-crossing partitions for 4 elements is given by  $\mathcal{C}_2 = 2$  and for 6 elements by  $\mathcal{C}_3 = 5$ .

### 3.4.3 Illustration of Proof - Example of Fourth Moment

In [P2], which is reprinted in Sec. 3.4.6, it is proven that for finite statistics, the distribution of eigenvalues of a tomographically obtained maximally mixed state tends towards a Wigner semicircle distribution [173]. To better understand the principle of the proof in detail, in this section an essential part of it will be illustrated.

As shown in Eq. (3.18), an ideal semicircle distribution results in statistical moments which depend on the radius and the Catalan numbers  $\mathcal{C}_n$ . If it can be shown that all moments of the eigenvalue distribution are identical to the moments of an ideal semicircle distribution, those distributions are shown to be equal, as a semicircle is defined on a compact support, see Sec. 2.2.6.2.

For that purpose, it can easily be shown that odd moments of the spectral distribution vanish. Here, to illustrate the method, the calculation of even moments will be illustrated by means of the fourth moment  $m_4^{\text{ev}}$  of the eigenvalues of the difference matrix  $\Delta\rho = \frac{1}{2^n} \sum_{\mu} \Delta T_{\mu} \sigma_{\mu}$  with  $\Delta T_{\mu} = \tilde{T}_{\mu} - T_{\mu}$ , where  $\tilde{T}_{\mu}$  denote the experimentally obtained correlations in the basis with index  $\mu$ . For the maximally mixed state, the theoretical correlations  $T_{\mu}$  vanish except of  $T_{0,\dots,0} = 1$ . Please note that the difference matrix  $\Delta\rho$  is considered instead of  $\rho$  as this simplifies the further calculation. As for the maximally mixed state all but  $T_{0,\dots,0}$  vanish, by subtracting the diagonal matrix  $\rho$  from  $\tilde{\rho}$ , all elements of the matrix  $\Delta\rho$  and therefore all eigenvalues are centered around 0 instead of the diagonal elements (and the eigenvalues) being centered around  $c = 1/2^n$ . Hence, considering the difference matrix  $\Delta\rho$  corresponds in this case to shifting the center of the eigenvalue distribution from  $c$  to 0.

Although the eigenvalues of a density matrix are due to the normalization condition  $\sum_i \lambda_i = 1$  not independent of each other, for multiqubit systems, one can still use the approximation that they are independent. The distribution of the eigenvalues as shown in Fig. 3.4 is supposed to describe the probability distribution for *any* eigenvalue, when picking one at random, and obviously differs from the distribution of only the, say, smallest eigenvalue. In order to incorporate this into the derivation, one can first approximate the joint spectral probability distribution  $p(\lambda_1, \lambda_2, \dots, \lambda_{2^n})$  by the product of the almost

independent probability distributions of each eigenvalue,

$$p(\lambda_1, \lambda_2, \dots, \lambda_{2^n}) \approx \prod_{i=1}^{2^n} p(\lambda_i). \quad (3.20)$$

To describe now the distribution of a randomly picked eigenvalue, one can marginalize over all but one specified eigenvalue and average of the specification,

$$\begin{aligned} p(\lambda) \approx \frac{1}{2^n} & \left[ p(\lambda_1) \prod_{i=2}^{2^n} \int_{-\infty}^{\infty} d\lambda_i p(\lambda_i) + p(\lambda_2) \prod_{i=1; i \neq 2}^{2^n} \int_{-\infty}^{\infty} d\lambda_i p(\lambda_i) + \right. \\ & \left. \dots + p(\lambda_{2^n}) \prod_{i=2}^{2^n-1} \int_{-\infty}^{\infty} d\lambda_i p(\lambda_i) \right]. \end{aligned} \quad (3.21)$$

As the integral over each eigenvalue is normalized, i.e., the probability to find the  $i$ -th eigenvalue *anywhere* is 1, one finally obtains

$$p(\lambda) \approx \frac{1}{2^n} \sum_{i=1}^{2^n} p(\lambda_i) \quad (3.22)$$

for the probability distribution for a randomly chosen eigenvalue. As the  $k$ -th moment of the  $i$ -th eigenvalue of the centralized matrix  $\Delta\varrho$  is given by

$$m_{k,i}^{\text{ev}} = \int_{-\infty}^{\infty} d\lambda_i p(\lambda_i) \lambda_i^k = \mathbb{E}[\lambda_i^k], \quad (3.23)$$

where  $p(\lambda_i)$  describes the probability distribution of the  $i$ -th eigenvalue, the  $k$ -th moment of the distribution of the randomly picked eigenvalue is given by

$$m_k^{\text{ev}} \approx \int_{-\infty}^{\infty} d\lambda p(\lambda) \lambda^k = \frac{1}{2^n} \sum_i m_{k,i}^{\text{ev}} = \frac{1}{2^n} \sum_i \mathbb{E}[\lambda_i^k], \quad (3.24)$$

which therefore is an expectation value (over the various eigenvalues) of an expectation value (over the various distributions).  $m_4^{\text{ev}}$  is then given by

$$\begin{aligned} m_4^{\text{ev}} & \approx \frac{1}{2^n} \sum_{i=1}^{2^n} \mathbb{E}[\lambda_i^4] = \frac{1}{2^n} \mathbb{E}[\text{tr}((\Delta\varrho)^4)] \\ & = \frac{1}{2^{5n}} \mathbb{E} \left[ \text{tr} \sum_{\mu, \nu, \gamma, \lambda} \Delta T_\mu \Delta T_\nu \Delta T_\gamma \Delta T_\lambda \sigma_\mu \sigma_\nu \sigma_\gamma \sigma_\lambda \right] \\ & = \frac{1}{2^{5n}} \sum_{\mu, \nu, \gamma, \lambda} \mathbb{E}[\Delta T_\mu \Delta T_\nu \Delta T_\gamma \Delta T_\lambda] \text{tr}(\sigma_\mu \sigma_\nu \sigma_\gamma \sigma_\lambda). \end{aligned} \quad (3.25)$$

As  $\Delta T_\mu$  is a random variable following a Gaussian distribution with  $\mathbb{E}[\Delta T_\mu] = 0$  and  $\mathbb{E}[\Delta T_\mu^2] > 0$ , the respective summand vanishes whenever one of the indices appears in an odd power. Hence,

$$m_4^{\text{ev}} \approx \frac{1}{2^{5n}} \sum_{\mu} [\mathbb{E}[\Delta T_\mu^4] \text{tr}(\sigma_\mu \sigma_\mu \sigma_\mu \sigma_\mu) + \frac{1}{2} \sum_{\nu; \mu \neq \nu} \mathbb{E}[\Delta T_\mu^2 \Delta T_\nu^2] \text{tr}(\sigma_\mu \sigma_\mu \sigma_\nu \sigma_\nu + \sigma_\mu \sigma_\nu \sigma_\mu \sigma_\nu + \sigma_\mu \sigma_\nu \sigma_\nu \sigma_\mu + \sigma_\nu \sigma_\nu \sigma_\mu \sigma_\mu + \sigma_\nu \sigma_\mu \sigma_\nu \sigma_\mu + \sigma_\nu \sigma_\mu \sigma_\mu \sigma_\nu)]. \quad (3.26)$$

The prefactor  $\frac{1}{2}$  in front of the second term takes the double counting of the different contributions into account. By defining

$$j(\sigma_\mu) \equiv \sum_{i=1}^n \delta_{\mu_i, 0}, \quad (3.27)$$

which counts the number of zeros in the index  $\mu$ , one finds

$$\mathbb{E}[\Delta T_\mu^2] = \frac{1}{3^{j(\sigma_\mu)} N}, \quad (3.28)$$

$$\mathbb{E}[\Delta T_\mu^4] = 3 \left( \frac{1}{3^{j(\sigma_\mu)} N} \right)^2 \quad (3.29)$$

for the second and the fourth moment of the Gaussian distributed random variable  $\Delta T_\mu$ . For more details on the Gaussian distribution of deviations of correlations, see [P2]. Consequently, Eq. (3.26) simplifies to

$$m_4^{\text{ev}} \approx \frac{1}{2^{5n}} \frac{1}{N^2} \sum_{\mu} \left[ \frac{3 \text{tr}(\sigma_\mu \sigma_\mu \sigma_\mu \sigma_\mu)}{(3^{j(\sigma_\mu)})^2} + \sum_{\nu; \mu \neq \nu} \frac{\text{tr}(\sigma_\mu \sigma_\mu \sigma_\nu \sigma_\nu + \sigma_\mu \sigma_\nu \sigma_\mu \sigma_\nu + \sigma_\mu \sigma_\nu \sigma_\nu \sigma_\mu + \sigma_\nu \sigma_\nu \sigma_\mu \sigma_\mu + \sigma_\nu \sigma_\mu \sigma_\nu \sigma_\mu + \sigma_\nu \sigma_\mu \sigma_\mu \sigma_\nu)}{2 \cdot 3^{j(\sigma_\mu)} \cdot 3^{j(\sigma_\nu)}} \right]. \quad (3.30)$$

The first expression depends solely on the index  $\mu$ , whereas the second one depends on  $\mu$  and  $\nu$ . Hence, the second expression occurs  $\mathcal{O}(4^n)$  times more often. As the magnitudes of the values are comparable, in the limit of many qubits, one can neglect the first term and is left with

$$m_4^{\text{ev}} \approx \frac{1}{2^{5n}} \frac{1}{N^2} \sum_{\mu, \nu; \mu \neq \nu} \frac{\text{tr}(\sigma_\mu \sigma_\mu \sigma_\nu \sigma_\nu + \sigma_\mu \sigma_\nu \sigma_\mu \sigma_\nu + \dots + \sigma_\nu \sigma_\mu \sigma_\nu \sigma_\mu + \sigma_\nu \sigma_\mu \sigma_\mu \sigma_\nu)}{2 \cdot 3^{j(\sigma_\mu)} \cdot 3^{j(\sigma_\nu)}}. \quad (3.31)$$

The crucial step is to realize that, while  $\sigma_\mu \sigma_\mu \sigma_\nu \sigma_\nu = \mathbb{1}$  holds for any choice of indices  $\{\mu, \nu\}$ ,  $\sigma_\mu \sigma_\nu \sigma_\mu \sigma_\nu$  can result in  $+\mathbb{1}$  or  $-\mathbb{1}$  depending on the commutation relation of  $\sigma_\mu$  and



FIGURE 3.7: Two of the six different permutations of  $\sigma_\mu\sigma_\nu\sigma_\mu\sigma_\nu$  are shown. The first one is called *crossing* as connecting lines between identical operators cross each other. Depending on the commutation relation of the operators, the argument of the trace equals  $\mathbb{1}$  or  $-\mathbb{1}$ . The other shown permutation, however, can directly be evaluated to  $\mathbb{1}$ , irrespective of the commutation relation of  $\sigma_\mu$  and  $\sigma_\nu$ .

$\sigma_\nu$ . One can find for every fixed index  $\mu$  two indices  $\nu$  and  $\nu'$  (with  $j(\sigma_\nu) = j(\sigma_{\nu'})$ ) such that  $\sigma_\mu\sigma_\nu = \sigma_\nu\sigma_\mu$  and  $\sigma_\mu\sigma_{\nu'} = -\sigma_{\nu'}\sigma_\mu$ . Hence, for a fixed  $\mu$ , only those summands within  $\text{tr}(\sigma_\mu\sigma_\mu\sigma_\nu\sigma_\nu + \sigma_\mu\sigma_\nu\sigma_\mu\sigma_\nu + \dots + \sigma_\nu\sigma_\mu\sigma_\nu\sigma_\mu + \sigma_\nu\sigma_\mu\sigma_\mu\sigma_\nu)$  have to be counted which do not depend on the commutation relation between the operators. This argument is illustrated in Fig. 3.7. Those summands which depend on these relations result in a positive and a negative value, respectively, half of the time, canceling each other. This simplifies Eq. (3.31) further and leads to

$$m_4^{\text{ev}} \approx \frac{1}{2^{5n}} \frac{1}{N^2} \sum_{\mu, \nu; \mu \neq \nu} \frac{\text{tr}(\sigma_\mu\sigma_\mu\sigma_\nu\sigma_\nu + \sigma_\mu\sigma_\nu\sigma_\nu\sigma_\mu + \sigma_\nu\sigma_\nu\sigma_\mu\sigma_\mu + \sigma_\nu\sigma_\mu\sigma_\mu\sigma_\nu)}{2 \cdot 3^{j(\sigma_\mu)} \cdot 3^{j(\sigma_\nu)}}. \quad (3.32)$$

Out of the six possible permutations of  $\sigma_\mu, \sigma_\mu, \sigma_\nu, \sigma_\nu$ , only four are left. Those four are also characterized by the fact that, when connecting mutually identical operators, no connecting lines are crossing as shown in Fig. 3.7. Please note, however, that each configuration is appearing twice, e.g.,  $\sigma_\mu\sigma_\mu\sigma_\nu\sigma_\nu$  and  $\sigma_\nu\sigma_\nu\sigma_\mu\sigma_\mu$  are identical up to relabeling of the indices  $\mu \leftrightarrow \nu$ .<sup>2</sup> Therefore, Eq. (3.32) can be simplified further by means of the Catalan number  $\mathcal{C}_2 = 2$ ,

$$m_4^{\text{ev}} \approx \frac{1}{2^{5n}} \frac{1}{N^2} \sum_{\mu, \nu; \mu \neq \nu} \frac{2^n \cdot 2 \cdot \mathcal{C}_2}{2 \cdot 3^{j(\sigma_\mu)} \cdot 3^{j(\sigma_\nu)}} \approx \frac{1}{2^{4n}} \frac{1}{N^2} \sum_{\mu, \nu} \frac{\mathcal{C}_2}{3^{j(\sigma_\mu)} \cdot 3^{j(\sigma_\nu)}}, \quad (3.33)$$

where the last expression neglects the restriction of  $\mu \neq \nu$  in the summation. Eventually, as the sums over  $\mu$  and  $\nu$  are now independent, this expression can be factorized and becomes

$$m_4^{\text{ev}} \approx \mathcal{C}_2 \left( \frac{1}{2^{2n}} \sum_{\mu} \frac{1}{3^{j(\sigma_\mu)} N} \right)^2. \quad (3.34)$$

<sup>2</sup>For higher moments, when  $k$  different indices are involved, the *relabeling factor* becomes  $k!$ , which cancels with the prefactor appearing in Eq. (3.26).

When comparing Eq. (3.34) with the second moment,

$$m_2^{\text{ev}} \approx \frac{1}{2^{2n}} \sum_{\mu} \frac{1}{3^{j(\sigma_{\mu})} N} \approx \frac{5^n}{6^n} \frac{1}{N}, \quad (3.35)$$

the fourth moment finally becomes

$$m_4^{\text{ev}} \approx \mathcal{C}_2 (m_2^{\text{ev}})^2. \quad (3.36)$$

Equivalently, all higher even moments can be calculated. In combination with the vanishing odd moments, it has thus been shown that the moments of the spectral distribution of a tomographically obtained maximally mixed state approximate the moments of an ideal semicircle distribution, Eq. (3.18). Furthermore, the radius can be determined from the second moment, leading to  $R = 2\sqrt{m_2^{\text{ev}}}$ . For more details, see appendix A of [P2].

### 3.4.4 Degenerate Eigenvalues

Although the maximally mixed state by itself is not a highly interesting state, it is a helpful model system to understand the effects due to finite statistics. The  $n$ -qubit maximally mixed state has  $2^n$  degenerate eigenvalues. The finite statistics will eventually lead to a lifting of the degeneracy of those eigenvalues such that they are distributed according to the Wigner semicircle distribution. This process of lifting the degeneracy, however, is not limited to the maximally mixed state and is also observed for other degenerate eigenvalues.

Often, experimentally relevant noise can be modeled via an admixture of the maximally mixed state, i.e., by adding white noise to a pure (or low rank) state. States described by the model  $p|\psi\rangle\langle\psi| + (1-p)\mathbb{1}_{\perp}/(2^n-1)$ , where  $\mathbb{1}_{\perp}$  denotes the maximally mixed state from the subspace perpendicular to  $|\psi\rangle$ , have a single eigenvalue  $p$  and  $2^n-1$  eigenvalues degenerate at  $(1-p)/(2^n-1)$ . In the same manner as discussed above, those degenerate eigenvalues tend to follow a Wigner semicircle distribution with a radius depending on the number of measurements performed.

This knowledge allows to discriminate *noise* eigenvalues from *signal* eigenvalues, estimate the essential rank of the state, and find a reasonable state estimate. Also, one can estimate before the actual measurement, how many measurements are required in order to avoid unphysical state estimates in the first place.

### 3.4.5 Error Estimation

The state estimators discussed in Sec. 3.2 are point estimators and do not provide any information about errors. In general, error analysis in quantum state estimation is a non-trivial task [108]. Using standard methods of error propagation commonly lead to a significant overestimation of statistical errors [108]. A very frequently used tool to obtain some information on the statistical scatter of the density matrix and quantities derived from it is the so-called *bootstrapping* method [174, 175]. This method uses the point estimate for the density matrix as a starting point to numerically simulate new density matrices

according to the statistical uncertainty of the experiment [176]. Using this set of density matrices, errors for the corresponding figure of merit such as the fidelity or the purity can be calculated. As bootstrapping involves a random process to obtain a set of density matrices close to the point estimate, it belongs to the class of *Monte Carlo* methods [108]. Besides the computational effort, this method implicitly relies on an unbiased estimate for the quantum state, which cannot be provided using, e.g., maximum likelihood estimation [176].

Other approaches allow, e.g., to obtain bounds for the fidelity depending on the number of measurements [177]. In Refs. [178–180], the likelihood function has been used to derive some confidence regions, whereas Refs. [181, 182] provide credibility regions as error regions. For some overview including a study of the computational complexity of optimal error regions, see [183].

Our ansatz of modeling the statistical uncertainty of the eigenvalues of the density matrix allows for a deep analysis of a tomographically obtained quantum state without relying on computationally demanding Monte Carlo methods or on biased state estimates. Using the knowledge of the distribution of the Wigner semicircle helps to identify noise eigenvalues, which are of no further interest as they are present solely due to finite statistics or they cannot be distinguished from this kind of noise. However, one is typically interested in the distribution of the other eigenvalues. For that purpose, the influence of finite statistics onto the estimated state can be examined using perturbation theory.

To illustrate the method, consider first an  $n$ -qubit state which has been shown to be of rank 1 by above mentioned methods, which are also described in Sec. 3.4.6, such that  $2^n - 1$  eigenvalues cannot be distinguished from the combination of white noise and effects of finite statistics. As the direct estimate of the state provides only a point estimate for this large eigenvalue  $\lambda_{\max} \equiv \lambda_{2^n}$ , the uncertainty is still unknown. In order to circumvent this shortcoming, statistical fluctuations to the density matrix are treated as small perturbations to the obtained eigendecomposition, i.e., the obtained density matrix with finite statistics is considered to be the result of the ideal density matrix together with a perturbation matrix. This perturbation matrix  $\Delta\rho$  can be decomposed as  $\Delta\rho = 1/2^n \sum_{\mu} \Delta T_{\mu} \sigma_{\mu}$  with a set of parameters  $\{\Delta T_{\mu}\}$  analog to the correlation tensor. Here,  $\Delta T_{\mu}$  are considered to be random variables following a normal distribution with expectation value 0 and variance  $[1 - T_{\mu}^2]/[3^{j(\mu)}N]$ , with  $N$  denoting the sample size and  $j$  defined as in Eq. (3.27). If now  $\Delta T_{\mu}$  is considered to be a small perturbation, one obtains for the first order deviation of the largest eigenvalue due to the perturbation

$$\begin{aligned} \Delta\lambda_{\max} &= \langle \Psi_{\max} | \Delta\rho | \Psi_{\max} \rangle = \text{tr} (|\Psi_{\max}\rangle \langle \Psi_{\max}| \Delta\rho) \\ &= \frac{1}{2^{2n}} \text{tr} \left( \sum_{\mu, \nu} \sigma_{\mu} \sigma_{\nu} \Delta T_{\mu} T_{\nu}^{\Psi_{\max}} \right) \\ &= \frac{1}{2^n} \sum_{\mu} \Delta T_{\mu} T_{\mu}^{\Psi_{\max}}, \end{aligned} \tag{3.37}$$

where  $T_{\nu}^{\Psi_{\max}}$  denotes the  $\nu$ -th correlation of the eigenstate corresponding to the largest eigenvalue, i.e.,  $T_{\nu}^{\Psi_{\max}} = \text{tr} (|\Psi_{\max}\rangle \langle \Psi_{\max}| \sigma_{\nu})$ .

This can be translated to the variance of the largest eigenvalue,

$$\begin{aligned}
\mathbb{V}[\lambda_{\max}] &= \mathbb{E}[(\Delta\lambda_{\max})^2] - \underbrace{\mathbb{E}[(\Delta\lambda_{\max})]^2}_{=0} \\
&= \frac{1}{2^{2n}} \sum_{\mu} \mathbb{E}[(\Delta T_{\mu})^2] \mathbb{E}[(T_{\mu}^{\Psi_{\max}})^2] \\
&= \frac{1}{4^n} \sum_{\mu} \frac{[1 - T_{\mu}^2]}{3^{j(\mu)}N} (T_{\mu}^{\Psi_{\max}})^2 \\
&\approx \frac{1}{4^n} \sum_{\mu} \frac{[1 - (\lambda_{\max} T_{\mu}^{\Psi_{\max}})^2]}{3^{j(\mu)}N} (T_{\mu}^{\Psi_{\max}})^2.
\end{aligned} \tag{3.38}$$

In the last step, one can approximate the  $\mu$ -th correlation of the state,  $T_{\mu}$ , by a scaled value of the correlation of the eigenstate to the largest eigenvalue, i.e.,  $T_{\mu} \approx \lambda_{\max} T_{\mu}^{\Psi_{\max}}$ .

In this case of a rank-1 state, the distribution of the largest eigenvalue is directly related to the distribution of the center of the noise eigenvalue, such that

$$\mathbb{V}[c] \approx \left(\frac{1}{2^n - 1}\right)^2 \mathbb{V}[\lambda_{\max}] \tag{3.39}$$

holds.

However, for cases with higher rank, i.e.,  $r > 1$ , the calculation becomes more involved. In Eq. (3.38), one distinguished between the one eigenstate containing information and the other eigenstates. Now, instead of a single eigenstate of interest, several have to be taken into account, resulting in the variance of the  $i$ -th relevant eigenvalue of

$$\mathbb{V}[\lambda_i] \approx \frac{1}{4^n} \sum_{\mu} \frac{[1 - \sum_l (\lambda_l T_{\mu}^{\Psi_l})^2]}{3^{j(\mu)}N} (T_{\mu}^{\Psi_i})^2. \tag{3.40}$$

All eigenvalues which are identified to be of relevance are incorporated in the summation  $\sum_l (\lambda_l T_{\mu}^{\Psi_l})^2$ , where, analogously to above, the  $\mu$ -th correlation of the  $l$ -th eigenstate is rescaled with the corresponding eigenvalue. This allows to directly propagate the uncertainty also to other quantities such as the purity.

Furthermore, one can now compare which deviations in the estimated state are caused by finite statistics when using the direct method of linear inversion and a scheme, which distinguishes noise eigenvalues from relevant ones. The impact of finite statistics can be estimated in terms of the Hilbert-Schmidt distance, which is defined by  $h(\Delta\rho) = \sqrt{1/2^n} \sqrt{\sum_{j,k} |\Delta\rho_{j,k}|^2} = \sqrt{1/2^n} \sqrt{\sum_{\mu} (\Delta T_{\mu})^2}$  for  $\Delta\rho$ , the difference of two states, parametrized by means of  $\Delta T_{\mu}$ . As in Sec. 3.4.3,  $\Delta\rho = \tilde{\rho} - \rho$  is defined as the deviation of the estimated state  $\tilde{\rho}$  from the original state  $\rho$ . The Hilbert-Schmidt distance will be used to estimate how large statistical errors in the state estimate will be. In fact, the squared

Hilbert-Schmidt distance corresponds to the mean squared error (MSE) of the estimator. For the linear inversion, the mean value and variance of the squared Hilbert-Schmidt norm are obtained by means of

$$\mu_{\text{MSE,lin}} = \mathbb{E} [h(\Delta\rho)^2] = \frac{1}{2^n} \sum_{\mu} \mathbb{E} [(\Delta T_{\mu})^2], \quad (3.41a)$$

$$\sigma_{\text{MSE,lin}}^2 = \frac{1}{4^n} \sum_{\mu,\nu} \mathbb{E} [(\Delta T_{\mu})^2 (\Delta T_{\nu})^2] - \mu_{\text{MSE,lin}}^2, \quad (3.41b)$$

$$\mu_{\text{HS,lin}} \approx \sqrt{\mu_{\text{MSE,lin}}} - \frac{1}{8} \frac{\sigma_{\text{MSE,lin}}^2}{\mu_{\text{MSE,lin}}^{3/2}}, \quad (3.41c)$$

$$\sigma_{\text{HS,lin}}^2 \approx \frac{\sigma_{\text{MSE,lin}}^2}{4\mu_{\text{MSE,lin}}}, \quad (3.41d)$$

where Eqs. (3.41c) and (3.41d) are computed by means of a Taylor expansion.

As an example, consider a random pure  $n = 6$  qubit state which is determined by means of tomography with  $N = 10\,000$  copies per basis setting. In the linearly obtained matrix, one large eigenvalue will be found and 63 eigenvalues will be distributed according to the Wigner semicircle distribution close to 0. The confidence region for linearly reconstructed states around this estimated state is given by the Hilbert-Schmidt distance, which can be approximated by a Gaussian distribution with mean  $\mu_{\text{HS,lin}} \approx 0.04593$  and standard deviation  $\sigma_{\text{HS,lin}} \approx 0.00713$ . The analytically obtained values are compared with 1 000 simulated tomographies around the reconstructed state, for which the mean value and standard deviation are 0.04594 and 0.00696, respectively, see Fig. 3.8. The distribution of the Hilbert-Schmidt norm, characterized by  $\mu_{\text{HS,lin}}$  and  $\sigma_{\text{HS,lin}}$ , is calculated as the distance from the linearly reconstructed state to other linearly reconstructed states.

If one uses the method proposed above and in [P2] (see also Sec. 3.4.6) to identify and discriminate signal eigenvalues from noise eigenvalues of the linearly reconstructed state and averages over the noise eigenvalues, the deviation of the state estimate from the theoretical state will be reduced, as shown by the distribution of the Hilbert-Schmidt norm in Fig. 3.8. After state reconstruction and estimation of eigenvalues, for this specific example the large eigenvalue is found to be  $\lambda_{64} = 0.99966$  with a standard deviation of 0.00071. For this reconstruction, the distribution of the Hilbert-Schmidt distance is described by the mean value  $\mu_{\text{MSE,recon}}$  and the standard deviation  $\sigma_{\text{MSE,recon}}$ , respectively, with

$$\mu_{\text{MSE,recon}} \approx \frac{4}{4^n} \sum_{\mu} \mathbb{E} [(\Delta T_{\mu})^2], \quad (3.42a)$$

$$\sigma_{\text{MSE,recon}}^2 \approx \frac{16}{16^n} \sum_{\mu,\nu} \mathbb{E} [(\Delta T_{\mu})^2 (\Delta T_{\nu})^2] - \mu_{\text{MSE,recon}}^2. \quad (3.42b)$$

One obtains for the approximation of the distribution of the Hilbert-Schmidt distance of reconstructed states the mean value  $\mu_{\text{MSE,recon}} \approx 0.01146$  and the standard deviation

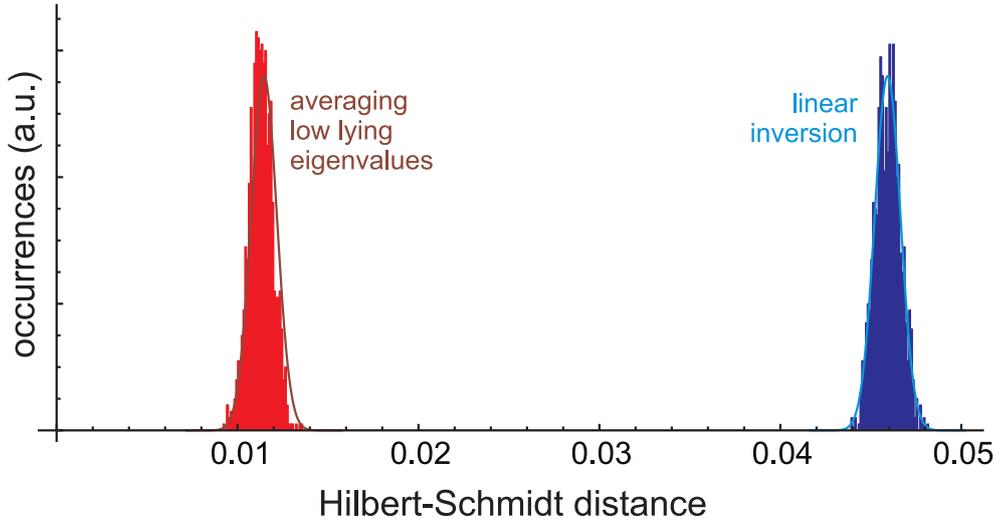


FIGURE 3.8: (Blue bars) Distribution of Hilbert-Schmidt norm in 1 000 numerical simulations of linear inversion compared with (blue line) Gaussian distribution with mean value and standard deviation as obtained from calculation based upon Eqs. (3.41c) and (3.41d). (Red bars) Distribution of Hilbert-Schmidt norm when using the reconstruction method proposed in [P2] and (red line) anticipated distributions based on Eqs. (3.42a) and (3.42b).

$\sigma_{\text{MSE, recon}} \approx 0.00071$ , respectively. Consequently, the state estimation is far less influenced by statistical noise than the linear inversion estimate as can be seen by the smaller Hilbert-Schmidt distances.

Other figures of merit can be either derived in an analogous way or by means of numerical simulations based on the random matrices  $\Delta \varrho$ .

In the following section, which is a reprint of [P2], these methods are discussed in detail and also applied for the evaluation of an experimentally prepared quantum state. The origin of the Wigner semicircle distribution for estimates of state tomography with finite statistics is explained together with tools for an in-depth analysis of tomography results. Using hypothesis tests, one can easily assess the actual rank of the state, although the tomographically obtained estimate may seem to be of approximately half rank. Furthermore, this theory allows to estimate the necessary tomographic effort for obtaining statistically meaningful results.

This publication arose from discussion with Tobias Moroder, Matthias Kleinmann and Otfried Guhne as well as with Geza Toth. Based on discussions with Harald Weinfurter, I performed the numerical simulations for this project and evaluated the data. Together with Nico Klein and Christian Schwemmer, we interpreted the results and worked on the spectral probability distribution for a single qubit. Jonas Reuter provided an initial proof that the spectral distribution will converge to a semicircle distribution, still neglecting some conditions of the tomographic setup. Later on, I could replace the proof by a very different approach which is specifically tailored for quantum state tomography. The experimental data used in the publication had been obtained by Christian Schwemmer in the course

of an earlier publication. The manuscript was mainly written by myself together with Christian Schwemmer, Géza Tóth, and Harald Weinfurter and edited by all authors. An earlier version of this manuscript was entitled “How long does it take to obtain a physical density matrix?” (arXiv:1512.06866).

## Multiqubit State Tomography with Finite Data

Lukas Knips,<sup>1,2</sup> Christian Schwemmer,<sup>1,2,\*</sup> Nico Klein,<sup>1,2</sup> Jonas Reuter,<sup>3</sup> Géza Tóth,<sup>4,5,6</sup> and Harald Weinfurter<sup>1,2</sup>

<sup>1</sup>*Max-Planck-Institut für Quantenoptik, Hans-Kopfermann-Strasse 1, D-85748 Garching, Germany*

<sup>2</sup>*Department für Physik, Ludwig-Maximilians-Universität, D-80797 München, Germany*

<sup>3</sup>*Bethe Center for Theoretical Physics, Universität Bonn, D-53115 Bonn, Germany*

<sup>4</sup>*Department of Theoretical Physics, University of the Basque Country UPV/EHU, P.O. Box 644, E-48080 Bilbao, Spain*

<sup>5</sup>*IKERBASQUE, Basque Foundation for Science, E-48013 Bilbao, Spain*

<sup>6</sup>*Wigner Research Centre for Physics, Hungarian Academy of Sciences, P.O. Box 49, H-1525 Budapest, Hungary*

We show that for finite set of data the statistical nature of measurements is an almost unavoidable reason for unphysical estimates in multiqubit quantum state tomography using linear inversion. The usual multinomial or Poissonian noise results in distributions of eigenvalues converging to the Wigner semicircle distribution for already a modest number of qubits. This fact has to be taken into account by all methods evaluating tomographic data. Here, we show how to determine which eigenvalues of the raw density matrix obtained via tomography are relevant and which ones are irrelevant as they are the result of statistical effects or cannot be distinguished from it. We introduce a method to avoid unphysical estimates to a great extend, without using constrained optimization, thus avoiding numerical artefacts such as the bias. This approach allows to directly obtain also error bars for interesting figures of merit such as the fidelity with minimal numerical effort.

PACS numbers: 03.65.Wj,03.67.Ac

### I. INTRODUCTION

Quantum state tomography (QST) became the standard tool for fully determining unknown multiqubit states [1, 2] and has been used to analyze experimental data for numerous systems like photons [3–5], ions [6], nuclear magnetic resonance systems [7], superconducting qubits [8], and cold gases [9].

In experiments, however, using straightforward linear inversion, i.e., solving a system of linear equations based on the measured data for the density matrix, typically, one obtains unphysical “raw estimates” [10]. Powerful numerical methods like the maximum likelihood estimation (ML), least squares (LS), or Bayesian concepts [4, 11] have been employed to cope with this fact and to enforce a physical result. Although reconstruction methods made great progress [12], the tremendous computational requirements limit these approaches for full QST. Recent developments enable to obtain confidence and credibility regions, respectively, instead of point estimates [13], yet, are confined to very few qubits. An even more pressing problem is that the methods using constrained optimization are precise only in the limit of an infinite number of copies, since, for real, finite data the constraint of physicality in optimization routines results in artefacts. For example, puzzling convergence behaviour has been reported [14] as well as biased estimation leading to distorted results, systematic errors and wrong information about the statistical uncertainty of the estimated parameters [15–17].

Here, analyzing the impact of statistical noise for a finite amount of data we find a characteristic distribu-

tion for the eigenvalues of matrices obtained via linear inversion. For Poissonian or multinomial measurement statistics and the standard overcomplete Pauli tomography scheme, the obtained eigenvalues can be described by a Wigner semicircle distribution [18, 19] already for a handful of qubits. The knowledge of the eigenvalue distribution enables one to identify statistically relevant components of a raw estimate and thus to put the analysis onto physical grounds. We propose a simple, but plausible physical model composed by the statistically relevant eigenvalues and with confidence regions obtained using perturbation theory.

Our approach compares well with standard data analysis methods in various areas of physics. Let us illustrate our general approach with the analogy of interferometry. If one analyzes an interference pattern knowing that data points will scatter due to statistics, one surely restrains from attributing physical significance to all the amplitudes of a Fourier transform, but rather keeps the most significant ones to model the data. The residues are then tested for consistency, equivalently to the hypothesis test proposed here.

### II. QUANTUM STATE ESTIMATION

There are several strategies of how to collect data for QST. The measurement scheme we first focus on is the so-called Pauli tomography scheme (other schemes will be discussed in Sec. III C). For characterizing an  $n$ -qubit system, the individual qubits are projected on the eigenstates of all  $3^n$  possible tensor products of local Pauli operators,  $\sigma_1$ ,  $\sigma_2$ , and  $\sigma_3$ . These  $3^n$  different measurement settings are denoted by  $\mathbf{s} \in \{(s_1, s_2, \dots, s_n) | s_j \in \{1, 2, 3\}, j = 1, \dots, n\}$  and correspond to measurements of  $\sigma_{s_1} \otimes \sigma_{s_2} \otimes \dots \otimes \sigma_{s_n}$ . For each setting  $\mathbf{s}$  one ob-

\* Current Address: IBM Research GmbH, Säumerstrasse 4, CH-8803 Rüschlikon, Switzerland.

tains  $2^n$  different outcomes,  $\mathbf{r} \in \{(r_1, r_2, \dots, r_n) | r_j \in \{-1, 1\}, j = 1, \dots, n\}$ . This leads to a tomographically overcomplete set of  $6^n$  relative frequencies  $f_{\mathbf{r}}^{\mathbf{s}} = c_{\mathbf{r}}^{\mathbf{s}}/N_{\mathbf{s}}$  with  $c_{\mathbf{r}}^{\mathbf{s}}$  the number of results  $\mathbf{r}$  for settings  $\mathbf{s}$  and with total number of counts  $N_{\mathbf{s}} = \sum_{\mathbf{r}} c_{\mathbf{r}}^{\mathbf{s}}$  of setting  $\mathbf{s}$ . We use the same total number of counts for all settings, i.e.,  $N_{\mathbf{s}} = N$  for all  $\mathbf{s}$ .

To determine the raw estimate  $\tilde{\varrho}$  of the density matrix  $\varrho$ , the decomposition

$$\varrho = \frac{1}{2^n} \sum_{\boldsymbol{\mu}} T_{\boldsymbol{\mu}} \sigma_{\boldsymbol{\mu}} \quad (1)$$

can be used, where now  $\sigma_{\boldsymbol{\mu}} = \sigma_{\mu_1} \otimes \sigma_{\mu_2} \otimes \dots \otimes \sigma_{\mu_n}$ ,  $\mu_i \in \{0, 1, 2, 3\}$ , includes also  $\sigma_0$ , the identity matrix. The correlation tensor, defined via  $T_{\boldsymbol{\mu}} = \langle \sigma_{\boldsymbol{\mu}} \rangle$ , includes both the so-called *full* correlations, where  $\mu_k \neq 0$  for all  $k$ , such that  $\boldsymbol{\mu} = \mathbf{s}$ , but also the *non-full* correlations with some  $\mu_k = 0$ , i.e., which describe correlations between only those qubits with  $\mu_k \neq 0$ .

In a measurement, the estimates  $\tilde{T}_{\boldsymbol{\mu}}$  are obtained from the respective frequencies from which  $\tilde{\varrho}$  is derived using Eq. (1) [20]. Note that the raw estimate  $\tilde{\varrho}$  is defined on the set of all Hermitian  $2^n \times 2^n$  matrices without the constraint of positive semi-definiteness. It is used here solely as a basis for calculating a physical state estimate  $\hat{\varrho}$ .

In practice, due to the finite number of measurements and the resulting statistical noise, the  $\tilde{T}_{\boldsymbol{\mu}}$  and consequently the elements of  $\tilde{\varrho}$  are random variables depending on both the observed state and the statistics. For a fixed number of measurements, the events  $c_{\mathbf{r}}^{\mathbf{s}}$  are multinomially distributed random values. If instead of a fixed number of measurements, a fixed measurement time per setting is chosen, the number of events follows a Poissonian distribution. In any case, when increasing the sample sizes  $N$ , the distribution of the correlation  $\tilde{T}_{\boldsymbol{\mu}}$  tends to a Gaussian distribution, which in the following will be denoted  $g(\tilde{T}_{\boldsymbol{\mu}})$ .

Note, in the overcomplete Pauli scheme, the different correlations  $\tilde{T}_{\boldsymbol{\mu}}$  and also the different elements of  $\tilde{\varrho}$  do not have all the same uncertainty. The correlation value of a full correlation  $\tilde{T}_{\mathbf{s}}$  is approximately distributed according to a Gaussian distribution with mean  $T_{\mathbf{s}}$ , i.e., the theoretically expected correlation, and with variance  $(1 - T_{\mathbf{s}}^2)/N$ . However, due to the inherent redundancy, all  $3 \times 2^n$  frequencies  $f_{\mathbf{r}}^{\mathbf{s}}$  can be used to calculate the non-full correlation  $\tilde{T}_{\boldsymbol{\mu}}$ . For example, for  $\mu_k = 0$ ,  $\tilde{T}_{\boldsymbol{\mu}}$  is obtained from three settings  $\mathbf{s} = (s_1, \dots, s_{k-1}, s_k, s_{k+1}, \dots, s_n)$  with  $s_k \in \{1, 2, 3\}$ . Thus, for those non-full correlations, more events are taken into account, clearly reducing its variance. If  $\sigma_{\boldsymbol{\mu}}$  contains  $j(\boldsymbol{\mu}) = \sum_{i=1}^n \delta_{\mu_i, 0}$  times the factor  $\sigma_0$  ( $\delta_{a,b}$  denotes the Kronecker delta), the distribution  $g(\tilde{T}_{\boldsymbol{\mu}})$  will be narrower compared to the one for full correlations, as  $3^{j(\boldsymbol{\mu})}N$  events instead of  $N$  are available for the evaluation. In the following, often the short-hand notation  $j$

will be used instead of  $j(\boldsymbol{\mu})$ . Altogether, we can approximate the distribution of the correlations by a Gaussian given by

$$g(\tilde{T}_{\boldsymbol{\mu}}) = \sqrt{\frac{3^{j(\boldsymbol{\mu})}N}{(1 - T_{\boldsymbol{\mu}}^2)2\pi}} \exp\left(-\frac{3^{j(\boldsymbol{\mu})}N(\Delta T_{\boldsymbol{\mu}})^2}{2(1 - T_{\boldsymbol{\mu}}^2)}\right) \quad (2)$$

with the variance  $\mathbb{V}[g(\tilde{T}_{\boldsymbol{\mu}})] = (1 - T_{\boldsymbol{\mu}}^2)/(3^{j(\boldsymbol{\mu})}N)$  and with  $\Delta T_{\boldsymbol{\mu}} = (\tilde{T}_{\boldsymbol{\mu}} - T_{\boldsymbol{\mu}})$ . The distribution of correlations can be approximated by the Gaussian of Eq. 2 as long as  $\mathbb{V}[g(\tilde{T}_{\boldsymbol{\mu}})] \ll 1 - T_{\boldsymbol{\mu}}$ . As a first fact, to note, for any data analysis the different variances of the observed correlations  $\tilde{T}_{\boldsymbol{\mu}}$  have to be taken into account.

### III. DISTRIBUTION OF EIGENVALUES

Due to finite statistics, the eigenvalues of the raw estimate  $\tilde{\varrho}$  will be randomly distributed around the true values. To understand these distributions we start with the completely mixed state (Sec. III A) as this state is, on one hand, least prone to result in unphysical estimates and, on the other hand, can analyze the effect of finite statistics to highly degenerate eigenvalues. Low rank states will be discussed in Sec. III D. We compare the derived distribution with the ones obtained from numerical simulations.

#### A. Maximally Mixed State with Many Qubits

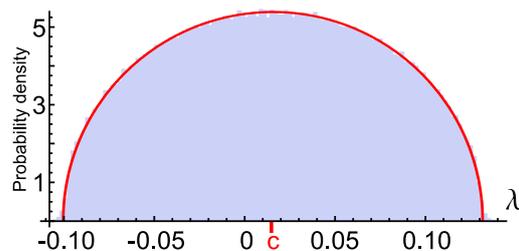


FIG. 1. The relative occurrences of eigenvalues for the  $n = 6$  qubit completely mixed state. 10 000 QST and linear state estimations have been performed on simulated data with  $N = 100$  expected counts per measurement basis yielding a characteristic eigenvalue distribution (blue). (red line) The Wigner semicircle centered at  $c = 2^{-6}$  with a radius of  $R = 0.116$ , see Eqs. (3) and (6). (The probability density of the histogram is obtained here by dividing probabilities by bin size.)

The completely mixed state  $\varrho_{\text{cm}}$  is characterized by  $T_{\boldsymbol{\mu}} = 0$  for all  $\boldsymbol{\mu}$  except  $T_{0, \dots, 0} = 1$ , hence  $\varrho_{\text{cm}} = \frac{1}{2^n} \sigma_{0, 0, \dots, 0} = \frac{1}{2^n} \mathbb{1}$  with  $2^n$ -fold degenerate eigenvalue  $1/2^n$ . However, the statistical fluctuations of the mea-

surement results  $c_r^s$  and thus of  $\tilde{T}_\mu$  lift the degeneracy, in fact, they cause repulsion of the eigenvalues [18], and consequently cause a wide distribution of observed eigenvalues. With this we come to another central result: provided the distribution of the correlations can be approximated by a Gaussian distribution with this specific pattern of variances as given in Eq. (2), in the limit of large matrices the eigenvalues follow a Wigner semicircle distribution [18, 19]. This is surprising since the different matrix elements have considerably different variances, which is different from the scenario usually assumed for random matrices, where all elements are identically distributed [21]. A proof of this central fact is given in App. A.

As shown in Fig. 1 we observe in a numerical simulation already for a modest number of  $n = 6$  qubits for the distribution of the eigenvalues of the completely mixed state an excellent agreement with the semicircle distribution. The center  $c$  of the Wigner semicircle distribution  $W(\lambda) \propto \sqrt{R^2 - (\lambda - c)^2}$  is given by the mean value of the eigenvalues  $\lambda_i$  and is, due to normalization of  $\tilde{g}$ , equal to

$$c = \frac{1}{2^n} \mathbb{E} \left[ \sum_{i=1}^{2^n} \lambda_i \right] = \frac{1}{2^n}, \quad (3)$$

while the radius  $R$  depends on the second moment of the correlation tensor elements,  $\mathbb{E}(\tilde{T}_\mu^2)$ , via the relation

$$\left(\frac{R}{2}\right)^2 = \frac{1}{4^n} \sum_{\mu} \mathbb{E}(\tilde{T}_\mu^2). \quad (4)$$

For the completely mixed state and multinomial noise, the correlation  $\tilde{T}_\mu$  will be distributed according to Eq. (2) as  $g(\tilde{T}_\mu) \approx \sqrt{3^{j(\mu)}N/(2\pi)} \exp(-\tilde{T}_\mu^2 3^{j(\mu)}N/2)$ . Consequently, we obtain for the second moment  $\mathbb{E}(\tilde{T}_\mu^2) = \int_{-\infty}^{\infty} d\tilde{T}_\mu g(\tilde{T}_\mu) \tilde{T}_\mu^2 \approx \frac{1}{3^{j(\mu)}N}$ . and thus for the right-hand side of Eq. (4)

$$\begin{aligned} \frac{1}{4^n} \sum_{\mu} \mathbb{E}(\tilde{T}_\mu^2) &= \frac{1}{4^n} \sum_{j'=0}^{n-1} \sum_{j(\mu)=j'} \frac{1}{3^{j(\mu)}N} = \\ \frac{1}{4^n} \left[ \frac{3^n}{N} + \frac{3^{n-1}}{3N} \binom{n}{1} + \dots + \frac{3}{3^{n-1}N} \binom{n}{n-1} \right] &= \\ \frac{1}{4^n} \sum_{j'=0}^{n-1} 3^{n-j'} \binom{n}{j'} \frac{1}{3^{j'}N} &= \frac{10^n - 1}{12^n N}. \end{aligned} \quad (5)$$

Here, the second line follows from counting the number of correlations with  $j'$  indices being 0. The first contribution in the sum considers all  $3^n$  full correlations. The second summand is due to the non-full correlations for  $n-1$  observers: each observers can measure in 3 different settings ( $3^{n-1}$ ) and the binomial counts the possibilities

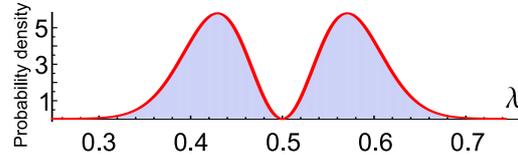


FIG. 2. The spectral probability distribution for the single qubit completely mixed state. QST was simulated  $10^6$  times, each with  $N = 100$  events per measurement setting. (blue area) Histogram of the eigenvalues with high-resolution bins. (red line) Expected distribution.

of  $n-1$  out of  $n$  possible observers. Hence, the radius  $R$  of the Wigner semicircle distribution for the approximation of the  $n$ -qubit completely mixed state is given by

$$R = 2\sqrt{\frac{10^n - 1}{12^n}} \frac{1}{\sqrt{N}} \approx 2\left(\frac{5}{6}\right)^{\frac{n}{2}} \frac{1}{\sqrt{N}}. \quad (6)$$

Note, the radius of the semicircle approximation decreases with the statistics  $N$  as  $\sqrt{1/N}$ . For the simulated data, we can now estimate the radius to be  $R = 0.116$ , giving the clear agreement with the distribution of eigenvalues visible in Fig. 1. Obviously, if the radius  $R$  is larger than the center  $c$ , i.e., the distribution extends to the negative range, there is a non-vanishing probability for negative eigenvalues (see App. C2).

### B. Maximally Mixed State with Few Qubits

The Wigner semicircle describes the spectral distribution in the limit of large matrices, i.e., of many qubits [18]. While already for six qubits the matrix is large enough such that the distribution of eigenvalues can be described very well by the Wigner semicircle distribution, let us also analyze the case of states with few qubits. For a single qubit, the eigenvalues and the spectral density can be determined analytically as  $g(\lambda) \propto \exp\left[-\frac{(1-2\lambda)^2 N}{2}\right] (1-2\lambda)^2$  (see App. B). Fig. 2 shows the simulated spectral probability of the  $n = 1$  qubit completely mixed state together with the probability density according to  $g(\lambda)$ . Evidently, it is highly unlikely to obtain the correct result with both eigenvalues at  $1/2$  since, due to  $\lambda_{1,2} = \frac{1}{2} \left(1 \pm \sqrt{T_1^2 + T_2^2 + T_3^2}\right)$ , all three correlations have to vanish. In particular, if  $N$  happens to be odd, it is even impossible. Yet, we see an exponential damping towards the boundaries of the distribution. Hence, for a reasonable number of measurements, the probability of negative eigenvalues is negligible.

The spectral distribution in the case of a single qubit clearly differs from a semicircular behavior. Fig. 3 shows that by increasing the number of qubits the distribution develops a comb structure. This effect is due to the repul-

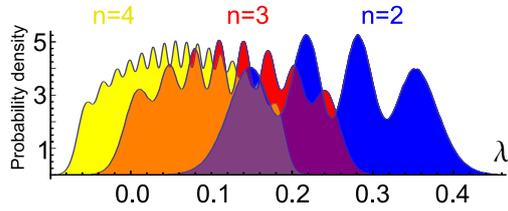


FIG. 3. Distribution of eigenvalues of simulated QSTs of the completely mixed state for  $N = 100$ . (blue) Results for  $n = 2$  qubits, displaying four peaks. The two-qubit state remains in almost all cases physical, i.e., all eigenvalues can be interpreted as probabilities. (red) For  $n = 3$  qubits, the distribution is shifted such that the probability for negative eigenvalues and thus for unphysical matrices increases strongly. (yellow) For  $n = 4$  qubits, the  $2^4$  subpeaks overlap strongly such that the comb structure is almost vanished and the shape of the Wigner semicircle function already becomes visible. One perceives the increased damping towards the boundaries of the distribution for increasing  $n$ .

sion of the eigenvalues, i.e., a reduced probability to find eigenvalues close to each other [21, 22]. The center of the distribution shifts, according to Eq. (3), closer to 0. As the probability density then extends to negative values, it becomes increasingly unlikely, even for the completely mixed state, to obtain physical results for a limited number of measurements per setting (here  $N = 100$ ). For  $n = 2$  qubits (blue histogram) only 6 out of  $10^6$  simulated states are unphysical, while for  $n = 3$  qubits, already 32% are unphysical. For  $n = 4$  all of the  $10^6$  simulated raw state estimates lack physicality. The tail of the distribution can be approximated by the Tracey-Widom distribution of the smallest and the largest eigenvalue of random matrices, respectively [23]. In our case already a Gaussian approximation is appropriate to estimate the likelihood for obtaining an unphysical matrix  $\tilde{\rho}$ , see App. C.

### C. Alternative Tomography Schemes

The overcomplete Pauli tomography discussed so far is one of several sampling strategies. Other schemes, for example, obtain the elements of the density matrix directly one by one by separate sets of measurements [24]. Tailoring the respective variances to be the same for all elements then will lead to a semicircle distribution for the eigenvalues (radius  $R = 2/\sqrt{N}$ , which is also discussed in the proof in App. A, see also [21, 25]). Similarly, also a symmetric sampling procedure [26] can yield such a distribution [27].

For the very frequently used informationally complete scheme proposed in [4], i.e., experimental data are obtained by projecting onto, e.g., all  $4^n$  tensor products of  $|0\rangle$ ,  $|1\rangle$ ,  $\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$ , and  $\frac{1}{\sqrt{2}}(|0\rangle + i|1\rangle)$ , the eigenvalue spectrum changes considerably. In this badly con-

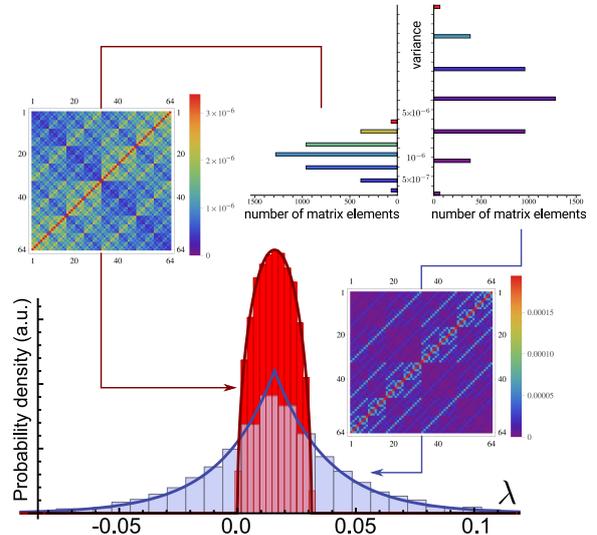


FIG. 4. (Histogram) Simulated spectrum of a 6 qubit completely mixed state obtained using the tomographically complete scheme [4] (blue bars) and the overcomplete scheme (red bars). The blue line corresponds to  $h(\lambda)$  given in Eq. (7). The dark red line corresponds to the semicircle model. While all eigenvalues in the latter scheme are expected to be positive, the asymmetric scheme [4] gives unphysical results in all cases. The matrix insets show the relative variances of the matrix elements encoded in adequately scaled colors. The pattern of the variances in the overcomplete scheme (left inset) lead to the semicircle distribution, as proven in the appendix. In the complete scheme (right inset), the antidiagonal entries of the density matrix strongly dominate, leading to a distorted eigenvalue distribution. While the variances of the diagonal entries are comparable for both schemes, the variances of the antidiagonal elements highly differ as visible from the paired histogram inset (upper right corner). Please note that the color codes are different for both methods for higher visibility.

ditioned [24] tomography scheme not all correlations are directly measured. Although the distribution of correlation values can still be approximated by Gaussian distributions, their variances are changed compared to Eq. (2). Consequently, the Catalan numbers (see App. A) cannot be reproduced by the moments of the distribution, leading to a different distribution than the Wigner semicircle distribution.

Fig. 4 compares complete and overcomplete sampling obtained from the spectral distribution of 100 simulation runs ( $n = 6$  qubits, completely mixed state, cf. Fig. 1). In each run, we choose the total number of measurements  $N_{\text{total}} = 4\,000\,000$ . Distributing all measurements over the required  $3^6$  settings for Pauli sampling ( $N = N_{\text{total}}/3^6 \approx 5487$ ) yields physical estimates with the eigenvalue distribution shown in red. Contrary to this, for complete sampling [4] with the same number of total measurements  $N_{\text{total}}$  one obtains the distribution shown

in blue. The second (centralized) moment of the heuristically found distribution  $h(\lambda) = \frac{\alpha}{2} \exp(-\alpha |\lambda - 2^{-n}|)$  is given by  $2/\alpha^2$  and has to be equal to the second (centralized) moment of the eigenvalue distribution, i.e.,  $4^n/N_{\text{total}}$ . Thus, the spectral probability distribution for the informationally complete tomography as proposed in Ref. [4] can be approximated by

$$h(\lambda) = \sqrt{\frac{N_{\text{total}}}{2 \cdot 4^n}} \exp\left(-\sqrt{\frac{2N_{\text{total}}}{4^n}} |\lambda - 2^{-n}|\right). \quad (7)$$

Evidently, this tomography scheme yields negative eigenvalues with high probability in the tails of the distribution far away from the center at  $1/64$ . For the dark blue line in Fig. 4, one obtains  $\alpha = \sqrt{2 \cdot 4 \cdot 10^6/4^6} \approx 44.2$ . Thus, another fact to note, it is highly recommended to always perform symmetric sampling as in the overcomplete scheme which yields physical estimates for a much lower number of measurements.

The problem of strongly differing variances is very likely to become even worse for sparse sampling methods such as, for example, compressed sensing [2, 28, 29], which introduce further disturbances to the spectrum due to its unavoidable fitting. These schemes and badly conditioned ones, such as photon number state reconstruction schemes [30] require a separate, detailed analysis.

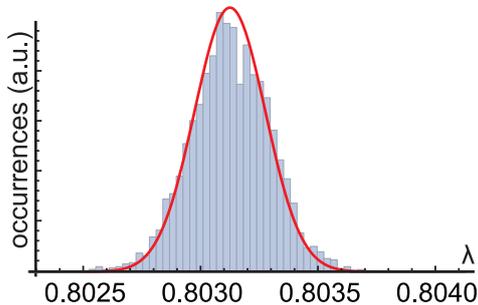


FIG. 5. (Bars) The large eigenvalue of a noisy 6 qubit GHZ state follows (red line) a Gaussian distribution with mean  $0.8 + 0.2/64 = 0.803125$  and standard deviation  $1.5 \cdot 10^{-4}$  estimated from Eq. (8). A more detailed discussion of this state is given in App. C and App. E, see also Fig. 10.

#### D. Low Rank States

Using first order perturbation theory, the variance of a large eigenvalue  $\lambda_i$  can be obtained as (App. E)

$$\text{V}[\lambda_i] \approx \frac{1}{4^n} \sum_{\mu} \frac{\left[1 - \sum_l (\lambda_l T_{\mu}^{\Psi_l})^2\right] (T_{\mu}^{\Psi_i})^2}{3^{j(\mu)} N}, \quad (8)$$

where the summation  $l$  is over the contributions from the  $r$  relevant (large) eigenvalues and  $T_{\mu}^{\Psi_l}$  denotes cor-

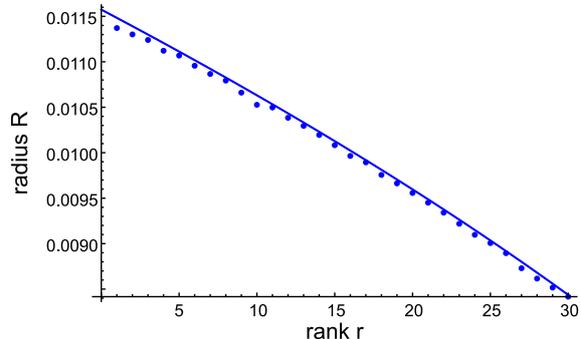


FIG. 6. QST (with  $N = 10000$ ) is simulated for the equal mixture of  $r$  pure (random)  $n = 6$  qubit states. The numerically obtained radii  $R$  for different values for the rank  $r$  are compared with the expected radius behaviour. For more details, see Sec. III D.

relations of the  $r$  eigenstates [36]. For the example of noisy 6 qubit Greenberger-Horne-Zeilinger state with  $\varrho = 0.8|\text{GHZ}\rangle\langle\text{GHZ}| + \frac{0.2}{64}\mathbb{1}$ , the large eigenvalue is expected to be  $\lambda_{64} = 0.8 + 0.2/64 \approx 0.803$ . However, due to finite statistics, the estimated eigenvalue will follow a Gaussian distribution with mean  $0.8 + 0.2/64$  and with variance according to Eq. (8). Fig. 5 shows a comparison between this approximate model and a numerical simulation.

In experiments, one typically aims to realize a specific quantum state, yet, the state preparation and analysis are never perfect. We thus make the ansatz that a low rank state  $\hat{\varrho}_r$  (rank  $r$ ) to be observed is mixed with white noise as  $\hat{\varrho} = q\hat{\varrho}_r + (1-q)\varrho_{\text{cm}} = f\hat{\varrho}_r + \frac{1-f}{2^n-r}\varrho_{\text{cm}}^{\perp}$ , where  $\varrho_{\text{cm}}^{\perp}$  denotes white noise in the subspace orthogonal to  $\hat{\varrho}_r$ . We use white noise often encountered in experiments as a first ansatz, which will be checked later with a hypothesis test.

For rank one ( $r = 1$ ,  $\varrho_r = |\psi\rangle\langle\psi|$ ), the probability of observing  $|\psi\rangle$  is  $q + (1-q)/2^n$ , which turns out to be the largest eigenvalue of  $\varrho$ . The other eigenvalues are degenerate, but, for finite statistics, due to level repulsion they follow again a semicircle distribution now centered at  $c_{q,r} = (1-q)/(2^n - r)$ , which is significantly closer to 0 or at 0 if there would not be any contribution from the orthogonal subspace [31, 32]. This also leads to a small reduction of the radius by approximately a factor of  $\sqrt{(2^n - r)/2^n}$  as

$$R \rightarrow R_r = R \cdot \sqrt{1 - r \cdot 2^{-n}} = 2 \left(\frac{5}{6}\right)^{\frac{\alpha}{2}} \frac{\sqrt{1 - r \cdot 2^{-n}}}{\sqrt{N}}, \quad (9)$$

as shown in Fig. 6.

Knowing the distribution of eigenvalues enables now to estimate the necessary amount of measurements such that the data directly result in a physical density ma-

trix (App. C). A physical solution is expected for  $R_r \leq c_{q,r}$ , thus the number of measurements necessary to obtain a physical state with high probability is  $N_0 = 4 \left(\frac{5}{6}\right)^n \left(\frac{2^n - 1}{1 - q}\right)^2$ . Note that  $N_0$  grows exponentially with the number of qubits and is inversely proportional to  $(1 - q)^2$ , i.e., to the squared infidelity. For high fidelity states, it is thus almost impossible to obtain a physical estimate without further data evaluation.

#### IV. PHYSICALLY MOTIVATED STATE ESTIMATION

##### A. Scheme

We have now the tools at hand to devise a new strategy to obtain a physical state estimate without constrained optimization. The strategy is to use clearly identifiable eigenstates and values of  $\tilde{\rho}$ . Provided the remaining eigenvalue spectrum agrees with the expected Wigner distribution, one should replace those eigenvalues by white noise. For a (noisy) low rank state sampled with finite statistics, we expect  $r$  eigenvalues and the  $2^n - r$  ideally degenerate eigenvalues to scatter within an interval of width  $w_r = 2 \cdot R_r$ . Knowing the width of this scatter enables to identify the rank of  $\varrho_r$  [33]. The proposed procedure is to (i) find  $r$  using  $|\{\lambda_j | \lambda_j > \min_i \lambda_i + w_r\}| = r$ , i.e., find the number of eigenvalues larger than  $\min_i \lambda_i + w_r$  (for an initial estimate, use the number of eigenvalues larger  $\min_i \lambda_i + 2R$ ), (ii) compose  $\varrho_r$  of  $r$  eigenstates with eigenvalues larger than  $\min_i \lambda_i + w_r$ , (iii) replace all  $2^n - r$  other eigenvalues by their center value  $c_r = \mathbb{E}(\{\lambda_j | \lambda_j < \min_i \lambda_i + w_r\})$ , (iv) estimate the state  $\varrho$  by averaging the lowest eigenvalues (ALE),  $\hat{\varrho}_{\text{ALE}} = \sum_{i=2^n - r + 1}^{2^n} \lambda_i |\psi_i\rangle \langle \psi_i| + \sum_{i=1}^{2^n - r} c_r |\psi_i\rangle \langle \psi_i|$ , and finally (v) check  $r$  as well as the form of the actual noise by a hypothesis test. As a cross check, repeat this procedure for neighboring values of  $r$ .

This scheme avoids overfitting [29] and has the advantage that no constraint, and thus no bias [15], influences the evaluation of the important eigenvalues [34]. For (almost) pure states all  $2^n - r$  eigenvalues will also be distributed according to a Wigner semicircle distribution with the respective radius.

Please note that approaches such as maximum likelihood estimation do not modify states which are physical in the first place. However, already in these situations, one will observe that the small eigenvalues are distributed according to a Wigner semicircle distribution due to the finite statistics. One cannot contribute any meaning to those eigenvalues.

##### B. Experimental Demonstration

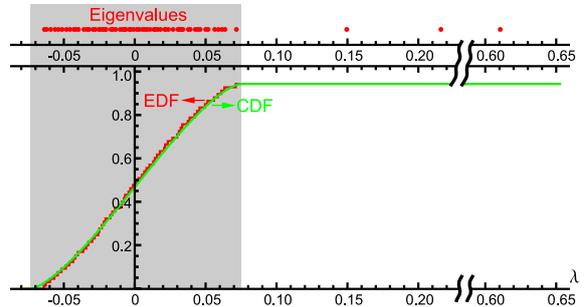


FIG. 7. Eigenvalue distribution for a tomography of a six-qubit Dicke state. (red dots) Eigenvalues and (gray shaded area) the support of the assumed Wigner semicircle distribution. (red stepped curve) The empirical distribution function (EDF) of the smallest 61 eigenvalues and the (green smooth curve) cumulative distribution function (CDF) of the Wigner semicircle distribution with  $R = 0.0745$  and  $c = 4 \times 10^{-4}$  in excellent agreement.

TABLE I. Center and radius of a hypothetical Wigner semicircle distribution for different rank values  $r$ . The  $P$ -values directly obtained by means of an Anderson-Darling (AD) hypothesis test are shown as well as *effective*  $P_{\text{eff}}$ -values. If more than  $r$  eigenvalues are found outside of the Wigner semicircle,  $P_{\text{eff}}$ -values is set to 0. For the hypothesis of the state being the maximally mixed state, one obtains the center  $c = 1/64$  with a radius of  $R = 0.076317$ . The  $P$ -value for this hypothesis is below  $10^{-5}$ .

rank $r$	center $c$	radius $R$	$P$ -value	$P_{\text{eff}}$ -value
1	0.006187	0.075719	0.0089	0
2	0.002803	0.075115	0.3553	0
3	0.000399	0.074507	$1 - 8 \cdot 10^{-7}$	$1 - 8 \cdot 10^{-7}$
4	-0.000790	0.073894	0.9998	0.9998
5	-0.001883	0.073275	0.9976	0.9976

Let us demonstrate the power of this scheme on real data of an overcomplete quantum state tomography of a 6 qubit Dicke state obtained in a 6 photon experiment [3]. On average,  $N = 230$  projection measurements per setting were performed, leading to  $R_r \approx 0.07632 \cdot \sqrt{1 - r \cdot 2^{-n}}$ . The 64 ordered eigenvalues of the unphysical raw matrix are found to be  $\{\lambda_i\} = \{-0.06368, -0.06223, \dots, 0.06371, 0.07171, 0.14949, 0.21595, 0.61024\}$  indicated by red dots in Fig. 7. First, the rank of  $\varrho_r$  is set to  $r = 3$ , and thus, the center is obtained as the average of the smallest 61 eigenvalues or, respectively, by  $c = \mathbb{E}(\{\lambda_i | i \leq 61\}) = (1 - \lambda_{64} - \lambda_{63} - \lambda_{62})/61 \approx 4 \times 10^{-4}$ . For  $R_{r=3} = 0.07452$  all the 61 lowest eigenvalues are within the given interval, hence  $\varrho_{r=3}$  now defines the estimated state. Using the Anderson-Darling hypothesis test [35], the assumption of a state with rank  $r = 3$  results in a  $P$ -value of about  $1 - 8.2 \cdot 10^{-7}$ , well above any prechosen significance level

(typically 0.05). For lower values of  $r$ , the  $P$ -value is either below the significance level or some eigenvalues are neither attributed to the signal nor the noise, thus ruling out this value for  $r$ , see Tab. I. Since one is looking for the simplest model explaining the data, one identifies the experimentally prepared state as a rank-3 state admixed with some noise. The next-to-perfect agreement between empirical and cumulative distribution functions (see Fig. 7) shows that the observed eigenvalue distribution cannot be distinguished from the eigenvalue distribution of white noise, thereby justifying to replace all  $\lambda_j$  ( $j \leq 2^n - r$ ) with  $c_r$ , i.e., to complement  $\varrho_r$  by  $\varrho_{\text{cm}}^\perp$ , white noise in the subspace orthogonal to  $\varrho_r$ , see App. D. The variance of the large eigenvalue  $\lambda_{64}$  is estimated by applying perturbation theory according to Eq. (8) and App. E. One obtains for the large eigenvalue and for the fidelity  $\lambda_{64} = 0.610 \pm 0.004$  and  $F(|D_6^{(3)}\rangle, \hat{\varrho}_{\text{ALE}}) = \langle D_6^{(3)} | \hat{\varrho}_{\text{ALE}} | D_6^{(3)} \rangle = 0.602 \pm 0.004$ , respectively, with 95% confidence level. Indeed, the remaining two large eigenvalues  $\lambda_{62} = 0.14949$  and  $\lambda_{63} = 0.21595$  can be identified as higher order noise due to the experimental preparation. The corresponding eigenstates correspond to the Dicke states  $|D_6^{(2)}\rangle$  and  $|D_6^{(4)}\rangle$ , respectively [3].

## V. CONCLUSIONS

For finite measurement statistics, the degenerate eigenvalues of the ideal state are distributed according to the Wigner semicircle distribution. Ideally, this holds in the limit of many qubits, but is useful as a very neat approximation already for a modest number of qubits ( $n \gtrsim 6$ ). The knowledge how statistical noise influences the eigenvalue spectrum enables now a new ansatz to obtain a physical state estimate together with confidence inter-

vals. Hypothesis or residual tests can be used to compare the distribution of eigenvalues with the semicircle distribution in order to possibly identify colored noise. One should not attribute any physical meaning to eigenvalues which are consistent with the Wigner semicircle distribution, but rather set equal as due to statistics they cannot be distinguished from white noise. This ansatz gives the number of measurements necessary to avoid negative eigenvalues altogether using a simple modelling. The physically meaningful eigenvalues and eigenvectors of the density matrix are not affected by an estimation bias even for a finite number of measurements [32]. Numerical optimization to merely obtain a physical state estimate is replaced by a simple, yet efficient strategy, directly taking statistical effects into account. Evidently, the same analysis method is highly beneficial for quantum process tomography, e.g., of quantum logic gates. This finally enables one to obtain reliable estimates also for the high fidelity multiqubit states of future quantum simulations or first quantum computers.

## ACKNOWLEDGMENTS

We like to thank J. Dziewior, O. Gühne, M. Kleinmann, W. Laskowski, J. Meinecke, T. Moroder, and T. Paterek for stimulating discussions. This work has been supported by the EU (QWAD, ERC AdG QOLAPS, ERC StG GEDENTQOPT), by the MINECO (Project No. FIS2012-36673-C03-03), the Basque Government (Project No. IT4720-10), the UPV/EHU Program No. UFI 11/55, the OTKA (Contract No. K83858), and the excellence cluster Nano-Initiative Munich (NIM). The work of J.R. is partially supported by the Bonn-Cologne Graduate school BCGS. L.K. and C.S. acknowledge support by the Elite Network of Bavaria (QCCC and ExQM).

- 
- [1] M. G. A. Paris and J. Řeháček, eds., *Quantum state estimation* (Springer Berlin Heidelberg, 2004); R. Schmied, *Journal of Modern Optics* **63**, 1744 (2016); For recent methods for an efficient tomography of large systems, see G. Tóth, W. Wiczorek, D. Gross, R. Krischek, C. Schwemmer, and H. Weinfurter, *Phys. Rev. Lett.* **105**, 250403 (2010); M. Cramer, M. B. Plenio, S. T. Flammia, R. Somma, D. Gross, S. D. Bartlett, O. Landon-Cardinal, D. Poulin, and Y.-K. Liu, *Nat. Commun.* **1**, 149 (2010).
- [2] D. Gross, Y.-K. Liu, S. T. Flammia, S. Becker, J. Eisert, *Phys. Rev. Lett.* **105**, 150401 (2010).
- [3] C. Schwemmer, G. Tóth, A. Niggebaum, T. Moroder, D. Gross, O. Gühne, and H. Weinfurter, *Phys. Rev. Lett.* **113**, 040503 (2014).
- [4] D. F. V. James, P. G. Kwiat, W. J. Munro, and A. G. White, *Phys. Rev. A* **64**, 052312 (2001).
- [5] M. W. Mitchell, C. W. Ellenor, S. Schneider, and A. M. Steinberg, *Phys. Rev. Lett.* **91**, 120402 (2003); K. J. Resch, P. Walther, and A. Zeilinger, *Phys. Rev. Lett.* **94**, 070402 (2005); N. Kiesel, C. Schmid, G. Tóth, E. Solano, and H. Weinfurter, *Phys. Rev. Lett.* **98**, 063604 (2007).
- [6] H. Häffner, W. Hänsel, C. F. Roos, J. Benhelm, D. Chekalkar, M. Chwalla, T. Körber, U. D. Rapol, M. Riebe, P. O. Schmidt, C. Becher, O. Gühne, W. Dür, and R. Blatt, *Nature* **438**, 643 (2005); J. P. Home, D. Hanneke, J. D. Jost, J. M. Amini, D. Leibfried, and D. J. Wineland, *Science* **325**, 1227 (2009).
- [7] M. A. Nielsen, E. Knill, and R. Laflamme, *Nature* **396**, 52 (1998); O. Mangold, A. Heidebrecht, and M. Mehring, *Phys. Rev. A* **70**, 042307 (2004).
- [8] L. DiCarlo, J. M. Chow, J. M. Gambetta, L. S. Bishop, B. R. Johnson, D. I. Schuster, A. Majer, J. Blais, L. Frunzio, S. M. Girvin, and R. J. Schoelkopf, *Nature* **460**, 240 (2009); A. Fedorov, L. Steffen, M. Baur, M. da Silva, and A. Wallraff, *Nature* **481**, 170 (2012).
- [9] J. Peise, I. Kruse, K. Lange, B. Lücke, L. Pezzè, J. Arlt, W. Ertmer, K. Hammerer, L. Santos, A. Smerzi, and C. Klempt, *Nat. Commun.* **6**, 8984 (2015);

- [10] D. T. Smithey, M. Beck, M. G. Raymer, and A. Faridani, *Phys. Rev. Lett.* **70**, 1244 (1993); G. M. D'Ariano, C. Macchiavello, and N. Sterpi, *Quantum Semiclass. Opt.* **9**, 929 (1997); U. Leonhardt, M. Munroe, T. Kiss, T. Richter, and M. Raymer, *Opt. Comm.* **127**, 144 (1996).
- [11] Z. Hradil, *Phys. Rev. A* **55**, R1561 (1997); R. Blume-Kohout, *New J. Phys.* **12**, 043034 (2010).
- [12] Z. Hou, H.-S. Zhong, Y. Tian, D. Dong, B. Qi, L. Li, Y. Wang, F. Nori, G.-Y. Xiang, C.-F. Li, and G.-C. Guo, *New J. Phys.* **18**, 083036 (2016); J. Shang, Z. Zhang, H. K. Ng, *Phys. Rev. A* **95**, 062336 (2017).
- [13] M. Christandl and R. Renner, *Phys. Rev. Lett.* **109**, 120403 (2012); R. Blume-Kohout, arXiv:1202.5270 [quant-ph] (2012); T. Sugiyama, P. S. Turner, and M. Mura, *Phys. Rev. Lett.* **111**, 160406 (2013); J. Shang, H. K. Ng, A. Sehwat, X. Li, and B.-G. Englert, *New J. Phys.* **15**, 123026 (2013); C. Granade, J. Combes, D. G. Cory, *New J. Phys.* **18**, 033024 (2016); X. Li, J. Shang, H. K. Ng, B.-G. Englert, *Phys. Rev. A* **94**, 062112 (2016); P. Faist, and R. Renner, *Phys. Rev. Lett.* **117**, 010404 (2016); D. Suess, L. Rudnicki, T. O. Maciel, D. Gross, *New J. Phys.* **19**, 093013 (2017).
- [14] M. D. de Burgh, N. K. Langford, A. C. Doherty, and A. Gilchrist, *Phys. Rev. A* **78**, 052122 (2008); D. H. Mahler, L. A. Rozema, A. Darabi, C. Ferrie, R. Blume-Kohout, and A. M. Steinberg, *Phys. Rev. Lett.* **111**, 183601 (2013).
- [15] C. Schwemmer, L. Knips, D. Richart, H. Weinfurter, T. Moroder, M. Kleinmann, and O. Gühne, *Phys. Rev. Lett.* **114**, 080403 (2015).
- [16] A. F. Mood, *Introduction to the theory of statistics* (McGraw-Hill Inc., 1974); T. Sugiyama, P. S. Turner, and M. Mura, *New J. Phys.* **14**, 085005 (2012).
- [17] C. Butucea, M. Guta, T. Kypraios, *New J. Phys.* **17**, 113050 (2015).
- [18] E. Wigner, *Ann. Math.* **62**, 548 (1955); E. Wigner, *Ann. Math.* **67**, 325 (1958); M. L. Mehta, *Random Matrices* (Academic, New York, 2004).
- [19] F. Hiai, D. Petz, *The Semicircle Law, Free Random Variables and Entropy* (AMS, 2000).
- [20] N. Kiesel, Ph.D. thesis, Ludwig-Maximilians-Universität München (2007).
- [21] A. Edelman and N. R. Rao, *Acta Numerica* **14**, 233 (2005).
- [22] J. Gustavsson, *Ann. I. H. Poincaré - PR* **41**, 151 (2004); W. König, *Probability Surveys* **2**, 385 (2005); G. Aubrun, arXiv:1011.0275v3 [math.PR] (2012); T. Tao and V. Vu, arXiv:1201.4789v4 [math.PR] (2013); G. Borot, *An introduction to random matrix theory*, Lecture notes (2015); B. Eynard, T. Kumura, and S. Ribault, arXiv:1510.04430v1 [math.PR] (2015).
- [23] M. Chiani, arXiv:1209.3394 [cs.IT] (2014).
- [24] A. Miranowicz, K. Bartkiewicz, J. Peřina Jr., M. Koashi, N. Imoto, and F. Nori, *Phys. Rev. A* **90**, 062123 (2014).
- [25] Z. D. Bai, *Statistica Sinica* **9**, 611 (1999).
- [26] J. M. Renes, R. Blume-Kohout, A. J. Scott, and C. M. Caves, *J. Math. Phys.* **45**, 2171 (2004); J. Řeháček, B.-G. Englert, and D. Kaszlikowski, *Phys. Rev. A* **70**, 052321 (2004).
- [27] H. Zhu, B.-G. Englert, *Phys. Rev. A* **84**, 022327 (2011).
- [28] S. T. Flammia, D. Gross, Y.-K. Liu, J. Eisert, *New J. Phys.* **14**, 095022 (2012); M. Guta, T. Kypraios, and I. Dryden, *New J. Phys.* **14**, 105002 (2012).
- [29] C. A. Riofrio, D. Gross, S. T. Flammia, T. Monz, D. Nigg, R. Blatt, J. Eisert, *Nature Comm.* **8**, 15305 (2017).
- [30] G. Brida, M. Genovese, M. Gramegna, A. Meda, F. Piacentini, P. Traina, E. Predazzi, S. Olivares, and M. G. A. Paris, *Adv. Sci. Lett.* **4**, 1 (2011); A. Steffens, C. A. Riofrio, R. Hübener, and J. Eisert, *New J. Phys.* **16**, 123010 (2014).
- [31] T. Tao, *Probab. Theory Related Fields* **155**, 231-263 (2013).
- [32] For pure states being eigenstates of the measurement basis, several outcomes are not random, leading to a distorted spectrum, which, however, is still within the support of the Wigner semicircle distribution. Similar statement can be made for low rank states, where distortions of the probability distributions of outcomes can introduce unavoidable bias.
- [33] Other possibilities for model selection and rank estimation are given in [17, 29] and in J. A. Smolin, J. M. Gambetta, and G. Smith, *Phys. Rev. Lett.* **108**, 070502 (2012); T. L. Scholten and R. Blume-Kohout arXiv:1609.04385v1 [quant-ph] (2016).
- [34] Constrained optimization, e.g., using ML and LS fitting, effectively pushes the unphysical solution onto the border of physical space. The negative eigenvalues are shifted to 0 resulting in a rank in the order of  $r \approx 2^{n-1}$ , i.e., half of the system dimensions. This is (a) not related to the properties of the real quantum state, and (b) results in systematic shifts of the solution since normalization reduces the largest eigenvalue.
- [35] T. W. Anderson, D. A. Darling, *JASA* **49**, 765 (1954).
- [36] Note, due to statistics also the center  $c$  will exhibit a variance of this order of magnitude potentially resulting in negative values for high purity states. If the eigenvalue distribution cannot be distinguished from the Wigner semicircle, we propose to set  $c = 0$  and model only the state  $\varrho_r$ , e.g., by Maximum likelihood optimization or by a variational ansatz.
- [37] J. G. Skellam, *J. R. Stat. Soc.* **109**, 296 (1946); M. Abramowitz, I. A. Stegun, *Handbook of mathematical functions with formulas, graphs, and mathematical tables* (Dover Publications, 1965).
- [38] H. Cramér, *Scan. Actuar. J.* **1**, 13 (1928); R. von Mises, *Wahrscheinlichkeit, Statistik und Wahrheit* (Julius Springer, Wien, 1928).
- [39] T. Koshy, *Catalan Numbers with Applications* (Oxford University Press, 2009).

## Appendix A: Proof of the semicircle distribution

### 1. Limit of approximation

Obviously, the correlation values are restricted to lie between  $-1$  and  $1$ , i.e.,  $\tilde{T}_\mu \in [-1, 1]$ , while the Gaussian distribution has support everywhere in  $\mathbb{R}$ . The above approximation is justified for sufficiently large  $N$ , i.e., for cases where enough counts are recorded such that it is fairly unlikely to obtain  $|T_\mu| > 1$  in the Gaussian approximation. For the maximally mixed state with  $T_\mu = 0$  and for already only  $N = 10$ , the probability of  $\tilde{T}_\mu \notin [-1, 1]$

is fairly low for the Gaussian approximation,

$$1 - \int_{-1}^1 g(\tilde{T}_\mu) d\tilde{T}_\mu = 1 - \text{erf}(N/2) \approx 0.0016, \quad (\text{A1})$$

where  $\text{erf}(x) = 2/\sqrt{\pi} \int_0^x \exp(-\tau^2) d\tau$  denotes the error function. Thus, one can approximate the correlation with a Gaussian distribution, where the variance ( $\mathbb{V}[\tilde{T}_\mu] = \mathbb{E}[\tilde{T}_\mu^2] - \mathbb{E}[\tilde{T}_\mu]^2 = \mathbb{E}[\tilde{T}_\mu^2]$  for  $\mathbb{E}[\tilde{T}_\mu] = 0$ ) is given by

$$\begin{aligned} \mathbb{E}[\tilde{T}_\mu^2] &= \int_{-1}^1 d\tilde{T}_\mu g(\tilde{T}_\mu) \tilde{T}_\mu^2 \approx \int_{-\infty}^{\infty} d\tilde{T}_\mu g(\tilde{T}_\mu) \tilde{T}_\mu^2 \\ &= \sqrt{\frac{N}{2\pi}} \int_{-\infty}^{\infty} d\tilde{T}_\mu \exp\left(-\frac{N\tilde{T}_\mu^2}{2}\right) \tilde{T}_\mu^2 = \frac{1}{N}. \end{aligned} \quad (\text{A2})$$

## 2. Idea of proof

We prove that the eigenvalue distribution is described by a Wigner semicircle distribution based on the moments of the correlations. The moments of the density matrices obtained by means of the used overcomplete Pauli scheme for quantum state tomography are compared to the moments of a semicircle distribution. As the semicircle distribution is defined on a compact support, the equality of moments suffices to prove the equality of the distributions.

In fact, the proof shows the eigenvalue distribution of the deviation matrix  $\Delta_\varrho \equiv \tilde{\varrho} - \varrho$ . For the highly relevant case of  $\varrho$  being the maximally mixed state,  $\Delta_\varrho$  and  $\tilde{\varrho}$  have the same spectral distribution up to a shift of its center.

## 3. Moments of semicircle distribution

First, let us calculate the moments of a function describing a semicircle. The function

$$f_{c,R}(x) = \frac{2}{\pi R^2} \sqrt{R^2 - (x - c)^2} \quad (\text{A3})$$

describes a normalized semicircle centered around  $c$  with Radius  $R$ . Without loss of generality, we focus on the central moments. Instead, the function  $f_{c,R}(x)$  itself can be centered by setting  $c = 0$ . By this, moments and central moments are becoming equal. Due to the symmetry, all odd moments vanish, i.e.,  $m_{2k+1}^{\text{sc}} = \int_{-\infty}^{\infty} f_{0,R}(x) x^{2k+1} dx = 0$ . For the even moments, one

finds

$$m_2^{\text{sc}} = \int_{-\infty}^{\infty} f_{0,R}(x) x^2 dx = \left(\frac{R}{2}\right)^2, \quad (\text{A4a})$$

$$m_4^{\text{sc}} = \int_{-\infty}^{\infty} f_{0,R}(x) x^4 dx = 2 \left(\frac{R}{2}\right)^4, \quad (\text{A4b})$$

$$m_6^{\text{sc}} = \int_{-\infty}^{\infty} f_{0,R}(x) x^6 dx = 5 \left(\frac{R}{2}\right)^6, \quad (\text{A4c})$$

$$m_8^{\text{sc}} = \int_{-\infty}^{\infty} f_{0,R}(x) x^8 dx = 14 \left(\frac{R}{2}\right)^8, \quad (\text{A4d})$$

where the coefficient can be found recursively and equals the *Catalan numbers*  $\mathcal{C}_j$ . Thus, the moments are

$$m_{2k+1}^{\text{sc}} = \int_{-\infty}^{\infty} f_{0,R}(x) x^{2k+1} dx = 0, \quad (\text{A5a})$$

$$m_{2k}^{\text{sc}} = \int_{-\infty}^{\infty} f_{0,R}(x) x^{2k} dx = \mathcal{C}_k \left(\frac{R}{2}\right)^{2k} = \mathcal{C}_k (m_2^{\text{sc}})^k. \quad (\text{A5b})$$

To show that the spectral probability distribution of tomographically obtained completely mixed states equals a semicircle function, one has to recover Eqs. (A5) for the distribution of the eigenvalues of white noise.

### a. Catalan numbers

The Catalan numbers are given by, e.g., a recursively defined sequence also appearing in various counting problems. The zeroth Catalan number is  $\mathcal{C}_0 = 1$ . The subsequent numbers are defined by

$$\mathcal{C}_{j+1} = \mathcal{C}_j \frac{2(2j+1)}{j+2}. \quad (\text{A6})$$

Consequently, the first elements of the (zerobased) sequence  $\mathcal{C}_j$  are 1, 1, 2, 5, 14, 42, ... counting, e.g., the possibilities of  $2j$  persons shaking hands at the same time under the constraint that the hands of two pairs do not cross. Another example is the number of possible paths on the  $xy$  plane from point  $(0,0)$  to  $(2j,0)$  going only steps of  $(1, \pm 1)$  under the constraint that the  $x$  axis can be touched, but not crossed. These so called Dyck paths are equivalent to the question of how many possibilities of correctly setting  $j$  pairs of parentheses, e.g., “ $()()$ ” and “ $(())$ ” are the only possibilities for  $j = 2$ . For those and more applications of the Catalan numbers, see [39].

#### 4. Moments of density matrices

The  $k$ -th moment of the  $i$ -th eigenvalue will be denoted by

$$m_k^{(i)} = \mathbb{E} [\lambda_i^k]. \quad (\text{A7})$$

Since in this case, one is only interested in the joint spectral probability distribution, one has to average over all  $2^n$  eigenvalues in order to obtain the  $k$ -th moment of the spectral probability distribution of  $\tilde{\rho}'$ ,

$$\begin{aligned} m_k^{\text{ev}} &= \frac{1}{2^n} \sum_{i=1}^{2^n} m_k^{(i)} = \frac{1}{2^n} \sum_{i=1}^{2^n} \mathbb{E} [\lambda_i^k] \\ &= \frac{1}{2^n} \mathbb{E} \left[ \sum_{i=1}^{2^n} \lambda_i^k \right] = \mathbb{E} \left[ \frac{1}{2^n} \text{Tr} (D^k) \right] \\ &= \mathbb{E} \left[ \frac{1}{2^n} \text{Tr} \left( (U^\dagger \Delta \rho U)^k \right) \right] = \mathbb{E} \left[ \frac{1}{2^n} \text{Tr} \left( (\Delta \rho)^k \right) \right], \end{aligned} \quad (\text{A8})$$

where  $U^\dagger \Delta \rho U = D$  corresponds to the eigendecomposition of  $\Delta \rho$  with diagonal  $D$  ( $D_{i,j} = \delta_{i,j} \lambda_i$ ). To show that the spectrum of random density matrices, namely the distribution of their eigenvalues, is semicircular, one can now prove the equality of Eq. (A8) with the moments of the compactly supported semicircle given in Eqs. (A5).

##### a. First moment

By considering  $\Delta \rho$  instead of  $\tilde{\rho}$ , the first moment of the joint spectral distribution directly vanishes.

$$\begin{aligned} m_1^{\text{ev}} &= \mathbb{E} \left[ \frac{1}{2^n} \text{Tr} (\Delta \rho) \right] = \mathbb{E} \left[ \frac{1}{2^n} \text{Tr} (\tilde{\rho} - \rho) \right] \\ &= \frac{1}{2^n} \mathbb{E} [\text{Tr} (\tilde{\rho}) - \text{Tr} (\rho)] = 0. \end{aligned} \quad (\text{A9})$$

##### b. Second moment

By using Eq. (A8) and  $\Delta \rho = \frac{1}{2^n} \sum_{\mu} \Delta T_{\mu} \sigma_{\mu}$  with  $\Delta T_{\mu} = \tilde{T}_{\mu} - T_{\mu}$ , one can calculate the second moment,

$$\begin{aligned} m_2^{\text{ev}} &= \frac{1}{2^n} \sum_{i=1}^{2^n} \mathbb{E} [\lambda_i^2] = \frac{1}{2^n} \mathbb{E} \left[ \text{Tr} \left( (\Delta \rho)^2 \right) \right] \\ &= \frac{1}{2^{3n}} \sum_{\mu, \nu} \mathbb{E} [\Delta T_{\mu} \Delta T_{\nu}] \text{Tr} (\sigma_{\mu} \sigma_{\nu}) \\ &= \frac{2^n}{2^{3n}} \sum_{\mu, \nu} \mathbb{E} [\Delta T_{\mu} \Delta T_{\nu}] \delta_{\mu, \nu} \\ &= \frac{1}{2^{2n}} \sum_{\mu} \mathbb{E} \left[ (\Delta T_{\mu})^2 \right]. \end{aligned} \quad (\text{A10})$$

For the maximally mixed state, this can easily be evaluated since  $T_{\mu} = 0$  for all  $\mu \neq (0, 0, \dots, 0)$  and thus  $\tilde{T}_{\mu} = \Delta T_{\mu}$ . Because one can infer non-full correlations with larger statistics, see Eq. 2, one has to take this into account,

$$m_2^{\text{ev}} = \frac{1}{4^n N} \sum_{j=0}^{n-1} \binom{n-1}{j} \frac{3^{n-j}}{3^j} = \frac{10^n - 1}{12^n} \frac{1}{N} \approx \frac{5^n}{6^n} \frac{1}{N}. \quad (\text{A11})$$

Here, one uses that the sum runs over  $\mu$ , where  $\mathbb{E} [\tilde{T}_{\mu}^2] = 1/(3^j N)$  depends on the number  $j$  of local measurements of  $\sigma_0$  of  $\mu$ . If instead of the used overcomplete Pauli scheme a tomography scheme with  $\mathbb{E} (\tilde{T}_{\mu}^2) = 1/N$  for all  $\mu$  is used, (approximately) all matrix elements have the same variance. In this case, the second moment is found to be  $m_2^{\text{ev}} = 1/N$ .

##### c. Third and higher odd moments

All odd (centralized) moments of the eigenvalue distribution vanish, as we will argue at the example of the third moment. The third moment

$$\begin{aligned} m_3^{\text{ev}} &= \frac{1}{2^n} \sum_{i=1}^{2^n} \mathbb{E} [\lambda_i^3] = \frac{1}{2^n} \mathbb{E} \left[ \text{Tr} \left( (\Delta \rho)^3 \right) \right] \\ &= \frac{1}{2^{4n}} \mathbb{E} \left[ \text{Tr} \left( \sum_{\mu, \nu, \gamma} \Delta T_{\mu} \Delta T_{\nu} \Delta T_{\gamma} \sigma_{\mu} \sigma_{\nu} \sigma_{\gamma} \right) \right] \\ &= \frac{1}{2^{4n}} \sum_{\mu, \nu, \gamma} \mathbb{E} [\Delta T_{\mu} \Delta T_{\nu} \Delta T_{\gamma}] \text{Tr} (\sigma_{\mu} \sigma_{\nu} \sigma_{\gamma}) \end{aligned} \quad (\text{A12})$$

vanishes due to the expression  $\mathbb{E} [\Delta T_{\mu} \Delta T_{\nu} \Delta T_{\gamma}]$ . For all choices of indices  $\{\mu, \nu, \gamma\}$ , the expectation value will be taken of at least one correlation value in an odd power, i.e., the summation contains  $\mathbb{E} [\Delta T_{\mu} \Delta T_{\nu} \Delta T_{\gamma}] = \mathbb{E} [\Delta T_{\mu}] \mathbb{E} [\Delta T_{\nu}] \mathbb{E} [\Delta T_{\gamma}]$  (for mutually distinct indices  $\{\mu, \nu, \gamma\}$ ),  $\mathbb{E} [(\Delta T_{\mu})^2] \mathbb{E} [\Delta T_{\gamma}]$  (for  $\mu = \nu \neq \gamma$ ), and  $\mathbb{E} [(\Delta T_{\mu})^3]$  ( $\mu = \nu = \gamma$ ). Because  $\mathbb{E} [\tilde{T}_{\mu}] = T_{\mu}$ ,  $\mathbb{E} [\Delta T_{\mu}] = \mathbb{E} [(\Delta T_{\mu})^3] = 0$  holds, and thus all odd moments vanish,

$$m_{2k+1}^{\text{ev}} = 0 \quad (\text{A13})$$

## d. Fourth moment

Extending the procedure of Eqs. (A10) and (A11) to the fourth moment  $m_4^{\text{ev}}$ , one obtains

$$\begin{aligned} m_4^{\text{ev}} &= \frac{1}{2^{5n}} \mathbb{E} \left[ \text{Tr} \sum_{\mu, \nu, \gamma, \lambda} \Delta T_\mu \Delta T_\nu \Delta T_\gamma \Delta T_\lambda \sigma_\mu \sigma_\nu \sigma_\gamma \sigma_\lambda \right] \\ &= \frac{1}{2^{5n}} \sum_{\mu, \nu, \gamma, \lambda} \mathbb{E} [\Delta T_\mu \Delta T_\nu \Delta T_\gamma \Delta T_\lambda] \text{Tr} (\sigma_\mu \sigma_\nu \sigma_\gamma \sigma_\lambda). \end{aligned} \quad (\text{A14})$$

By the parity argument, only those terms contribute that contain indices in even power. One is now left to count the number of contributing summands in Eq. (A14). For illustration, consider the case of  $n = 2$  qubits and two different factors in the trace, each appearing twice. For commuting factors, e.g.,  $\sigma_{1,1}$  and  $\sigma_{2,2}$

$$\text{Tr} (\sigma_{1,1} \sigma_{1,1} \sigma_{2,2} \sigma_{2,2}) = \text{Tr} (\sigma_{1,1} \sigma_{2,2} \sigma_{1,1} \sigma_{2,2}) = \dots = 2^2 \quad (\text{A15})$$

holds for all 6 permutations of  $\sigma_{1,1} \sigma_{1,1} \sigma_{2,2} \sigma_{2,2}$ . On the other hand, one notices that for anticommuting factors  $\sigma_{1,1}$  and  $\sigma_{1,2}$ , i.e.,  $\sigma_{1,1} \sigma_{1,2} = -\sigma_{1,2} \sigma_{1,1}$ , different contributions occur.

$$\begin{aligned} \text{Tr} (\sigma_{1,1} \sigma_{1,1} \sigma_{1,2} \sigma_{1,2}) &= \text{Tr} (\sigma_{1,1} \sigma_{1,2} \sigma_{1,2} \sigma_{1,1}) = \\ \text{Tr} (\sigma_{1,2} \sigma_{1,1} \sigma_{1,1} \sigma_{1,2}) &= \text{Tr} (\sigma_{1,2} \sigma_{1,2} \sigma_{1,1} \sigma_{1,1}) = 2^2 \end{aligned} \quad (\text{A16})$$

and

$$\text{Tr} (\sigma_{1,1} \sigma_{1,2} \sigma_{1,1} \sigma_{1,2}) = \text{Tr} (\sigma_{1,2} \sigma_{1,1} \sigma_{1,2} \sigma_{1,1}) = -2^2 \quad (\text{A17})$$

are valid for the permutations of two *anticommuting* factors, each of those appearing twice. Hence, if the matrices  $\sigma_\mu$  and  $\sigma_\nu$  commute, 6 terms are contributing while anticommuting  $\sigma_\mu$  and  $\sigma_\nu$  lead to an effective contribution of only  $4 - 2 = 2$  summands.

We already could argue that the four indices  $\mu, \nu, \gamma, \lambda$  in Eq. (A14) reduce to two indices, e.g.  $\mu$  and  $\nu$ , where the sum in Eq. (A14) is now running over those two indices and all permutations of  $\sigma_\mu \sigma_\mu \sigma_\nu \sigma_\nu$ . By a simple counting argument, one realizes that after choosing  $\mu$  out of approximately  $4^n$  possibilities, one has approximately  $4^n/2$  possibilities for  $\nu$  such that  $[\sigma_\mu, \sigma_\nu] = 0$ , while the other  $4^n/2$  choices of  $\nu$  lead to  $\{\sigma_\mu, \sigma_\nu\} = 0$ . The expressions

$$\begin{aligned} \text{Tr} (\sigma_\mu \sigma_\mu \sigma_\nu \sigma_\nu) &= \text{Tr} (\sigma_\mu \sigma_\nu \sigma_\nu \sigma_\mu) \\ &= \text{Tr} (\sigma_\nu \sigma_\mu \sigma_\mu \sigma_\nu) = \text{Tr} (\sigma_\nu \sigma_\nu \sigma_\mu \sigma_\mu) = 2^2 \end{aligned} \quad (\text{A18})$$

give a positive contribution independently of the commutation relation between  $\sigma_\mu$  and  $\sigma_\nu$ .

$$\text{Tr} (\sigma_\mu \sigma_\nu \sigma_\mu \sigma_\nu) = \text{Tr} (\sigma_\nu \sigma_\mu \sigma_\nu \sigma_\mu) = \pm \text{Tr} \sigma_{0,0} = \pm 2^2 \quad (\text{A19})$$

leads in half of the choices of  $\mu$  and  $\nu$  to a positive, in the other half to a negative contribution. Contributions of those permutations cancel for different choices of indices.

$$m_4^{\text{ev}} = \frac{1}{2^n} \sum_{i=1}^{2^n} \mathbb{E} [\lambda_i^4] = \frac{1}{2^{5n}} \sum_{\mu, \nu, \gamma, \lambda} \mathbb{E} [\Delta T_\mu \Delta T_\nu \Delta T_\gamma \Delta T_\lambda] \text{Tr} (\sigma_\mu \sigma_\nu \sigma_\gamma \sigma_\lambda) = \quad (\text{A20a})$$

$$\frac{1}{2^{5n}} \frac{1}{2!} \sum_{\mu} \left[ \sum_{\nu: \{\nu \neq \mu\}} \mathbb{E} [(\Delta T_\mu \Delta T_\nu)^2] \text{Tr} \left( \sum_{i=1}^6 \mathcal{P}_i (\sigma_\mu \sigma_\mu \sigma_\nu \sigma_\nu) \right) + \mathbb{E} [(\Delta T_\mu)^4] \text{Tr} (\sigma_\mu \sigma_\mu \sigma_\mu \sigma_\mu) \right] \approx \quad (\text{A20b})$$

$$\frac{1}{2^{5n}} \frac{1}{2!} \sum_{\mu} \sum_{\nu: \{\nu \neq \mu\}} \mathbb{E} [(\Delta T_\mu \Delta T_\nu)^2] \text{Tr} \left( \sum_{i=1}^6 \mathcal{P}_i (\sigma_\mu \sigma_\mu \sigma_\nu \sigma_\nu) \right), \quad (\text{A20c})$$

where  $\sum_{i=1}^6 \mathcal{P}_i (\sigma_\mu \sigma_\mu \sigma_\nu \sigma_\nu) = \sigma_\mu \sigma_\mu \sigma_\nu \sigma_\nu + \sigma_\mu \sigma_\nu \sigma_\mu \sigma_\nu + \dots + \sigma_\nu \sigma_\nu \sigma_\mu \sigma_\mu$  denotes the summation over all permutations of the tensor products of Pauli matrices. By means of Eq. (2), one finds for the maximally mixed state  $\mathbb{E} [(\Delta T_\mu \Delta T_\nu)^2] = \mathbb{E} [(\Delta T_\mu)^2] \mathbb{E} [(\Delta T_\nu)^2] = 1/(3^{j_\mu} N) 1/(3^{k_\nu} N)$  and  $\mathbb{E} [(\Delta T_\mu)^4] = 3/(3^{j_\mu} N)^2$ , where  $j_\mu$  and  $k_\nu$  denote the number of  $\sigma_0$  operators appearing in  $\sigma_\mu$  and  $\sigma_\nu$ , respectively. In addition to the approximation in Eq. (A11), here we can neglect the second term in Eq. (A20b) for many qubits, i.e.,  $n \rightarrow \infty$  since the first summand is occurring  $\mathcal{O}(4^n)$  times more often. With the aforementioned argumentation,  $\text{Tr} \left( \sum_{i=1}^6 \mathcal{P}_i (\sigma_\mu \sigma_\mu \sigma_\nu \sigma_\nu) \right) = 4 \cdot 2^n$

holds, leading to

$$m_4^{\text{ev}} = \frac{1}{2^n} \sum_{i=1}^{2^n} \mathbb{E} [\lambda_i^4] = \frac{1}{2} \frac{1}{2^{5n}} \sum_{\mu} \sum_{\nu: \{\nu \neq \mu\}} \frac{1}{N^2} 4 \cdot 2^n = \frac{1}{2} \frac{1}{2^{5n}} 4^n 4^n \left( \frac{5^n}{6^n N} \right)^2 4 \cdot 2^n = 2 \cdot \left( \frac{5^n}{6^n N} \right)^2 = 2 (m_2^{\text{ev}})^2. \quad (\text{A21})$$

Using instead a tomography scheme with  $\mathbb{E} [(\Delta T_{\mu})^2] = 1/N$  for all  $\mu$  leads to  $m_4^{\text{ev}} = 2/N^2$ .

e. Sixth moment

Analogously, the sixth moment reads

$$m_6^{\text{ev}} = \frac{1}{2^n} \sum_{i=1}^{2^n} \mathbb{E} [\lambda_i^6] \approx \frac{1}{2^{7n}} \frac{1}{3!} \sum_{\mu} \sum_{\nu: \{\nu \neq \mu\}} \sum_{\gamma: \{\gamma \neq \nu, \gamma \neq \mu\}} \mathbb{E} [(\Delta T_{\nu})^2 (\Delta T_{\mu})^2 (\Delta T_{\gamma})^2] \cdot \text{Tr} \left( \sum_{i=1}^{90} \mathcal{P}_i (\sigma_{\mu} \sigma_{\nu} \sigma_{\mu} \sigma_{\nu} \sigma_{\gamma} \sigma_{\gamma}) \right), \quad (\text{A22})$$

where in the approximation terms of the form  $\mathbb{E} [(\Delta T_{\nu})^4 (\Delta T_{\mu})^2]$  and  $\mathbb{E} [(\Delta T_{\nu})^6]$  can be neglected for many qubits because terms of the form  $\mathbb{E} [(\Delta T_{\nu})^2 (\Delta T_{\mu})^2 (\Delta T_{\gamma})^2]$  dominate the expression by appearing  $\mathcal{O}(4^n)$  and  $\mathcal{O}(16^n)$  times more often, respectively. The factor  $1/3!$  compensates for multiple counting of exchanging indices. A specific summand of that expression, e.g.,  $\text{Tr} (\sigma_{\mu} \sigma_{\nu} \sigma_{\gamma} \sigma_{\mu} \sigma_{\nu} \sigma_{\gamma})$  contributes positively for half of the choices of  $\{\mu, \nu, \gamma\}$ , while the other choices lead to a negative contribution. Thus, permutations of that form cancel each other and can thus be neglected. By extending this argument, all *crossing* partitions do not have to be considered. Hence, one has to count only *noncrossing* partitions. Consequently, Eq. (A22) can for the maximally mixed state be further simplified leading to

$$m_6^{\text{ev}} = \frac{1}{2^n} \sum_{i=1}^{2^n} \mathbb{E} [\lambda_i^6] \approx \frac{1}{2^{7n}} \frac{1}{3!} 4^n \cdot 4^n \cdot 4^n \left( \frac{5^n}{6^n N} \right)^3 2^n 30 = 5 \left( \frac{5^n}{6^n N} \right)^3 = 5 (m_2^{\text{ev}})^3. \quad (\text{A23})$$

For a scheme with  $\mathbb{E} [(\Delta T_{\mu})^2] = 1/N$  for all  $\mu$ , one obtains  $m_6^{\text{ev}} = 5/N^3$ .

f. Crossing and noncrossing partitions - solution for arbitrary moments

Fig. 8 illustrates four examples of permutations occurring in the calculation of the sixth moment in Eq. (A22). For permutations of the six operators  $\sigma_{\mu}, \sigma_{\nu}, \sigma_{\mu}, \sigma_{\nu}, \sigma_{\gamma},$  and  $\sigma_{\gamma}$ , which are according to Fig. 8 “noncrossing” partitions, the commutation relation between the operators does not play a role. However, the trace of a permutation of those operators, which is a “crossing partition”, can result in  $2^n$  or  $-2^n$ , depending on the commutation relation of the operators. Since the number of crossing partitions resulting in  $2^n$  is the same as those resulting in  $-2^n$ , crossing partitions do not contribute for computing the moments. Only noncrossing permutations, where the (anti-)commutation relations between the operators do not influence the value of the trace, have to be considered. Thus, we are left to count the number of noncrossing partitions. For our case of  $k/2$  different operators, each of those appearing twice, the number of noncrossing partitions is given by the *Catalan number*  $\mathcal{C}_{k/2}$ , see Eq. (A6). All odd moments vanish, see Eq. (A13), while

for the  $2k$ -th moment

$$m_{2k}^{\text{ev}} = \frac{1}{2^n} \sum_{i=1}^{2^n} \mathbb{E} [\lambda_i^{2k}] = \mathcal{C}_k (m_2^{\text{ev}})^k = \mathcal{C}_k \left( \frac{5^n}{6^n N} \right)^k \quad (\text{A24})$$

holds, which equals the expressions of Eqs. (A5) for  $R = 2 (5/6)^{n/2} / \sqrt{N}$ . Eventually, we could prove that the spectral probability distribution of the completely mixed state converges to the Wigner semicircle in the limit of many qubits, although the different random variables, correlations in our case, have different variances.

For a tomography scheme with equal variances for all correlations (and all matrix elements), one as well obtains

$$m_{2k}^{\text{ev}} = \frac{1}{2^n} \sum_{i=1}^{2^n} \mathbb{E} [\lambda_i^{2k}] = \mathcal{C}_k (m_2^{\text{ev}})^k, \quad (\text{A25})$$

which leads to a Wigner semicircle distribution also in this case. However, due to Eq. (A4a), it holds  $R = 2\sqrt{1/N}$ , where  $N$  is given by the number of measurements for determining each correlation (and thus each matrix element).  $\square$

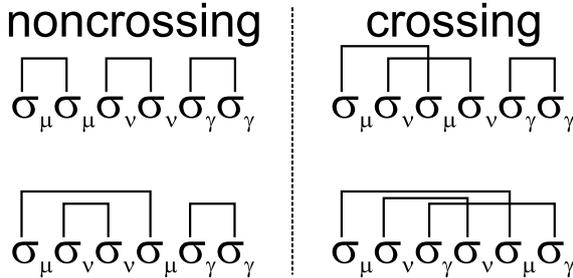


FIG. 8. Examples of *crossing* and *noncrossing* partitions. Four different permutations of  $\sigma_\mu\sigma_\mu\sigma_\nu\sigma_\nu\sigma_\gamma\sigma_\gamma$  are depicted, where equal operators are connected by lines. The permutations in the left column are called *noncrossing*, which corresponds to the graphical representation of noncrossing lines. Thus, the trace of those products is independent whether  $\sigma_\mu$ ,  $\sigma_\nu$  and  $\sigma_\gamma$  mutually commute or anticommute. In contrast, the trace of permutations shown in the right column depends on the (anti-)commutation relations. E.g.,  $\text{Tr}(\sigma_\mu\sigma_\nu\sigma_\mu\sigma_\nu\sigma_\gamma\sigma_\gamma)$  is positive if  $[\sigma_\mu, \sigma_\nu] = 0$ . These terms are positively contributing to the sum in Eq. (A22), but cancel with the negative values for  $\nu$  with  $\{\sigma_\mu, \sigma_\nu\} = 0$ .

#### Appendix B: Spectral probability distribution of one qubit density matrices

The spectral probability distribution  $g(\lambda)$  for the single qubit case as stated in the main text can be derived analytically. The eigenvalues are given by  $\lambda_{1,2} = \frac{1}{2} \left( 1 \pm \sqrt{\tilde{T}_1^2 + \tilde{T}_2^2 + \tilde{T}_3^2} \right)$ . Using the assumption of Gaussian distributed correlations, one obtains for the density  $g$  of eigenvalues  $\lambda$

$$g(\lambda) = \int_{-\infty}^{\infty} [\delta(\lambda - \lambda_1) + \delta(\lambda - \lambda_2)] \prod_{i=1}^3 g(\tilde{T}_i) d\tilde{T}_i \\ = 4\pi \int_0^{\infty} r^2 dr g(r) [\delta(\lambda - \lambda_1) + \delta(\lambda - \lambda_2)], \quad (\text{B1})$$

where the integration is performed over spherical coordinates, i.e. we substitute  $r = \sqrt{\tilde{T}_1^2 + \tilde{T}_2^2 + \tilde{T}_3^2}$ , and with  $\delta$  denoting the Dirac delta distribution. Since  $g(r) = g(\tilde{T}_1)g(\tilde{T}_2)g(\tilde{T}_3)$  with  $g(\tilde{T}_i) \propto \exp\left(-\frac{\tilde{T}_i^2 N}{2}\right)$ , we obtain

$$g(r) \propto \exp\left(-\frac{r^2 N}{2}\right) \quad (\text{B2})$$

with the number of counts  $N$ . Solving Eq. (B1), we finally obtain

$$g(\lambda) \propto \exp\left[-\frac{(1-2\lambda)^2 N}{2}\right] (1-2\lambda)^2. \quad (\text{B3})$$

The proportionality constant is given by normalization.

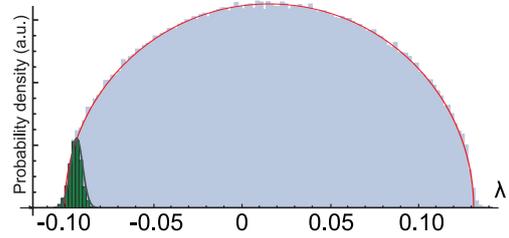


FIG. 9. The eigenvalue distribution as shown in Fig. 1 ( $n = 6$ ,  $N = 100$ ) in the main text with the distribution of the smallest eigenvalue (green bars) and the approximated Gaussian distribution (green line) with mean and standard deviation as given by Eqs. (C2) and (C3), respectively.

#### Appendix C: Probability of physical results

##### 1. Distribution of the smallest eigenvalue

It is known from Random Matrix Theory that in cases similar to the one considered here the smallest and largest eigenvalue follow the Tracy-Widom distribution, which solves a Painlevé differential equation of type II [21, 23]. This distribution can be approximated if mean, variance and skewness are known [23]. For determining the probability of negative eigenvalues, our numerical analysis showed that already the Gaussian distribution  $\mathcal{N}(\mu_{\lambda_1}, \sigma_{\lambda_1})$  with mean  $\mu_{\lambda_1}$  and standard deviation  $\sigma_{\lambda_1}$  is appropriately describing the distribution of the smallest eigenvalue. Thus, our approximation is based solely on  $\mu_{\lambda_1}$  and  $\sigma_{\lambda_1}$  with the probability density

$$\frac{1}{\sqrt{2\pi}\sigma_{\lambda_1}} \exp\left(-\frac{(\lambda - \mu_{\lambda_1})^2}{2\sigma_{\lambda_1}^2}\right), \quad (\text{C1})$$

of a Gaussian distribution (see Fig. 9) where the expectation value is

$$\mu_{\lambda_1} \approx c - R + 2\sigma_{\lambda_1} = c - 2 \left(\frac{5}{6}\right)^{n/2} \sqrt{\frac{1}{N}} + 2\sigma_{\lambda_1} \quad (\text{C2})$$

with center  $c$  and radius  $R$  of the noise eigenvalue distribution. The contribution from the standard deviation  $\sigma_{\lambda_1}$  is negligibly small, especially for larger  $n$ . For  $\sigma_{\lambda_1}$ , we obtain from a numerical analysis

$$\sigma_{\lambda_1} \approx \left(\frac{9}{14}\right)^{n+2} \cdot R = \left(\frac{9}{14}\right)^{n+2} 2 \left(\frac{5}{6}\right)^{n/2} \sqrt{\frac{1}{N}}. \quad (\text{C3})$$

The distribution of the largest eigenvalue can be approximated analogously.

## 2. Estimating the probability of physical results

The probability for the smallest eigenvalue being not negative can be obtained using Eq. (C1), thus

$$\begin{aligned}
 p(\lambda_1 > 0) &= 1 - \int_{-\infty}^0 \frac{1}{\sqrt{2\pi}\sigma_{\lambda_1}} \exp\left(-\frac{(\lambda - \mu_{\lambda_1})^2}{2\sigma_{\lambda_1}^2}\right) d\lambda \\
 &= 1 - \frac{1}{2} \operatorname{erfc}\left(\frac{\mu_{\lambda_1}}{\sqrt{2}\sigma_{\lambda_1}}\right) \\
 &\approx 1 - \frac{1}{2} \operatorname{erfc}\left[\frac{1}{\sqrt{8}} \left[4 + \left(\frac{14}{9}\right)^{n+2} \left(\left(\frac{6}{5}\right)^{\frac{n}{2}} \frac{c}{\sqrt{N}} - 2\right)\right]\right]
 \end{aligned} \tag{C4}$$

with the complementary error function  $\operatorname{erfc}(x) = 1 - \operatorname{erf}(x)$ . Obviously, with the Gaussian approximation, one can never guarantee that the result will be physical. Nevertheless, due to the small standard deviation of the distributions, unphysical results will get rather unlikely. For the example of the maximally mixed  $n = 6$  qubit, with  $N = 4371$  counts, the probability of a physical result is  $p(\lambda_1 > 0) \approx 0.046$ . For  $N = 4865$  counts, the probability already increased to  $p(\lambda_1 > 0) \approx 0.5$ , while for  $N = 5387$  counts, one expects a physical result with  $p(\lambda_1 > 0) \approx 0.954$ . Thus, the transition from the regime where unphysical estimates are highly likely to the regime of physical estimates can be estimated easily.

Applying Eq. (C4) to a state defined as  $\varrho = q|\psi\rangle\langle\psi| + (1-q)\varrho_{\text{cm}}$  with  $|\psi\rangle = |\text{GHZ}\rangle$ ,  $n = 6$ , and  $q = 0.8$  leads to the blue line in Fig. 10, being in good agreement with the simulated data as shown by points in this figure.

Similar considerations can be used to determine the number of total measurements required to separate small eigenvalues from statistical noise.

## Appendix D: Reconstruction of experimental state

### 1. Modification of radius for low rank states

According to App. A, the spectral probability density of the maximally mixed state converges to the Wigner semicircle. For low rank states, the radius of the spectral probability distribution of the insignificant eigenvalues can be approximated by the radius as given in Eq. (6) in the main text. The assumption that the radius does not depend on the rank  $r$  is only an approximation valid for  $r \ll 2^n$ . In Fig. 6, the radius  $R_r$  in dependence of the rank  $r$  is shown for 100 QST estimates ( $N = 10000$ ), where the underlying state is obtained by equally mixing  $r$  orthogonal random states.  $R$  can be estimated by means of the second moment  $m_2^{\text{ev}} = 1/2^n \mathbb{E}[\sum_i \lambda_i^2] = (R/2)^2$ . For a low rank state, the summation over the eigenvalues in Eq. (A10) has to

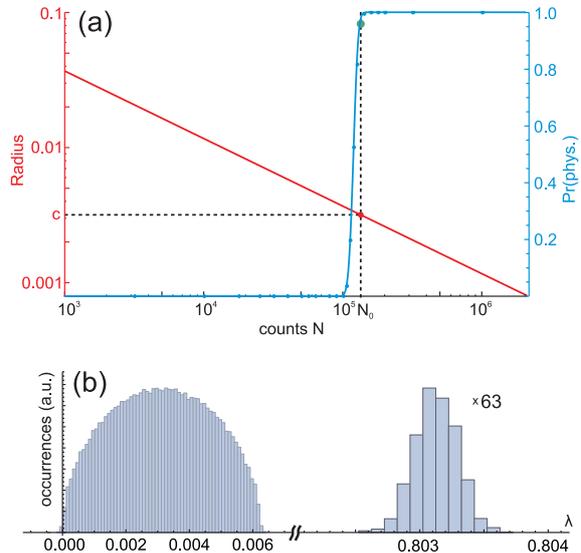


FIG. 10. Verifying our estimate  $N_0$  for the number of events needed for a physical state, as given in the main text. (a) (red line, left  $y$ -axis) The radius scales with  $1/\sqrt{N}$  with the number of events per basis setting  $N$  according to Eq. (6) in the main text, shown here for  $n = 6$  qubits. For a given admixture of white noise  $1 - q$ ,  $N_0$  is obtained by equating the radius with the center given in the main text as  $c_{q,r} = \frac{1-q}{2^n - r}$ , being a function of  $q$ . For  $q = 0.8$ , we obtain  $c = 0.0032$  and  $N_0 = 132921$ . (blue dots) The ratio of physical states for simulated QST for the given noise parameter. (blue line) Theoretical prediction based on the estimated distribution of the smallest eigenvalue, see Eq. (C4). (b) Histogram of eigenvalues for a GHZ state admixed with white noise for  $N_0$  events per setting [green circle in (a)]. The largest eigenvalues of each simulation correspond (approximately) to the GHZ state. Note that each simulated state results in 63 eigenvalues distributed within the semicircle and a single eigenvalue around  $0.8 + 0.2/64 \approx 0.803$ .

be modified as

$$m_2^{\text{ev}} \rightarrow m_2^{\text{ev}'} \approx m_2^{\text{ev}} \cdot \left(\frac{2^n - r}{2^n}\right) = m_2^{\text{ev}} \cdot (1 - r \cdot 2^{-n}) \tag{D1}$$

and

$$R \rightarrow R_r = R \cdot \sqrt{1 - r \cdot 2^{-n}} = 2 \left(\frac{5}{6}\right)^{\frac{n}{2}} \frac{\sqrt{1 - r \cdot 2^{-n}}}{\sqrt{N}}. \tag{D2}$$

For a rank  $r = 1$  state with  $n = 6$  qubits, the radius is changed by a factor of  $\sqrt{1 - r \cdot 2^{-n}} \approx 0.992$ , which can be neglected.

## 2. Hypothesis test

In the main text, the obtained spectrum of the experimentally prepared Dicke state [3] is tested against its anticipated distribution. For this hypothesis test, we use the Anderson-Darling test [35]. The distance between the hypothetical cumulative distribution function (CDF)  $F(x)$  and the empirical distribution function (EDF)  $\tilde{F}(\lambda)$  is calculated by means of

$$\mathcal{D} = \bar{n} \int_{-\infty}^{\infty} w(\lambda) \left[ \tilde{F}(\lambda) - F(\lambda) \right]^2 dF(\lambda), \quad (\text{D3})$$

where  $w(\lambda)$  is a weighting function and where  $\bar{n}$  denotes the sample size (number of eigenvalues in the empirical data).

The function  $w(\lambda) = 1/\{F(\lambda)[1-F(\lambda)]\}$  used for the Anderson-Darling test [35] weights the tails of the distribution higher than, e.g., the Cramér-von-Mises test [38]. The CDF of a semicircle with radius  $R$  and center  $c$  reads

$$F(\lambda) = \frac{1}{2} + \frac{(\lambda - c) \sqrt{1 - \frac{(\lambda - c)^2}{R^2}}}{\pi R} + \frac{\text{ArcSin}\left(\frac{\lambda - c}{R}\right)}{\pi} \quad (\text{D4})$$

for  $c - R < \lambda < c + R$ . For eigenvalues outside the interval,  $F(\lambda) = 0$  for  $\lambda \leq c - R$ , and  $F(\lambda) = 1$  for  $\lambda \geq c + R$ . Please note that the difference between the empirical and the cumulative distribution function contributes only for  $dF(\lambda) \neq 0$  to the test statistic from which the  $P$ -value is derived. Thus, if an eigenvalue is outside of the support of the Wigner semicircle distribution, it does not matter how far it is off. Consequently, it is recommended to check in the first place whether all the data are in the support and set the  $P$ -value to 0 otherwise.

As an additional check of the state estimation strategy introduced above, one can compare the  $P$ -values for various rank values  $r$  assumed in the tests. In Tab. I the center  $c$  and the corresponding  $P$ -values are given for low rank states with admixed white noise for different rank values  $r$ . Please note that for the cases where eigenvalues are found outside of the support of the Wigner semicircle distribution the corresponding hypothesis that statistical noise can be an explanation for these eigenvalues has to be rejected, effectively leading to a vanishing  $P$ -value. Evidently, for  $r \geq 4$ , the center of the assumed semicircle is shifted to negative values, being incompatible with our model. Thus,  $P_{\text{eff}}$  should be set to 0, too. Consequently, the analysis results in a rank  $r = 3$  state with a small amount of admixed white noise. Please note that the high  $P$ -value of  $1 - 8 \cdot 10^{-7}$  for  $r = 3$  alone is an indicator that the true state is rank 3, but due to the nature of hypothesis tests cannot be considered to be a proof. Instead, all other possibilities, namely  $r < 3$  and  $r > 3$  have to be ruled out, leading to  $r = 3$ . For high purity states one has to be careful. For  $r = 1$  there will be a certain probability, depending on the state relative to the projection of the Pauli tomography scheme, that

the largest eigenvalue is larger one,  $\lambda_{\text{max}} > 1$ , and consequently  $c < 0$ . Similarly, for low rank states with small noise admixed, the center can be negative, too. These are statistical effects, again, due to the finite number of measurements  $N$ . It is thus essential to estimate the variance of  $c$ , see App. E, Eq. (E3).

The comparison between CDF and EDF gives a direct handle to estimate the type of noise. The residues should follow a Student-t distribution for white noise. Different types of noise become visible easily, however, the particular dependence of characteristic parameters on various noise models requires more detailed investigation.

## 3. Reconstructed experimental state

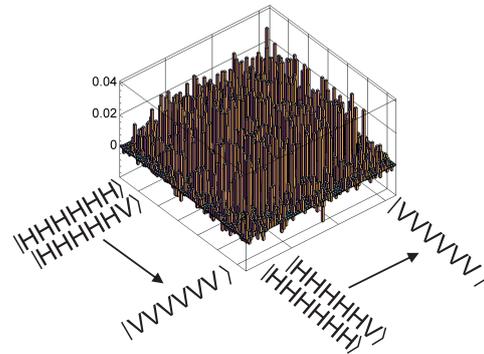


FIG. 11. Real part of the density matrix of the experimentally prepared  $n = 6$  qubit Dicke state  $|D_6^{(3)}\rangle$  after linear inversion.

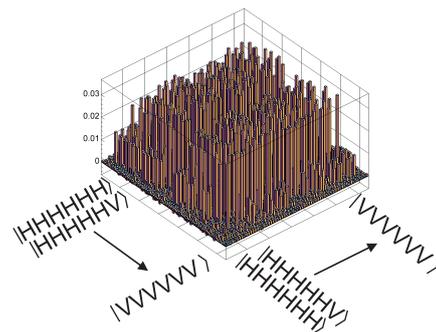


FIG. 12. Real part of the density matrix of the reconstructed state  $\rho_{r=3}$  after using the reconstruction scheme as explained in the main text.

With the aforementioned reconstruction scheme for the experimentally prepared and measured Dicke state  $|D_6^{(3)}\rangle$  [3], one obtains a state with rank  $r = 3$ . For the directly obtained, i.e., linearly reconstructed

state, the real part of the density matrix is shown in Fig. 11. One is barely able to recognize a pattern of bars of equal height, which are additionally strongly influenced by noise. After applying the reconstruction scheme introduced in the main part of this letter, i.e., taking into account the eigenstates of the relevant eigenvalues and average over the background noise, one obtains the density matrix, whose real part is shown in Fig. 12. The resulting eigenvalues are  $\{0.0004, 0.0004, \dots, 0.0004, 0.1495, 0.2160, 0.6102\}$  with the center  $c$  of the noise eigenvalues of about 0.0004. The eigenvalues are estimated by the eigenvalues of the raw matrix. The contribution of 8 photon events leads to higher order noise as seen by the eigenvalues 0.1495 and 0.2160. While we expect also noise from 10 photon events, the corresponding eigenvalues are too small to be resolved, and are thus within the distribution of white noise. Averaging the 61 low lying eigenvalues lead here to a large reduction of statistical noise, but still the structure of the state becomes evident.

#### Appendix E: Obtaining error bars using perturbation theory

By means of perturbation theory, the influence of finite statistics onto the estimated state can be assessed. In this section, we focus, for the sake of simplicity, mainly on estimated states with  $r = 1$ , i.e., a single eigenstate containing information, whereas the remaining  $2^n - 1$  eigenstates are regarded as background noise only. While the noise eigenvalues are of no further interest, one wants to assess an error interval for the large eigenvalue  $\lambda_{\max} \equiv \lambda_{2^n}$ .

For that purpose, consider a random matrix  $\Delta\varrho = 1/2^n \sum_{\mu} \Delta T_{\mu} \sigma_{\mu}$  with a set of parameters  $\{\Delta T_{\mu}\}$  analog to the correlation tensor.  $\Delta T_{\mu}$  are here taken to be random variables as defined in App. A with expectation value 0. The variances  $\Delta T_{\mu}$  are given according to App. A by  $[1 - T_{\mu}^2]/[3^{j(\mu)}N]$ . Let us consider  $\Delta T_{\mu}$  and consequently  $\Delta\varrho$  as a small perturbation resulting in a small perturbation of eigenstates and eigenvalues. The first order deviation of the large eigenvalue  $\lambda_{\max}$  is then obtained as

$$\begin{aligned} \Delta\lambda_{\max} &= \langle \Psi_{\max} | \Delta\varrho | \Psi_{\max} \rangle = \text{Tr}(|\Psi_{\max}\rangle \langle \Psi_{\max}| \Delta\varrho) \\ &= \frac{1}{2^{2n}} \text{Tr} \left( \sum_{\mu, \nu} \sigma_{\mu} \sigma_{\nu} \Delta T_{\mu} T_{\nu}^{\Psi_{\max}} \right) \\ &= \frac{1}{2^n} \sum_{\mu} \Delta T_{\mu} T_{\mu}^{\Psi_{\max}} \end{aligned} \quad (\text{E1})$$

with  $T_{\nu}^{\Psi_{\max}} = \text{Tr}(|\Psi_{\max}\rangle \langle \Psi_{\max}| \sigma_{\nu})$ .

The variance of  $\lambda_{\max}$  is thus obtained as

$$\begin{aligned} \mathbb{V}[\lambda_{\max}] &= \mathbb{E} [(\Delta\lambda_{\max})^2] \\ &= \frac{1}{2^{2n}} \sum_{\mu} \mathbb{E} [(\Delta T_{\mu})^2] \mathbb{E} [(T_{\mu}^{\Psi_{\max}})^2] \\ &= \frac{1}{4^n} \sum_{\mu} \frac{[1 - T_{\mu}^2]}{3^{j(\mu)}N} (T_{\mu}^{\Psi_{\max}})^2 \\ &\approx \frac{1}{4^n} \sum_{\mu} \frac{[1 - (\lambda_{\max} T_{\mu}^{\Psi_{\max}})^2] (T_{\mu}^{\Psi_{\max}})^2}{3^{j(\mu)}N}, \end{aligned} \quad (\text{E2})$$

with the approximation that the correlation  $T_{\mu}$  equals  $\lambda_{\max} T_{\mu}^{\Psi_{\max}}$ . Accordingly, the variance of the center of the noise eigenvalues can be estimated,

$$\mathbb{V}[c] \approx \left( \frac{1}{2^n - 1} \right)^2 \mathbb{V}[\lambda_{\max}]. \quad (\text{E3})$$

For estimates with rank  $r > 1$ , Eq. (E2) has to be modified such that it corresponds to the expected correlation, i.e., for a state with rank  $r$  and background noise, the variance of the  $i$ -th relevant eigenvalue

$$\mathbb{V}[\lambda_i] \approx \frac{1}{4^n} \sum_{\mu} \frac{[1 - \sum_l (\lambda_l T_{\mu}^{\Psi_i})^2] (T_{\mu}^{\Psi_i})^2}{3^{j(\mu)}N}, \quad (\text{E4})$$

where the summation index  $l$  labels all relevant eigenvalues and eigenstates. Consequently, statistical deviations of the measured matrix as described by  $\Delta\varrho$  lead to deviations of the obtained large eigenvalue. Thus, one uses a Gaussian distribution with mean value  $\lambda_{\max}$  and variance  $\mathbb{V}[\lambda_{\max}]$  in order to assess the confidence interval of  $\lambda_{\max}$ .

Please note that Eq. (E2) is state dependent due to the dependence on correlations of the corresponding eigenstate  $|\Psi_{\max}\rangle$ .

Obviously, if the pure target state  $|\phi\rangle$  is close to the eigenstate of the largest eigenvalue,  $|\Psi_{\max}\rangle$ , the variance of  $\langle \phi | \Delta\Psi_{\max} \rangle$  is far smaller than the variance of the large eigenvalue itself. Thus, to estimate the confidence interval of the fidelity it is sufficient to consider  $\mathbb{V}[\lambda_{\max}]$ , which is the dominating contribution. Deviations of the eigenstate can be neglected for the case of  $|\phi\rangle \approx |\Psi_{\max}\rangle$ . Error bars for some other figures of merit can directly be obtained as well. Using error propagation, the variance of the large eigenvalue given in Eq. (E2) results, for example, for the purity  $\mathcal{P}$  in

$$\mathbb{V}[\mathcal{P}] = 4\mathbb{V}[\lambda_{\max}] \left( \lambda_{\max} + \frac{\lambda_{\max} - 1}{2^n - 1} \right). \quad (\text{E5})$$

For the large eigenvalue of a noisy  $n = 6$  qubit GHZ state, as discussed in App. C, one obtains for  $N = 132921$  events a standard deviation of about  $1.5 \cdot 10^{-4}$ , see Fig. 5. Thus, the confidence interval for the large eigenvalue

with confidence level of 95% has a width of  $3 \cdot 10^{-4}$ . A similar analysis can be performed for the eigenvalues of the experimentally measured Dicke state. The largest eigenvalue is affected by the statistics according to a Gaussian distribution with standard deviation

of about 0.002. Thus, the confidence interval for  $\lambda_{64}$  is  $\lambda_{64} = 0.610 \pm 0.004$  when using a 95% confidence level ( $1.96 \sigma$ ). The estimated fidelity is  $F(|D_6^{(3)}\rangle, \tilde{\varrho}) = \langle D_6^{(3)} | \tilde{\varrho} | D_6^{(3)} \rangle = 0.602 \pm 0.004$ .

### 3.5 Conclusion

Quantum state tomography is a powerful tool for understanding, developing, improving, and verifying experimental setups to prepare quantum states. Together with quantum process tomography, state tomography thus provides powerful means for experimental analysis. On the other hand, those insights are costly in terms of labor and time. For full quantum state tomography, the experimental measurement effort scales exponentially with the number of qubits and can only be reduced for special states with, for example, some symmetry or with additional assumptions. Also, each measurement has to be performed with adequately large sample size in order to reduce the probability of unphysical results and for reducing statistical errors in the first place. Finally, many popular state estimators including the linear inversion and maximum likelihood estimation are point estimators and do not provide any error information at all. Those restrictions hold even more for quantum process tomography, which in turn requires quantum state tomography for a tomographically complete set of initial states.

In the previous sections, those obstacles have been discussed, pointing out some severe pitfalls of tomography. Our publication [P1] pinpoints these problems and assesses the size of systematic deviations for parameters typical in quantum state estimation scenarios. Moreover, in this publication, methods are derived to obtain upper or lower bounds for concave and convex functions of the density matrix, respectively, which do not suffer from any systematic effect.

Beyond that, in the publication [P2], common reconstruction methods of quantum state estimation have been analyzed. It is shown that for the widely used method of overcomplete Pauli tomography, degenerate eigenvalues of the state estimate will follow a Wigner semicircle distribution. Our tools allow to precisely describe this distribution based only on a few parameters such as the number of performed measurements, enabling to gauge the needed measurement effort. These tools nicely allow to estimate the influence of statistical effects onto the estimated quantum state and can thus be used to identify, for example, systematic errors in the experimental setup. Only if an eigenvalue is clearly outside the support of the Wigner semicircle distribution for the used sample size, this eigenvalue can be attributed some physical meaning. Hence, our methods enable to distinguish and separate the physically and statistically significant contributions from the statistical noise.

# Chapter 4

## Characterizing Entanglement

### Contents

---

<b>4.1</b>	<b>Introduction</b>	<b>104</b>
<b>4.2</b>	<b>Efficient Multipartite Entanglement Detection</b>	<b>105</b>
4.2.1	Need for Efficient Methods	105
4.2.2	Commuting and Anticommuting Operators	106
4.2.3	Nonlinear Witness for Bell States	108
4.2.4	Nonlinear Witness for Three-Qubit GHZ State	110
4.2.5	[P3] Multipartite Entanglement Detection with Minimal Effort	113
4.2.6	Application for Higher Dimensional Systems	124
4.2.7	[P4] Quadratic Entanglement Criteria for Qutrits	126
<b>4.3</b>	<b>Entanglement and Random Measurements</b>	<b>131</b>
4.3.1	Distribution of Expectation Values	132
4.3.2	Certifying Genuine Multipartite Entanglement	135
4.3.3	[P5] Multipartite Entanglement Analysis From Random Correlations	140
<b>4.4</b>	<b>Entanglement without Correlations</b>	<b>148</b>
4.4.1	Entanglement and Correlations	148
4.4.2	Correlations and Predictability	150
4.4.3	Postulates for Correlations	151
4.4.4	Even and Odd Number of Qubits	152
4.4.5	Obtaining a State without Correlations	152
4.4.6	[P6] Gen. Multipart. Ent. without Multipartite Correlations	154
4.4.7	[P7] Gen. $N$ -part. Ent. without $N$ -partite Correlation Functions	172
<b>4.5</b>	<b>Conclusion</b>	<b>179</b>

---

## 4.1 Introduction

Entanglement is one of the characterizing features of quantum systems. It is not only a fascinating characteristic by itself, stimulating research in quantum foundations, but also powers various quantum information protocols, including some variations of quantum key distribution [34], quantum teleportation [37], and quantum computing [40, 79]. Furthermore, there have been reports that entanglement might play a role in biological systems such as for the avian compass [184] or that coherent transport processes might be crucial for photosynthetic processes [185].

One naturally tries to understand phenomena in terms of our classical perception. Yet, this basic approach fails for entangled quantum systems. Although observers performing local measurements on classical systems may as well see correlated results, those correlations are different from the correlations produced by entangled systems. Consider the maximally entangled Bell singlet state  $|\psi^-\rangle$ . When two observers, Alice and Bob, perform local measurements in the same respective basis, they will obtain outcomes which look completely random. However, when both compare their results, they will realize that their outcomes were always opposite to each other. Any classical system, i.e., a system which can be described by local variables, can show perfect anticorrelations for *specific* measurements of Alice and Bob. Though, those perfect anticorrelations will only be found if Alice and Bob perform their measurements in one particular direction. Thus, only the (entangled) quantum system can give rise to this peculiar type of correlations.

Yet, while entanglement in bipartite systems can be understood fairly well, multipartite systems are much more difficult to describe. In bipartite systems, the presence of entanglement can for example be directly proven when full information about the quantum state is given. For multipartite systems, however, identifying entanglement is in general a hard task due to the exponentially increasing size of the Hilbert space and the multitude of combining subsystems (see Sec. 2.2.2). Additionally, the perfect anticorrelations in the bipartite example from above allow a maximal predictability. When Alice observes a specific measurement outcome, she can predict that Bob, when measuring in the same direction, will observe the opposite outcome. In contrast, in the multipartite scenario, maximal (anti-)correlations allow predictions with certainty in general only when all but one observer communicate about their results. The tripartite  $|W_3\rangle \propto |001\rangle + |010\rangle + |100\rangle$  state for example has a perfect anticorrelation  $T_{zzz} = -1$ , when all three observers Alice, Bob, and Charlie align their measurement apparatuses along  $\sigma_z$ . If Alice and Bob obtain, e.g., the respective results  $-1$  ( $\sigma_z|1\rangle = -|1\rangle$ ) and  $1$  ( $\sigma_z|0\rangle = |0\rangle$ ), they can together conclude that Charlie will measure  $1$ . As will be discussed on the basis of the so-called *no-correlation* state, genuinely multipartite entangled states exist which do not show any (full) correlation.

In the following sections of this chapter, first, an efficient and constructive way how to reveal genuine multipartite entanglement is discussed. Afterwards, correlations and entanglement are considered in the context of systems with noisy channels and without (stable) local reference frames. Finally, the role and connection of correlations, entanglement, and predictability are subject of the last part.

## 4.2 Efficient Multipartite Entanglement Detection

In Sec. 2.2.2, some methods had been introduced which can be employed for entanglement detection, namely the the PPT criterion (2.2.2.4), fidelity witnesses (2.2.2.5), and Bell inequalities (2.2.2.6). However, those methods are either restricted to small systems, sub-optimal or inefficient. Here, an efficient method is presented which allows to certify genuine multipartite entanglement based on the minimal possible number of measurements, i.e., measurements in only two bases. This method relies on prior knowledge of the quantum state similar to other entanglement witnesses, but it allows to construct adequate entanglement identifiers for various prominent quantum states including GHZ, Cluster and Dicke states.

### 4.2.1 Need for Efficient Methods

Consider two black box experiments, producing an arbitrary amount of identical copies of a quantum state with two qubits. Furthermore, let us assume one wants to identify whether the produced state is entangled and can thus be used as a resource for quantum information protocols such as quantum teleportation or entanglement swapping. One possibility to certify the entanglement is based on quantum state tomography, i.e., determination of the whole density matrix of the state, see also chapter 3. The first experiment might produce copies of the state

$$\varrho_1 = \begin{pmatrix} 0.362 & -0.223 - 0.005i & -0.054 + 0.217i & -0.096 + 0.349i \\ -0.223 + 0.005i & 0.138 & 0.030 - 0.134i & 0.054 - 0.217i \\ -0.054 - 0.217i & 0.030 + 0.134i & 0.138 & 0.223 + 0.005i \\ -0.096 - 0.349i & 0.054 + 0.217i & 0.223 - 0.005i & 0.362 \end{pmatrix} \quad (4.1)$$

described in a product basis, while the second experiment is shown to deliver

$$\varrho_2 = \begin{pmatrix} 0.042 & -0.173 + 0.012i & 0.022 + 0.007i & -0.095 - 0.024i \\ -0.173 - 0.012i & 0.718 & -0.091 - 0.036i & 0.384 + 0.124i \\ 0.022 - 0.007i & -0.091 + 0.036i & 0.013 & -0.055 + 0.004i \\ -0.095 + 0.024i & 0.384 - 0.124i & -0.055 - 0.004i & 0.227 \end{pmatrix}. \quad (4.2)$$

Although one might have some intuition by considering the ratio of moduli of coherences to populations, it is hard to tell if  $\varrho_1$  and  $\varrho_2$  are entangled. Using the PPT criterion (see Sec. 2.2.2.4), one directly obtains the eigenvalues  $\{-1/2, 1/2, 1/2, 1/2\}$  for the partial transpose of  $\varrho_1$ , while the partial transpose of  $\varrho_2$  results in  $\{0, 0, 0, 1\}$ . Thus,  $\varrho_1$  corresponds to a maximally entangled state, while  $\varrho_2$  is a product state. However, this analysis is limited to small systems and requires performing quantum state tomography, which in general requires an exponentially scaling effort with increasing number of qubits.

Fortunately, in most cases, the experimental apparatus does not need to be considered as a black box. Using prior knowledge, one can find more efficient means to reveal

entanglement, for example using entanglement witnesses. In order to reveal entanglement with properly constructed fidelity witnesses, one needs to measure the fidelity of the experimentally prepared state with respect to a reference state. Depending on the state, the amount of required measurements might also scale exponentially. For example, the  $n$ -qubit Greenberger-Horne-Zeilinger state, see Sec. 2.2.4.3, has  $2^{n-1}$  nonvanishing correlations in the subspace of locally observing  $\sigma_x$  or  $\sigma_y$ , i.e.,  $2^{n-1}$  of the expressions  $\langle \text{GHZ} | \sigma_{i_1} \otimes \sigma_{i_2} \otimes \cdots \otimes \sigma_{i_n} | \text{GHZ} \rangle$  with  $i_j \in \{x, y\}$  are nonvanishing, see also Sec. 2.2.1.3. In the subspace spanned by  $\sigma_0$  and  $\sigma_z$ ,  $2^{n-1} - 1$  nonvanishing correlations can be found, which all can be determined by measurements of only one measurement setting, namely, when all parties observe along  $\sigma_z$  direction, as this measurement setting allows to deduce also all corresponding non-full correlations. For example, a three-qubit GHZ state shows nonvanishing correlations for the nontrivial measurements of  $\sigma_x \otimes \sigma_x \otimes \sigma_x$ ,  $\sigma_x \otimes \sigma_y \otimes \sigma_y$ ,  $\sigma_y \otimes \sigma_x \otimes \sigma_y$ ,  $\sigma_y \otimes \sigma_y \otimes \sigma_x$ , and for  $\sigma_0 \otimes \sigma_z \otimes \sigma_z$ ,  $\sigma_z \otimes \sigma_0 \otimes \sigma_z$ ,  $\sigma_z \otimes \sigma_z \otimes \sigma_0$ , where latter three correlations can be obtained by the single measurement setting  $\sigma_z \otimes \sigma_z \otimes \sigma_z$ . Thus, in general correlations in  $2^{n-1} + 1$  measurement settings have to be evaluated to obtain the fidelity with respect to the GHZ state. However, for this very state, a more efficient strategy for detecting genuine multipartite entanglement based on only two measurement settings is known [56]. While this derivation was based on clever guessing, here, a method to constructively find nonlinear entanglement witnesses based on solely two settings is introduced [P3], see also [186].

## 4.2.2 Commuting and Anticommuting Operators

A sufficient criterion to show that a quantum state  $\rho$  with correlation tensor  $T$  is not bi-separable is given by violating the condition

$$\max_{T^{\text{bi-sep}}} (T^{\text{bi-sep}}, T)_M < (T, T)_M \equiv \|T\|_M^2, \quad (4.3)$$

where the left hand side contains the maximization over correlation tensors of bi-separable quantum states and  $M$  denotes some metric tensor for the scalar product between two tensors with

$$(X, Y)_M = \sum_{\mu_1, \dots, \mu_n; \nu_1, \dots, \nu_n} X_{\mu_1, \dots, \mu_n} Y_{\nu_1, \dots, \nu_n} M_{\mu_1, \dots, \mu_n, \nu_1, \dots, \nu_n}. \quad (4.4)$$

A derivation of expressions similar to Eq. (4.3) are given in [187, 188]. While it is sufficient to find any metric tensor  $M$  such that the corresponding maximization of the left-hand side gives a smaller value than the right-hand side of Eq. (4.3), it is hard to find a suitable tensor  $M$  and to perform the maximization over all bi-separable states.

However, an equivalent formulation of the criterion in Eq. (4.3) can be derived more easily based on the properties of commuting and anticommuting operators. Let  $C = \{C_1, C_2, \dots, C_N\}$  denote a set of  $N$  mutually *commuting*, traceless Hermitian operators with eigenvalues  $\{-1, 1\}$ . Then, the sum of the squares of the respective expectation

values is trivially bounded by  $N$  [189],

$$\sum_{j=1}^N \langle C_j \rangle^2 \leq N. \quad (4.5)$$

Furthermore, let  $A = \{A_1, A_2, \dots, A_N\}$  denote a set of  $N$  mutually *anticommuting*, traceless Hermitian operators with eigenvalues from the set  $\{-1, 1\}$ . For any quantum state,

$$\sum_{j=1}^N \langle A_j \rangle^2 \leq 1 \quad (4.6)$$

is satisfied [189]. Thus, the measurement results of mutually anticommuting operators obey a complementarity relation. The most straightforward example of the complementarity given in Eq. (4.6) is the Bloch sphere shown in Fig. 2.1. All physical states obey the relation  $\langle \sigma_x \rangle^2 + \langle \sigma_y \rangle^2 + \langle \sigma_z \rangle^2 \leq 1$ , which describes a unit sphere.

According to [189], two operators are called *cut-commuting* and *cut-anticommuting* if those operators are commuting or anticommuting on a subspace, respectively. Consider two operators  $O_1 = O_1^A \otimes O_1^B$  (with  $O_1^A \in \mathcal{H}_A$  and  $O_1^B \in \mathcal{H}_B$ ) and correspondingly  $O_2$ . If  $[O_1^A, O_2^A] = 0$  or  $[O_1^B, O_2^B] = 0$  holds, where  $[a, b] = ab - ba$  denotes the commutation bracket, the two operators  $O_1$  and  $O_2$  are said to be cut-commuting with respect to the bipartition  $A|B$ , or, for short,  $A|B$ -commuting. On the other hand, if the operators  $O_1 = O_1^A \otimes O_1^B$  and  $O_2 = O_2^A \otimes O_2^B$  are defined such that  $\{O_1^A, O_2^A\} = 0$  or  $\{O_1^B, O_2^B\} = 0$  for the anticommutation bracket  $\{a, b\} = ab + ba$ , they are said to be  $A|B$ -anticommuting.

According to Eq. (4.5), the sum of the squared expectation values of two commuting operators  $O_1 = O_1^A \otimes O_1^B$  and  $O_2 = O_2^A \otimes O_2^B$ , both with eigenvalues from  $\{-1, 1\}$ , is bounded for all physical quantum states,

$$\langle O_1 \rangle^2 + \langle O_2 \rangle^2 \leq 2. \quad (4.7)$$

Using the property that the expectation value of  $O = O^A \otimes O^B$  factorizes for a product state  $\varrho = \varrho_A \otimes \varrho_B$ , i.e.,  $\text{tr}(O\varrho) = \text{tr}(O^A\varrho_A)\text{tr}(O^B\varrho_B)$ , and using two operators  $O_1$  and  $O_2$  which not only commute, but also cut-anticommute with respect to the bipartition  $A|B$ , one can show that Eq. (4.7) is not tight for a product state. Instead, for a product state, one obtains

$$\langle O_1 \rangle^2 + \langle O_2 \rangle^2 = \text{tr}(O_1^A\varrho_A)^2 \text{tr}(O_1^B\varrho_B)^2 + \text{tr}(O_2^A\varrho_A)^2 \text{tr}(O_2^B\varrho_B)^2 \quad (4.8a)$$

$$\leq \underbrace{\left( \text{tr}(O_1^A\varrho_A)^2 + \text{tr}(O_2^A\varrho_A)^2 \right)}_{\leq 1} \underbrace{\left( \text{tr}(O_1^B\varrho_B)^2 + \text{tr}(O_2^B\varrho_B)^2 \right)}_{\leq 1} \quad (4.8b)$$

$$\leq 1. \quad (4.8c)$$

The expressions in Eq. (4.8b) are upper-bounded by 1 because of their complementarity due to Eq. (4.6). Therefore, for a product state, a large value for  $\langle O_1 \rangle^2$  excludes a large

value for  $\langle O_2 \rangle^2$ , while large values simultaneously for both operators indicate that the state is not a product state (and also not separable) along this specific bipartition.

Consequently, it was shown that

$$\langle O_1 \rangle^2 + \langle O_2 \rangle^2 \leq \begin{cases} 1 & \text{for product (and all separable) states,} \\ 2 & \text{in general.} \end{cases} \quad (4.9)$$

In the following, such an equation will be given in the form

$$\langle O_1 \rangle^2 + \langle O_2 \rangle^2 \leq_{\text{SEP}_{A|B}} 1 \quad (4.10)$$

A violation of the given bound then indicates non- $A|B$ -separability.

If all possible bi-partitions have been tested, a witness for genuine multipartite entanglement can be found. Such a combined witness, i.e., a single indicator to exclude all bi-partitions, can be written in the form

$$\mathcal{W} = \frac{1}{G_0} \sum_{j \in \mathcal{S}} v_j T_j^2 \leq_{\text{BISEP}} \frac{G}{G_0}, \quad (4.11)$$

where  $G_0$  denotes some normalization,  $G/G_0$  is the bound for bi-separable states,  $\mathcal{S}$  is the set of operator labels, and  $v_j$  the respective weighting for the correlation  $T_j$ . By comparing Eq. (4.11) with Eq. (4.3), one directly recognizes the analogies. The metric tensor  $M$  of Eq. (4.3) corresponds to the weights  $\{v_i\}_i$  and the choice of the set  $\mathcal{S}$ , whereas the maximization corresponds to finding a proper  $G/G_0$ . However, those tasks can now be performed with the aforementioned method. This enables to find entanglement witnesses for various quantum states of interest including GHZ and Cluster states as well as Dicke states. Since a measurement of a full correlation allows to obtain several non-full correlations, the scheme becomes highly efficient and can detect entanglement with the minimal number of measurement settings<sup>1</sup> for a non-probabilistic entanglement witness. For details, see [P3], which is reprinted in Sec. 4.2.5. The idea of the procedure will be illustrated by means of first a simple two-qubit example and will be further illustrated afterwards with a 3-qubit GHZ state.

### 4.2.3 Nonlinear Witness for Bell States

Consider the Bell state  $|\psi^-\rangle = \frac{1}{\sqrt{2}}(|01\rangle - |10\rangle)$ , see Sec. 2.2.4.2. Its nonvanishing correlations are found to be  $T_{00} = 1$  and  $T_{xx} = T_{yy} = T_{zz} = -1$ , i.e., the measurement results of two observers are always anticorrelated, as long as both measure in the same basis. The latter (nontrivial) correlations are given by the expectation values of the operators  $\sigma_x \otimes \sigma_x$ ,

<sup>1</sup>There are also approaches for detecting entanglement with a single measurement setting [190, 191]. However, those schemes are not only influenced by the finite statistics of measurements in each setting, but they also depend on the random choice of the measurement basis. In short, they are based on the fact that it is much more likely to find a large correlation in a highly entangled state than in a product state.

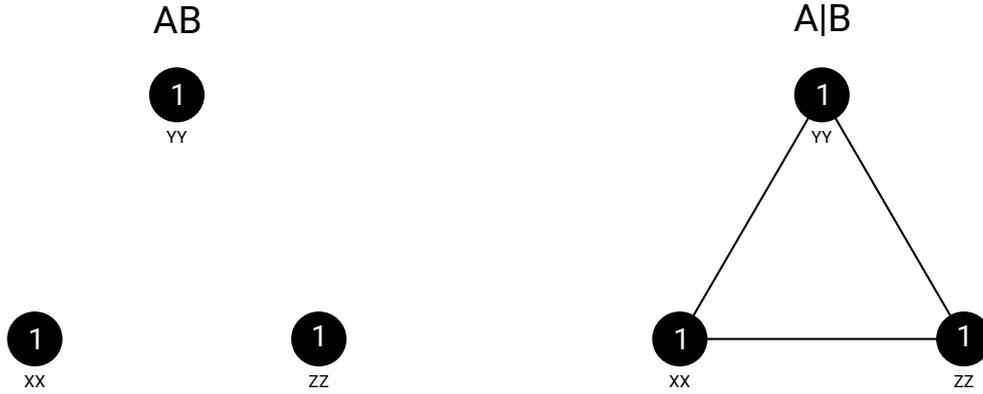


FIGURE 4.1: Graphical representation of derivation of bounds. (Left) Since the three operators  $\sigma_x \otimes \sigma_x$ ,  $\sigma_y \otimes \sigma_y$ , and  $\sigma_z \otimes \sigma_z$  commute, all of them can simultaneously take the maximal value of 1. The sum of squared expectation values is then 3. (Right) For obtaining the bound for separable states, one connects all mutually cut-anticommuting operators. The task is now to attribute the value 1 to as many operators as possible with the constraint that two mutually connected operators cannot reach 1 at the same time. Thus, one can assign this value to only one operator, leading to the bound 1.

$\sigma_y \otimes \sigma_y$ , and  $\sigma_z \otimes \sigma_z$ . Those three operators can be used for entanglement detection since the corresponding correlation values are nonvanishing. All three operators are mutually commuting, but cut-anticommuting (with respect to the only possible bi-partition), as one can verify by finding, e.g.,  $[\sigma_x \otimes \sigma_x, \sigma_y \otimes \sigma_y] = 0$  (commutation) and  $\{\sigma_x, \sigma_y\} = 0$  (cut-anticommutation). Thus, when considering all three operators, a straightforward nonlinear witness reads

$$\langle \sigma_x \otimes \sigma_x \rangle^2 + \langle \sigma_y \otimes \sigma_y \rangle^2 + \langle \sigma_z \otimes \sigma_z \rangle^2 = T_{xx}^2 + T_{yy}^2 + T_{zz}^2 \leq_{\text{BISEP}} 1, \quad (4.12)$$

whereas the general bound for physical states is 3, see also Fig. 4.1. The Bell state  $|\psi^-\rangle$  achieves that maximum of 3 and thus violates the bound of Eq. (4.12). It thus clearly indicates entanglement.

To study the noise sensitivity, consider the mixture of  $|\psi^-\rangle$  with some white noise,

$$\varrho = p|\psi^-\rangle\langle\psi^-| + \frac{1-p}{4}\mathbb{1}. \quad (4.13)$$

For this state, one observes  $T_{xx} = T_{yy} = T_{zz} = -p$ . Thus, as long as  $p > 1/\sqrt{3}$  holds, Eq. (4.12) can detect entanglement. Obviously, for the Bell state already two measurement settings are sufficient, e.g.,  $T_{xx}^2 + T_{zz}^2 \leq_{\text{BISEP}} 1$ . This further reduces the measurement effort, however, this witness is more prone to noise ( $p > 1/\sqrt{2}$  needed) than the one based upon three measurement settings.

#### 4.2.4 Nonlinear Witness for Three-Qubit GHZ State

Contrary to the two qubit example, for multipartite states several possibilities for bipartitions have to be inspected. To illustrate the principle, a nonlinear entanglement witness for the three-qubit GHZ state will be derived. The state  $|\text{GHZ}_3\rangle = \frac{1}{\sqrt{2}}(|000\rangle + |111\rangle)$  exhibits nonvanishing correlations for four measurements involving all three observers as well as three measurements involving only two of the observers, as shown in Tab. 4.1.

TABLE 4.1: All nonvanishing correlations of  $|\text{GHZ}_3\rangle$ . The first column gives the shorthand notation for the correlation values measured in the respective basis.

shorth. not.	basis	correlation
111	$\sigma_0 \otimes \sigma_0 \otimes \sigma_0$	1
1ZZ	$\sigma_0 \otimes \sigma_z \otimes \sigma_z$	1
Z1Z	$\sigma_z \otimes \sigma_0 \otimes \sigma_z$	1
ZZ1	$\sigma_z \otimes \sigma_z \otimes \sigma_0$	1
XXX	$\sigma_x \otimes \sigma_x \otimes \sigma_x$	1
XYY	$\sigma_x \otimes \sigma_y \otimes \sigma_y$	-1
YXY	$\sigma_y \otimes \sigma_x \otimes \sigma_y$	-1
YYX	$\sigma_y \otimes \sigma_y \otimes \sigma_x$	-1

Fig. 4.2 shows the trade-off relation for correlations that can be deduced when performing measurements in the settings  $\sigma_z \otimes \sigma_z \otimes \sigma_z$ ,  $\sigma_x \otimes \sigma_x \otimes \sigma_x$ ,  $\sigma_x \otimes \sigma_y \otimes \sigma_y$ ,  $\sigma_y \otimes \sigma_x \otimes \sigma_y$ , and  $\sigma_y \otimes \sigma_y \otimes \sigma_x$ . Please note that from the first measurement, also the non-full correlations shown in Tab. 4.1 can be deduced. The lines in Fig. 4.2 indicate a cut-anticommutation relation between the connected operators and thus that these operators cannot exhibit large expectation values at the same time for states being separable along this bipartition. Naturally, according to Eq. (4.5) all physical quantum states have to fulfill the inequality

$$\begin{aligned} &\langle \sigma_0 \otimes \sigma_z \otimes \sigma_z \rangle^2 + \langle \sigma_z \otimes \sigma_0 \otimes \sigma_z \rangle^2 + \langle \sigma_z \otimes \sigma_z \otimes \sigma_0 \rangle^2 + \\ &\langle \sigma_x \otimes \sigma_x \otimes \sigma_x \rangle^2 + \langle \sigma_x \otimes \sigma_y \otimes \sigma_y \rangle^2 + \langle \sigma_y \otimes \sigma_x \otimes \sigma_y \rangle^2 + \\ &\langle \sigma_y \otimes \sigma_y \otimes \sigma_x \rangle^2 \leq 7. \end{aligned} \quad (4.14)$$

Biseparable states, however, are constrained by the cut-anticommutation relations of those operators. From Fig. 4.2, one observes that for a biseparable state no more than 3 operators can show large expectation values simultaneously. Therefore, all biseparable states are bounded by

$$\begin{aligned} &\langle \sigma_0 \otimes \sigma_z \otimes \sigma_z \rangle^2 + \langle \sigma_z \otimes \sigma_0 \otimes \sigma_z \rangle^2 + \langle \sigma_z \otimes \sigma_z \otimes \sigma_0 \rangle^2 + \\ &\langle \sigma_x \otimes \sigma_x \otimes \sigma_x \rangle^2 + \langle \sigma_x \otimes \sigma_y \otimes \sigma_y \rangle^2 + \langle \sigma_y \otimes \sigma_x \otimes \sigma_y \rangle^2 + \\ &\langle \sigma_y \otimes \sigma_y \otimes \sigma_x \rangle^2 \leq_{\text{BISEP}} 3. \end{aligned} \quad (4.15)$$

One can directly deduce the white noise tolerance of  $\sqrt{3/7}$ , i.e., the three-qubit GHZ state is detected to be genuinely tripartite entangled as long as no more than  $1 - \sqrt{3/7}$  white noise is added.

In the publication *Multipartite Entanglement Detection with Minimal Effort* [P3], which is reprinted in the following section, this method is introduced. It is also applied to different experimentally prepared four-qubit states. Additionally, a method will be introduced allowing to use optimal weights for the measured correlations. In order to reduce the measurement effort, those witnesses are based upon only two different measurement settings, which is the smallest amount of possible measurements to certify genuine  $n$ -partite entanglement. As shown in [P3], for Greenberger-Horne-Zeilinger states as well as for cluster states, genuine  $n$ -partite entanglement can be detected with always two measurement settings irrespective of the number of qubits.

This publication is based on an idea of Marcin Wieśniak, with whom Christian Schwemmer, Harald Weinfurter and myself discussed the initial theoretical approach for entanglement witnessing and its experimental feasibility. I could significantly increase the efficiency of the initial methods by introducing different weights to the measurement outcomes. Together with Christian Schwemmer and Harald Weinfurter, I conceived the experimental apparatus. With Christian Schwemmer and later also with Nico Klein, I set up the experiment, conducted the measurements and evaluated the results. I also provided the experimental software for operating the experiment and conceived a systematic means to optimize the entanglement witness. The manuscript was written and edited by all authors.

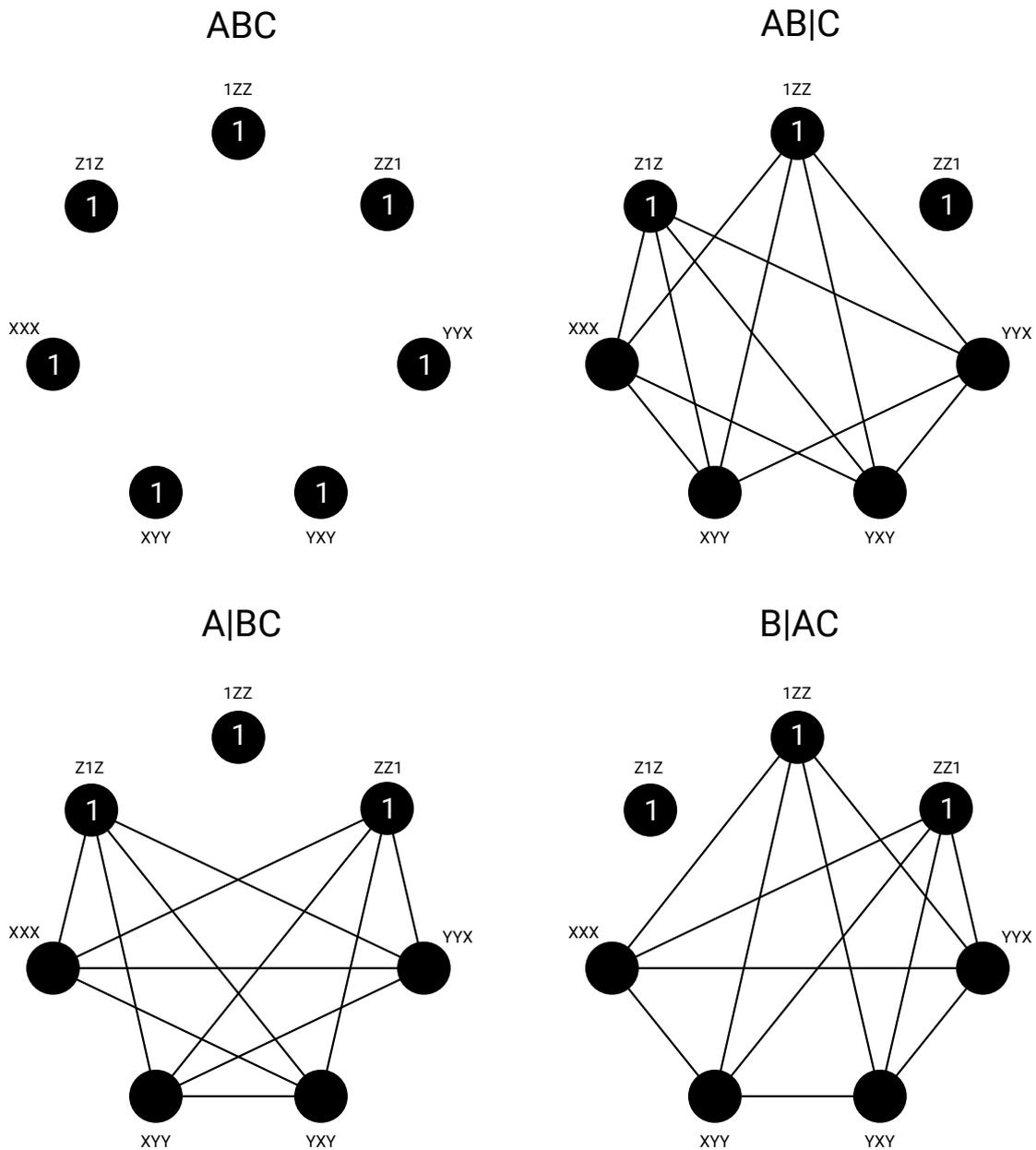


FIGURE 4.2: Correlations and cut-anticommutation relations of nonvanishing correlations of a three-qubit GHZ state. The upper left figure shows the 7 nonvanishing correlations, which are the expectation values of commuting operators. Thus, all of them can be 1 at the same time. When considering one of the 3 possible bipartitions, the mutual cut-anticommutation relations are as shown by the solid lines. A state separable along the given bipartition cannot obtain a large expectation for two mutually connected operators, leading to a trade-off relation between them.

### Multipartite Entanglement Detection with Minimal Effort

Lukas Knips,<sup>1,2</sup> Christian Schwemmer,<sup>1,2,\*</sup> Nico Klein,<sup>1,2</sup> Marcin Wieśniak,<sup>3</sup> and Harald Weinfurter<sup>1,2</sup>

<sup>1</sup>Max-Planck-Institut für Quantenoptik, Hans-Kopfermann-Strasse 1, D-85748 Garching, Germany

<sup>2</sup>Fakultät für Physik, Ludwig-Maximilians-Universität, D-80797 München, Germany

<sup>3</sup>Institute for Informatics, Faculty of Mathematics, Physics, and Informatics, University of Gdańsk, ul. Wita Stwosza 57, 80-308 Gdańsk, Poland

(Received 23 December 2015; published 18 November 2016; corrected 14 March 2017)

Certifying entanglement of a multipartite state is generally considered a demanding task. Since an  $N$  qubit state is parametrized by  $4^N - 1$  real numbers, one might naively expect that the measurement effort of generic entanglement detection also scales exponentially with  $N$ . Here, we introduce a general scheme to construct efficient witnesses requiring a constant number of measurements independent of the number of qubits for states like, e.g., Greenberger-Horne-Zeilinger states, cluster states, and Dicke states. For four qubits, we apply this novel method to experimental realizations of the aforementioned states and prove genuine four-partite entanglement with two measurement settings only.

DOI: 10.1103/PhysRevLett.117.210504

*Introduction.*—Entanglement is a fascinating feature of strictly quantum nature. It was first studied for the bipartite case [1,2] and has already been applied for first quantum communication tasks like quantum cryptography and quantum teleportation [3]. The generalization to multipartite entanglement comes with a whole new set of features providing, relative to separable states, information processing advantages for quantum computation and simulation or for quantum metrology. It is thus crucial to have tools at hand which allow us to identify *genuinely multipartite entangled* states [4–6].

Proving genuine multiparty entanglement is in general a complex task. Full quantum state tomography (QST) can be used for detecting and even for quantifying entanglement, but requires the determination of exponentially many parameters. Even when using simplified procedures [7–9], the effort is still significant. Thus, the goal was to find a direct measurement procedure for witnessing entanglement [5,10–13]. The only systematic method known today for constructing entanglement witnesses uses the fidelity relative to a chosen reference state. However, depending on the state, this as well leads to a rapidly increasing number of measurements required to infer the fidelity. Remarkably, specifically for the cluster and Greenberger-Horne-Zeilinger (GHZ) states, witnesses based on the stabilizer formalism [14] have been found incidentally which require only two measurements for any number of qubits [15]. Still, a systematic method, also not restricted to stabilizer states, is missing.

In this Letter, we introduce a constructive scheme to derive efficient multipartite entanglement witnesses, i.e., witnesses which can be evaluated from only a very small number of measurements. Our scheme employs basic properties of operators and their expectation values to construct witnesses for many relevant quantum states which require only two measurement settings, *independent* of the number of qubits. We show a way to enhance the

concept of finding measurements that are complementary for separable states [16] by introducing weights and providing the alternative scheme of testing violation of a set of inequalities in order to further increase the sensitivity. We demonstrate how to derive these efficient entanglement criteria for several of the most prominent quantum states, encompassing GHZ and cluster states, Dicke and  $W$  states, and the multipartite singlet state.

Every quantum mechanical  $N$ -qubit state  $\rho$  is uniquely described by its correlation tensor  $T$ ,

$$\rho = \frac{1}{2^N} \sum_{j \in \mathcal{I}} T_j \sigma_j, \quad (1)$$

where the set  $\mathcal{I} = \{0\dots00, 0\dots01, \dots, 3\dots33\}$  labels all indices  $j = (j_1 \dots j_N)$ ,  $j_i \in \{0, 1, 2, 3\}$  of the correlation tensor with  $\sigma_j = \sigma_{j_1} \otimes \dots \otimes \sigma_{j_N}$  and with Pauli matrices  $\sigma_0, \sigma_1, \sigma_2$ , and  $\sigma_3$ . The correlation tensor elements (for short called *correlations*) are given by  $T_j = \langle \sigma_j \rangle = \text{Tr}[\rho \sigma_j]$ . Since the eigenvalues of  $\sigma_j$  are  $\pm 1$ , the correlations are constrained to lie in the interval  $[-1, 1]$  and consequently  $T_j^2 \leq 1$ . These constraints, together with the physicality condition  $\rho \geq 0$  imply various bounds on the summed squares of correlations, which are helpful for the construction of efficient witness operators. Consider a set of  $n$  pairwise commuting operators  $\{\sigma_j : j \in \mathcal{C} \subset \mathcal{I}\}$ . These operators have common eigenstates, for which  $T_j = \pm 1$  holds. Consequently, the sum of squared correlations is bounded by  $\sum_{j \in \mathcal{C}} T_j^2 \leq n$ . On the contrary, for a set of pairwise anticommuting operators, e.g.,  $\{\sigma_j : j \in \mathcal{A} \subset \mathcal{I}\}$ , the threshold is [16]

$$\sum_{j \in \mathcal{A}} T_j^2 \leq 1, \quad (2)$$

establishing a complementarity relation between the correlations [17].

*Separability.*—Consider the bipartition (*cut*)  $\mathcal{B} = A|B$  of a multipartite quantum system into parts  $A$  and  $B$ . Two operators given by  $\sigma_{ab} = \sigma_a \otimes \sigma_b$  and  $\sigma_{a'b'} = \sigma_{a'} \otimes \sigma_{b'}$  anticommute with respect to the bipartition  $\mathcal{B}$  if  $\{\sigma_a, \sigma_{a'}\} = 0$  or  $\{\sigma_b, \sigma_{b'}\} = 0$ , i.e., if they locally anticommute on  $A$  or on  $B$ . According to Ref. [16] this property is called *cut-anticommutativity* or, more specifically,  $A|B$ -anticommutativity. Since for states separable with respect to  $\mathcal{B}$  the correlation tensor factorizes,  $T_{ab} = T_a T_b$ , these states fulfill

$$T_{ab}^2 + T_{a'b'}^2 \leq_{\text{SEP}} 1. \quad (3)$$

However, cut-anticommuting operators can also commute, i.e.,  $[\sigma_{ab}, \sigma_{a'b'}] = 0$ , allowing the common (entangled) eigenstates of  $\sigma_{ab}$  and  $\sigma_{a'b'}$  to exhibit  $T_{ab}^2 + T_{a'b'}^2 > 1$ . Therefore, violation of Eq. (3) rules out separability with respect to cut  $\mathcal{B}$ .

*Testing entanglement.*—To prove genuine multipartite entanglement of a state, Eq. (3) has to be violated for every possible bipartition. One starts with a list  $\{\sigma_j\}$  of all operators with nonvanishing expectation value,  $T_j \neq 0$  (all nonvanishing correlations). For the construction of the efficient entanglement criterion for a bipartition  $\mathcal{B}$ , one then chooses from that list two operators which are mutually commuting, but also cut-anticommuting relative to the bipartition  $A|B$ . One repeats this, until all bipartitions are tested.

The scheme becomes highly efficient if the correlation values of several  $\sigma_j$  can be obtained from the same measurement setting. In detail, this means that one makes use of the observation that from a single measurement setting  $\mathcal{M}_k$  with  $k = (k_1, k_2, \dots, k_N)$  and  $k_i \in \{1, 2, 3\}$  labeling the local Pauli measurements, all  $2^N$  correlations  $T_j$  with  $j \in \{(0, 0, \dots, 0), (0, 0, \dots, k_N), \dots, (k_1, k_2, \dots, k_N)\}$  can be inferred. Depending on the symmetry of the state, two measurement settings can suffice to prove genuine multipartite entanglement if one finds for each bipartition operators in the set that are commuting, but cut-anticommuting for the given bipartition.

*Combined entanglement witness.*—Combining the above criteria into a single witness facilitates the practical application (only a single value has to be calculated), though at the expense of a lower sensitivity, i.e., a reduced robustness against (white) noise. Compared to Ref. [16], the sensitivity can be considerably improved by using a weighted sum,

$$\mathcal{W} = \frac{1}{G_0} \sum_{j \in \mathcal{S}} v_j T_j^2 \leq_{\text{BISEP}} \frac{G}{G_0}, \quad (4)$$

where  $\mathcal{S} \subset \mathcal{I}$  labels the set of correlations that can be determined by the given set of measurements and where  $\leq_{\text{BISEP}}$  denotes that the inequality is valid for all biseparable states. The weights  $v_j$  and the (normalization) constants  $G$  ( $G_0$ ) are determined as follows:

(i) Depict the operators defined by  $\mathcal{S}$  as vertices of a graph (*anticommutativity graph*).

(ii) Assign weights  $v_j > 0$  to the vertices.

(iii) Choose bipartition  $\mathcal{B}_r$  and connect all vertices for which the corresponding operators cut-anticommute by edges. (If all operators indexed by  $\mathcal{S}$  mutually commute, no

edges will occur.) Distribute values  $c_j^{(m)} = \{0, 1\}$  among vertices under the constraint that any two “1’s” are not connected by an edge and calculate for each of the  $m$  possible distributions of 1’s the sum  $G_r^{(m)} = \sum_{j \in \mathcal{S}} c_j^{(m)} v_j$ . The case of no partition will be labeled by  $r = 0$ . Repeat step (iii) for all bipartitions  $\mathcal{B}_r$ .

(iv) Every choice of weights  $v_j$  in Eq. (4) defines a witness with  $G = \max_{r>0, m} G_r^{(m)}$  and  $G_0 = \max_m G_0^{(m)}$ . The ratio  $G/G_0$  determines the noise robustness of the criterion. To optimize the witness in terms of its noise robustness, one has to choose the weights  $v_j$  according to  $\arg \min_{\{v_j\}} G/G_0$ .

*Example.*—Let us consider the four-party GHZ state  $1/\sqrt{2}(|0000\rangle + |1111\rangle)$ , whose nonvanishing correlations are listed in Table I. As one can see, the measurement of the single setting  $\mathcal{M}_{3333}$  provides seven correlations with squared value 1. Since the operators of these correlations exhibit the same cut-anticommutation relation with any operator corresponding to the other eight correlations of Table I, the second measurement can be chosen arbitrarily out of those remaining eight. For example, the choice  $\mathcal{M}_{1221}$  for the second measurement setting results in the set of operators  $\{\sigma_{3333}, \sigma_{3300}, \sigma_{0033}, \sigma_{3003}, \sigma_{0330}, \sigma_{3030}, \sigma_{0303}, \sigma_{1221}\}$ , i.e.,  $\mathcal{S} = \{3333, 3300, \dots, 1221\}$ .

States that are, e.g.,  $A|BCD$ -separable fulfill, according to Eq. (3),

$$T_{3333}^2 + T_{1221}^2 \leq_{\text{SEP } A|BCD} 1. \quad (5)$$

Since  $\sigma_{1221}$  not only  $A|BCD$  anticommutes with  $\sigma_{3333}$ , but also with  $\sigma_{3030}, \sigma_{3003}, \sigma_{3300}$  from our list, a natural choice is to average over the expectation values of those four possibilities. Nonseparability against the partition  $A|BCD$  can then be detected with

$$\begin{aligned} \mathcal{W}_{A|BCD}^{\text{GHZ}} &= \frac{1}{2} \left[ \frac{1}{4} (T_{3030}^2 + T_{3003}^2 + T_{3300}^2 + T_{3333}^2) + T_{1221}^2 \right] \\ &\leq_{\text{SEP } A|BCD} \frac{1}{2}, \end{aligned} \quad (6)$$

where the additional normalization constant of  $1/2$  is introduced to ensure that  $\mathcal{W}_{A|BCD}^{\text{GHZ}} = 1$  holds for the ideal GHZ state, where all squared expectation values are one. The criteria for the remaining six bipartitions are derived analogously. For the list of criteria for the four-qubit Dicke, singlet, and  $W$  state see the Supplemental Material [18].

To derive a combined entanglement witness for the GHZ state, we use all eight operators labeled by  $\mathcal{S}$  (see Table I).

TABLE I. All nonvanishing correlations of the four-qubit GHZ state. The correlations in the first two rows can be inferred from the measurement setting  $\mathcal{M}_{3333}$  and the last correlation is obtained from the setting  $\mathcal{M}_{1221}$ .

$T_{0000}$	1	$T_{0033}$	1	$T_{0303}$	1	$T_{0330}$	1
$T_{3003}$	1	$T_{3030}$	1	$T_{3300}$	1	$T_{3333}$	1
$T_{2112}$	-1	$T_{2121}$	-1	$T_{2211}$	-1	$T_{2222}$	1
$T_{1111}$	1	$T_{1122}$	-1	$T_{1212}$	-1	$T_{1221}$	-1

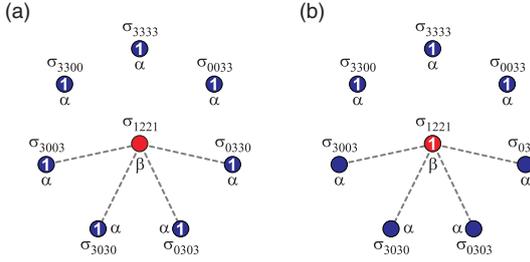


FIG. 1. The operators used to construct the witness  $\mathcal{W}^{\text{GHZ}}$ , cf. Table I. As an example, the cut-anticommutation relations for the cut  $AB|CD$  are indicated by dashed lines. One realizes that for each bipartition four of the seven operators obtained from the measurement of setting  $\mathcal{M}_{3333}$  cut-anticommute with  $\sigma_{1221}$ . Thus, the same weights  $\alpha$  are assigned to them, while  $\sigma_{1221}$  is weighted with  $\beta$ . Depending on the distribution of 1's, the sum for this bipartition is found to be either (a)  $\tilde{G}_r^{(1)} = 7\alpha$  or (b)  $\tilde{G}_r^{(2)} = 3\alpha + \beta$ . The best weights are obtained when the two assignments are equally good, i.e.,  $7\alpha = 3\alpha + \beta$ .

We assign equal weights to the seven operators obtained from the measurement setting  $\mathcal{M}_{3333}$ , i.e.,  $\alpha = v_{3333} = v_{0033} = \dots = v_{3300}$  since these mutually commute and behave similarly with regard to the cut-anticommutation relations with  $\sigma_{1221}$  for the different bipartitions. The weight of the remaining operator will be denoted by  $\beta = v_{1221}$ . From the anticommutativity graph (one without any edges) one obtains  $G_0 = 7\alpha + \beta$ . Depending on the distribution of 1's, the sums for all bipartitions are either  $G_r^{(1)} = 7\alpha$  or  $G_r^{(2)} = 3\alpha + \beta$ , see Fig. 1. For optimal noise robustness, one has to find the weights  $v_j$  by minimizing  $G/G_0$ . The minimum is achieved for  $G_r^{(1)} = G_r^{(2)}$ , thus  $7\alpha = 3\alpha + \beta$ , which leads, by arbitrarily setting  $\alpha = 1$ , to  $G_0 = 7\alpha + \beta = 7 + 4 = 11$  and  $G = 7\alpha = 3\alpha + \beta = 7$ . Then, the optimized two-measurement witness for the GHZ state reads

$$\mathcal{W}^{\text{GHZ}} = \frac{1}{11} (T_{3333}^2 + T_{3300}^2 + T_{0033}^2 + T_{3003}^2 + T_{0330}^2 + T_{3030}^2 + T_{0303}^2 + 4T_{1221}^2) \leq_{\text{BISEP}} \frac{7}{11}. \quad (7)$$

Analogously, for the cluster state  $|\mathcal{C}_4\rangle \propto (|0000\rangle + |0011\rangle - |1100\rangle + |1111\rangle)$  one obtains the witness

$$\mathcal{W}^{\mathcal{C}_4} = \frac{1}{6} (T_{3300}^2 + T_{3011}^2 + T_{0311}^2 + T_{1130}^2 + T_{1103}^2 + T_{0033}^2) \leq_{\text{BISEP}} \frac{2}{3}. \quad (8)$$

For details on the derivation, see the Supplemental Material [18].

*Extensions.*—Similar criteria can also be formulated for more qubits. The two-measurement-witness for the  $N$ -qubit GHZ state is based upon the measurements of  $\mathcal{M}_{3333\dots 3}$  and, e.g.,  $\mathcal{M}_{2211\dots 1}$  since one is able to find operators whose expectation value can be determined by those measurements such that Eq. (3) can be violated for each bipartition. Then, genuine multipartite entanglement is detected by violation of

$$\mathcal{W}^{\text{GHZ}_N} = \frac{1}{2^{N-1} + 2^{N-2} - 1} [T_{3333\dots 3}^2 + T_{0033\dots 3}^2 + T_{0303\dots 3}^2 + \dots + T_{33\dots 300}^2 + 2^{N-2} T_{2211\dots 1}^2] \leq_{\text{BISEP}} \frac{2^{N-1} - 1}{2^{N-1} + 2^{N-2} - 1} \xrightarrow{N \rightarrow \infty} \frac{2}{3}. \quad (9)$$

The extension of the criterion for the  $N$  qubit cluster state  $|\tilde{\mathcal{C}}_N\rangle$  ( $N$  even) is based on the correlations  $\{T_j\}_{j \in \mathcal{S}_{1313\dots 13} \cup \mathcal{S}_{3131\dots 31}}$  where the set  $\mathcal{S}_k$  indexes all non-vanishing correlations of the cluster state that can be determined from the measurement setting  $\mathcal{M}_k$ . Please note that  $|\tilde{\mathcal{C}}_4\rangle$  as defined via the stabilizer formalism [14] equals  $|\mathcal{C}_4\rangle$  up to LU transformations. Genuine multipartite entanglement of  $|\tilde{\mathcal{C}}_N\rangle$  is then identified by violation of

$$\mathcal{W}^{\tilde{\mathcal{C}}_N} = \frac{\sum_{j \in \mathcal{S}_{1313\dots 13}} T_j^2 + \sum_{j \in \mathcal{S}_{3131\dots 31}} T_j^2}{2(2^{N/2} - 1)} \leq_{\text{BISEP}} \frac{2^{N/2-1} + 2^{N/2} - 2}{2(2^{N/2} - 1)} \xrightarrow{N \rightarrow \infty} \frac{3}{4}. \quad (10)$$

*Analysis of experimental data.*—In order to experimentally demonstrate the applicability of our new entanglement criteria, we prepare a series of superpositions of GHZ and

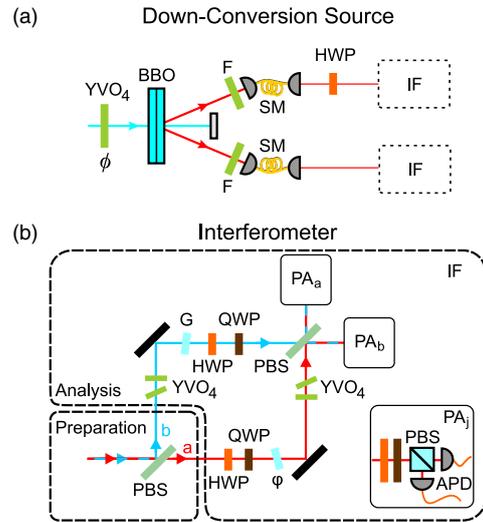


FIG. 2. Scheme of the experimental setup. In a first step (a) a type-I SPDC source together with a half wave plate (HWP) at angle  $\theta$  is used to prepare states of the form  $(|H\rangle(\cos 2\theta|H\rangle + \sin 2\theta|V\rangle) + e^{i\phi}|V\rangle(-\cos 2\theta|V\rangle + \sin 2\theta|H\rangle))/\sqrt{2}$ . The phase  $\phi$  can be set by a birefringent yttrium-vanadate crystal (YVO<sub>4</sub>). Interference filters (F) are applied for spectral filtering and spatial filtering is performed by coupling into single mode fibers (Supplemental Material [18]). In a second step (b), the state preparation is completed by increasing the Hilbert space by polarizing beam splitters (PBSs). Overlap at a beam splitter and polarization analysis allows us to measure all Pauli settings  $\sigma_i$  and to perform QST. YVO<sub>4</sub> crystals and glass plates (G and  $\phi$ ) inside the interferometer are used for phase and path length compensation, respectively.

TABLE II. Experimental values of the individual criteria and combined witnesses for the considered states. All values for a specific bipartition are clearly above the threshold of  $1/2$ , indicating genuine four-partite entanglement in all cases. The thresholds for the combined criteria are  $7/11$  (GHZ),  $2/3$  (cluster),  $4/5$  ( $D_4^{(2)}$ ), and  $3/5$  ( $\Psi_4$ ) respectively. The Dicke state cannot be significantly proven to be genuinely four-partite entangled by means of the combined witness, see Supplemental Material [18]. Hence, one has to resort to the individual criteria in this case.

Partition	$ \text{GHZ}\rangle$	$ \mathcal{C}_4\rangle$	$ \mathcal{D}_4^{(2)}\rangle$	$ \Psi_4\rangle$
$A BCD$	$0.894 \pm 0.007$	$0.922 \pm 0.006$	$0.819 \pm 0.013$	$0.804 \pm 0.019$
$B ACD$	$0.906 \pm 0.006$	$0.940 \pm 0.004$	$0.819 \pm 0.013$	$0.804 \pm 0.019$
$C ABD$	$0.906 \pm 0.006$	$0.940 \pm 0.004$	$0.819 \pm 0.013$	$0.804 \pm 0.019$
$D ABC$	$0.906 \pm 0.006$	$0.928 \pm 0.006$	$0.819 \pm 0.013$	$0.804 \pm 0.019$
$AB CD$	$0.904 \pm 0.006$	$0.922 \pm 0.006$	$0.627 \pm 0.013$	$0.608 \pm 0.017$
$AC BD$	$0.906 \pm 0.006$	$0.948 \pm 0.004$	$0.620 \pm 0.013$	$0.594 \pm 0.021$
$AD BC$	$0.901 \pm 0.006$	$0.943 \pm 0.004$	$0.625 \pm 0.013$	$0.622 \pm 0.021$
Combined	$0.916 \pm 0.005$	$0.940 \pm 0.004$	$0.801 \pm 0.017$	$0.683 \pm 0.014$

cluster states with variable weights. Different linear optical setups to prepare either four-qubit GHZ [21] or cluster states [22] are known. To have the flexibility to prepare superpositions of GHZ and cluster states in a single setup, we resort to a two photon experiment using two degrees of freedom per photon, namely polarization and path [23]. This approach enables one to prepare states with both high fidelity and high count rates. From now on, the computational basis states  $|0\rangle$  and  $|1\rangle$  are encoded either in polarization or in the path degree of freedom, i.e.,  $|0\rangle \rightarrow |H\rangle$  and  $|1\rangle \rightarrow |V\rangle$  for horizontal ( $H$ ) and vertical ( $V$ ) polarization and  $|0\rangle \rightarrow |a\rangle$  and  $|1\rangle \rightarrow |b\rangle$  for paths  $a$  and  $b$ .

The photon source shown in Fig. 2(a) uses spontaneous parametric down-conversion and allows us to prepare states of the form  $(|H\rangle(\cos(2\theta)|H\rangle + \sin(2\theta)|V\rangle) + e^{i\phi}|V\rangle(\sin(2\theta)|H\rangle - \cos(2\theta)|V\rangle))/\sqrt{2}$  (see the Supplemental Material [18] for details). In order to achieve the intended four-qubit state, coupling to the path degree of freedom is required. Thus, the polarization dependence of the output of a polarizing beam splitter is used; i.e., photons are transformed as  $|H\rangle \rightarrow |Ha\rangle$  and  $|V\rangle \rightarrow |Vb\rangle$  with  $a$  and  $b$  denoting the corresponding output modes of the PBS, see Fig. 2(b). Consequently, four-qubit states parametrized by  $\theta$  and  $\phi$ ,  $|\Psi(\theta, \phi)\rangle = (\cos(2\theta)|HaHa\rangle + \sin(2\theta)|HaVb\rangle + e^{i\phi}\sin(2\theta)|VbHa\rangle - e^{i\phi}\cos(2\theta)|VbVb\rangle)/\sqrt{2}$ , are obtained. Prominent members of  $|\Psi(\theta, \phi)\rangle$  are for example the GHZ states  $(|HaHa\rangle \mp |VbVb\rangle)/\sqrt{2}$  for  $\theta = 0$  and  $\phi = 0, \pi$ , respectively, or the cluster states  $(|HaHa\rangle + |HaVb\rangle \pm |VbHa\rangle \mp |VbVb\rangle)/2$  obtained for  $\theta = \pi/8$  and  $\phi = 0, \pi$ .

The prepared states are characterized by means of QST, proving full control of the experimental apparatus. This can be achieved with an interferometer setup as shown in Fig. 2(b), overlapping the modes  $a$  and  $b$  together with a polarization analysis and coincidence detection in the outputs.

*Experimental results.*—Thirteen states were prepared with  $\phi = \pi$  and  $\theta$  being increased from  $0$  (GHZ) to  $\pi/8$  (cluster) and to  $\pi/4$  (GHZ') in equidistant steps. The coincidence rate was approximately  $100 \text{ s}^{-1}$  with a measurement time of  $40 \text{ s}$  for each basis setting, resulting in

3700–4400 counts per setting and a measurement time of about  $12 \text{ h}$  to perform QST for all states. A measure for the quality of a prepared state  $\rho_{\text{exp}}$  with respect to a pure target state  $|\psi\rangle$  is the fidelity  $\mathcal{F} = \text{Tr}(\rho_{\text{exp}}|\psi\rangle\langle\psi|)$ . For the GHZ state, we observed a fidelity of  $\mathcal{F} = 0.958 \pm 0.004$ , while for the cluster state it was  $\mathcal{F} = 0.962 \pm 0.003$ . For the other states, see Table IV in the Supplemental Material [18].

Genuine four-partite entanglement could be tested using two measurement settings only. Let us start to determine the witnesses for the GHZ state from measuring two settings  $\mathcal{M}_{3333}$  and  $\mathcal{M}_{1221}$ . The values of the respective measured correlations (Table III in the Supplemental Material [18]) lead to a violation of all seven criteria by at least 56 standard deviations for all cuts, see Table II. Also, the combined criterion  $\mathcal{W}^{\text{GHZ}} = 0.916 \pm 0.005 > \frac{7}{11}$  certifies genuine four-partite entanglement. For the cluster state,

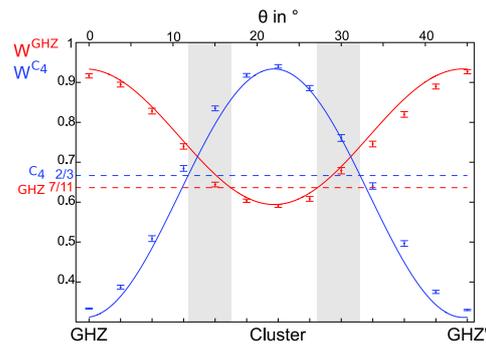


FIG. 3. The entanglement criterion for the GHZ states allows us to detect most of the superpositions of GHZ and the cluster state to be genuinely four-partite entangled (red, starting at 0.92) whereas the criterion for the cluster state detects states around  $\theta = 22.5^\circ$  to be genuinely four-partite entangled (blue, starting at 0.33). States within the gray shaded areas can be detected to be genuinely four-partite entangled by means of both criteria. The solid lines show the theoretically expected values for the target states  $|\Psi(\theta, \phi)\rangle$  mixed with white noise such that on average the fidelities correspond to the measured values.

according to our entanglement criterion, the measurement settings  $\mathcal{M}_{1133}$  and  $\mathcal{M}_{3311}$  were used (see Supplemental Material [18]), resulting in  $\mathcal{W}^{\mathcal{C}_4} = 0.940 \pm 0.004 > \frac{2}{3}$  for the combined criterion.

Using the combined witnesses, we analyze the entanglement for all states  $|\Psi(\theta, \phi)\rangle$  (Fig. 3). As can be seen, 10 of 13 states can be detected as genuinely four-partite entangled by the criterion  $\mathcal{W}^{\text{GHZ}}$ , the 6 states close to the cluster state can be determined by means of  $\mathcal{W}^{\mathcal{C}_4}$ . Some states can be shown to be truly four-partite entangled by means of both criteria as both are above their respective threshold. Genuine four-partite entanglement could be proven with experimental data of the Dicke state  $|D_4^{(2)}\rangle$  [24] and the singlet state [25], see Table II. For more details see the Supplemental Material [18].

*Conclusion.*—We have introduced a novel scheme for the systematic construction of entanglement witnesses, which need a minimal number of measurements for their evaluation independent of the number of qubits. We believe that such a minimal multipartite entanglement detection will become a handy diagnostic procedure as it is fast and simple. An interesting question is what other states can reveal their multipartite quantum correlations in two measurements. Another challenge is to find even stronger criteria, which, by possibly going to few more measurements, will detect multipartite entanglement with a higher robustness against noise.

We thank D. Richart and S. Habib for stimulating discussions. M.W. was supported by FNP project HOMING-PLUS/2011-4/14 and later by MNiSW Grant No. IdP2011 00361 (Ideas Plus) and NCN Grant No. 2015/19/B/ST2/01999. We acknowledge the support of this work by the EU (QWAD, ERC QOLAPS, FP7 BRISQ2). C. S. and L. K. thank QCCC and ExQM, respectively, of the Elite Network of Bavaria for support.

\*Present address: IBM Research GmbH, Säumerstrasse 4, CH-8803 Rüschlikon, Switzerland.

- [1] A. Einstein, B. Podolski, and N. Rosen, *Phys. Rev.* **47**, 777 (1935).
- [2] J. S. Bell, *Physics* (Long Island City, N.Y.) **1**, 195 (1964).
- [3] J.-W. Pan, Z.-B. Chen, C.-Y. Lu, H. Weinfurter, A. Zeilinger, and M. Żukowski, *Rev. Mod. Phys.* **84**, 777 (2012).
- [4] R. Horodecki, P. Horodecki, M. Horodecki, and K. Horodecki, *Rev. Mod. Phys.* **81**, 865 (2009).
- [5] O. Gühne and G. Tóth, *Phys. Rep.* **474**, 1 (2009).
- [6] Č. Brukner, M. Żukowski, J.-W. Pan, and A. Zeilinger, *Phys. Rev. Lett.* **92**, 127901 (2004).
- [7] G. Tóth, W. Wieczorek, D. Gross, R. Krischek, C. Schwemmer, and H. Weinfurter, *Phys. Rev. Lett.* **105**, 250403 (2010); T. Moroder, P. Hyllus, G. Tóth, C. Schwemmer, A. Niggebaum, S. Gaile, O. Gühne, and H. Weinfurter, *New J. Phys.* **14**, 105001 (2012); C. Schwemmer, G. Tóth, A. Niggebaum, T. Moroder, D. Gross, O. Gühne, and H. Weinfurter, *Phys. Rev. Lett.* **113**, 040503 (2014).
- [8] M. Cramer, M. B. Plenio, S. T. Flammia, R. Somma, D. Gross, S. D. Bartlett, O. Landon-Cardinal, D. Poulin, and Y.-K. Liu, *Nat. Commun.* **1**, 149 (2010); O. Landon-Cardinal and D. Poulin, *New J. Phys.* **14**, 085004 (2012).
- [9] D. Donoho, *IEEE Trans. Inf. Theory* **52**, 1289 (2006); E. Candès and T. Tao, *IEEE Trans. Inf. Theory* **52**, 5406 (2006); E. Candès and M. Wakin, *IEEE Signal Process. Mag.* **25**, 21 (2008); D. Gross, Y.-K. Liu, S. T. Flammia, S. Becker, and J. Eisert, *Phys. Rev. Lett.* **105**, 150401 (2010); S. T. Flammia, D. Gross, Y.-K. Liu, and J. Eisert, *New J. Phys.* **14**, 095022 (2012).
- [10] M. Lewenstein, B. Kraus, J. I. Cirac, and P. Horodecki, *Phys. Rev. A* **62**, 052310 (2000).
- [11] M. Bourennane, M. Eibl, C. Kurtsiefer, S. Gaertner, H. Weinfurter, O. Gühne, P. Hyllus, D. Brub, M. Lewenstein, and A. Sanpera, *Phys. Rev. Lett.* **92**, 087902 (2004).
- [12] A. Peres, *Phys. Rev. Lett.* **77**, 1413 (1996); M. Horodecki, P. Horodecki, and R. Horodecki, *Phys. Lett. A* **223**, 1 (1996).
- [13] E. Y.-Z. Tan, D. Kaszlikowski, and L. C. Kwek, *Phys. Rev. A* **93**, 012341 (2016); E. Verbanis, A. Martin, D. Rosset, C. C. W. Lim, R. T. Thew, and H. Zbinden, *Phys. Rev. Lett.* **116**, 190501 (2016).
- [14] D. Gottesman, *Phys. Rev. A* **54**, 1862 (1996); Ph.D. Thesis, California Institute of Technology, Pasadena, CA, 1997; [arXiv:quant-ph/9705052](https://arxiv.org/abs/quant-ph/9705052).
- [15] G. Tóth and O. Gühne, *Phys. Rev. Lett.* **94**, 060501 (2005).
- [16] M. Wieśniak and K. Maruyama, *Phys. Rev. A* **85**, 062315 (2012).
- [17] P. Kurzyński, T. Paterek, R. Ramanathan, W. Laskowski, and D. Kaszlikowski, *Phys. Rev. Lett.* **106**, 180402 (2011).
- [18] See Supplemental Material at <http://link.aps.org/supplemental/10.1103/PhysRevLett.117.210504> for entanglement criteria for more states and a more detailed description of the experimental setup containing Refs. [19,20].
- [19] P. Trojek, Ph.D. thesis, Ludwig-Maximilians-Universität München (2007); D. F. V. James, P. G. Kwiat, W. J. Munro, and A. G. White, *Phys. Rev. A* **64**, 052312 (2001).
- [20] N. Kiesel, Ph.D. thesis, Ludwig-Maximilians-Universität München (2007).
- [21] A. Zeilinger, M. A. Horne, H. Weinfurter, and M. Żukowski, *Phys. Rev. Lett.* **78**, 3031 (1997).
- [22] N. Kiesel, C. Schmid, U. Weber, G. Tóth, O. Gühne, R. Ursin, and H. Weinfurter, *Phys. Rev. Lett.* **95**, 210502 (2005); P. Walther, K. J. Resch, T. Rudolph, E. Schenck, H. Weinfurter, V. Vedral, M. Aspelmeyer, and A. Zeilinger, *Nature* (London) **434**, 169 (2005).
- [23] M. Barbieri, F. DeMartini, P. Mataloni, G. Vallone, and A. Cabello, *Phys. Rev. Lett.* **97**, 140407 (2006); A. Rossi, G. Vallone, A. Chiuri, F. De Martini, and P. Mataloni, *Phys. Rev. Lett.* **102**, 153902 (2009); W.-B. Gao, C.-Y. Lu, X.-C. Yao, P. Xu, O. Gühne, A. Goebel, Y.-A. Chen, C.-Z. Peng, Z.-B. Chen, and J.-W. Pan, *Nat. Phys.* **6**, 331 (2010); H.-X. Lu, L.-Z. Cao, J.-Q. Zhao, Y.-D. Li, and X.-Q. Wang, *Scientific reports* **4**, 4476 (2014).
- [24] C. Schwemmer, L. Knips, M. C. Tran, A. de Rosier, W. Laskowski, T. Paterek, and H. Weinfurter, *Phys. Rev. Lett.* **114**, 180501 (2015); N. Kiesel, C. Schmid, G. Tóth, E. Solano, and H. Weinfurter, *Phys. Rev. Lett.* **98**, 063604 (2007).
- [25] W. Wieczorek, C. Schmid, N. Kiesel, R. Pohlner, O. Gühne, and H. Weinfurter, *Phys. Rev. Lett.* **101**, 010503 (2008).

## Multipartite entanglement detection with minimal effort

Lukas Knips,<sup>1,2</sup> Christian Schwemmer,<sup>1,2,\*</sup> Nico Klein,<sup>1,2</sup> Marcin Wieśniak,<sup>3</sup> and Harald Weinfurter<sup>1,2</sup>

<sup>1</sup>Max-Planck-Institut für Quantenoptik, Hans-Kopfermann-Strasse 1, D-85748 Garching, Germany

<sup>2</sup>Fakultät für Physik, Ludwig-Maximilians-Universität, D-80797 München, Germany

<sup>3</sup>Institute of Theoretical Physics and Astrophysics,  
University of Gdańsk, PL-80-952 Gdańsk, Poland

### SUPPLEMENTAL MATERIAL

#### SM 1: CONSTRUCTING OPTIMAL CRITERIA

Criteria to detect genuine n-partite entanglement are specifically designed for individual states. Here, we will describe the construction of the criteria for the cluster state  $|\mathcal{C}_4\rangle \propto (|0000\rangle + |0011\rangle - |1100\rangle + |1111\rangle)$ . Furthermore, we will derive criteria for the symmetric four-qubit Dicke state  $|D_4^{(2)}\rangle \propto |1100\rangle + |1010\rangle + |1001\rangle + |0110\rangle + |0101\rangle + |0110\rangle$ , the four-qubit singlet state  $|\Psi_4\rangle \propto |0011\rangle + |1100\rangle - 1/2(|0110\rangle + |1001\rangle + |0101\rangle + |1010\rangle)$  and for  $|W_4\rangle \propto |1000\rangle + |0100\rangle + |0010\rangle + |0001\rangle$ .

#### A. Cluster state

The general procedure of finding the optimal entanglement criteria has already been described in the main text, together with an illustrative example for the GHZ state. According to the scheme given in the main text, Tab. I lists the non-vanishing correlations of the cluster state. One notices that the settings  $\mathcal{M}_{1133}$  and  $\mathcal{M}_{3311}$  are sufficient to infer six of the non-zero correlations of the state, indicated by bold letters in Tab. I. We use the corresponding operators to build the set  $\{\sigma_{1103}, \sigma_{1130}, \sigma_{0033}, \sigma_{0311}, \sigma_{3011}, \sigma_{3300}\}$ . Indeed, all six operators mutually commute and we are able to find cut-anticommutation relations for each bipartition. Therefore, it is possible to

TABLE I. The correlations of the cluster state. All other correlations vanish in the Pauli basis. Those with blue characters can be inferred from the setting  $\mathcal{M}_{1133}$ , the red colored ones from  $\mathcal{M}_{3311}$ .

$T_{0000}$	1	$T_{0033}$	1	$T_{0311}$	1	$T_{0322}$	-1
$T_{1103}$	-1	$T_{1130}$	-1	$T_{1212}$	-1	$T_{1221}$	-1
$T_{2112}$	-1	$T_{2121}$	-1	$T_{2203}$	1	$T_{2230}$	1
$T_{3011}$	1	$T_{3022}$	-1	$T_{3300}$	1	$T_{3333}$	1

\* Current Address: IBM Research GmbH, Säumerstrasse 4, CH-8803 Rüschlikon, Switzerland.

disprove separability along each cut by using an inequality like Eq. (3) of the main text. For example, the first two operators  $A|BCD$ -anticommute with the last two, so  $A|BCD$ -nonseparability is proven by violation of

$$\mathcal{W}_{A|BCD}^{C_4} = (T_{1103}^2 + T_{3011}^2) \leq_{A|BCD}^{\text{SEP}} 1. \quad (1)$$

or any other combination of squared expectation values of the operators of these two groups. Averaging over all possible combinations of  $A|BCD$ -anticommuting operators of the given set and normalization leads to

$$\mathcal{W}_{A|BCD}^{C_4} = \frac{1}{2} \left[ \frac{1}{2} (T_{1103}^2 + T_{1130}^2) + \frac{1}{2} (T_{3011}^2 + T_{3300}^2) \right] \leq_{A|BCD}^{\text{SEP}} \frac{1}{2}. \quad (2)$$

Permutations of the indices used in Ineq. (2) are used for other 1:3 cuts and for  $AB|CD$ . For cuts  $\mathcal{B}^* = \{AC|BD, AD|BC\}$ , the separability is refuted more efficiently with

$$\mathcal{W}_{\mathcal{B}^*}^{C_4} = \frac{1}{6} \left[ T_{0033}^2 + T_{0311}^2 + T_{1103}^2 + T_{1130}^2 + T_{3011}^2 + T_{3300}^2 \right] \leq_{\mathcal{B}^*}^{\text{SEP}} \frac{1}{2}. \quad (3)$$

The combined entanglement witness for the cluster state uses all six operators. Therefore, the anticommutativity graphs consist of six vertices, each representing one of the operators of the given set. For no bipartition ( $r = 0$ ), there are no edges and ‘1’s can be assigned to all vertices. Thus, this graph leads to the sum  $G_0 = v_{0311} + v_{3011} + v_{3300} + v_{1103} + v_{1130} + v_{0033}$ . The graph of the bipartition  $AB|CD$  is depicted in Fig. 1 a) and b), where the dashed lines indicate the cut-anticommutation relations. Because neither  $\sigma_{0033}$  nor  $\sigma_{3300}$   $AB|CD$ -anticommutes with any of the operators, they can be assigned ‘1’ in any case. Besides this, one can distribute ‘1’s according to Fig. 1 a) leading to  $G_{AB|CD}^{(1)} = v_{3300} + v_{1103} + v_{1130} + v_{0033}$ . A different distribution of ‘1’s results in  $G_{AB|CD}^{(2)} = v_{0311} + v_{3011} + v_{3300} + v_{0033}$ , see Fig. 1 b). The operators  $\sigma_{3300}$  and  $\sigma_{0033}$ , appearing in both distributions in this cut, seem to be superior [not connected with any operator in Fig. 1 a) and b)] to the other operators. In contrast, e.g., for the bipartition  $A|BCD$ , the operators  $\sigma_{0311}$  and  $\sigma_{0033}$  do not cut-anticommute with any other operators and are, thus, superior for this bipartition. Considering the bipartition  $AC|BD$ , as it

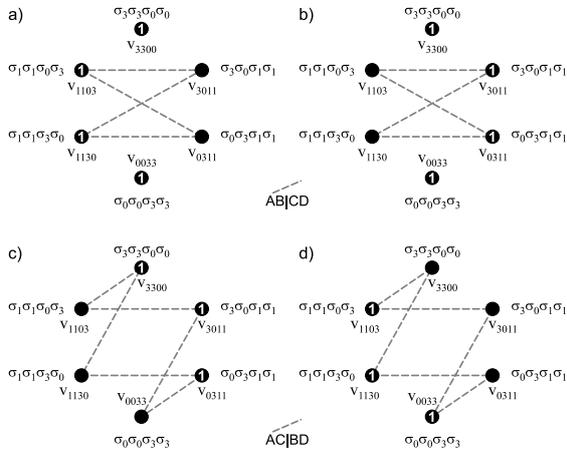


FIG. 1. The cut-anticommutativity graphs for the cluster state are shown for  $AB|CD$  (upper row) and  $AC|BD$  (lower row). For those bipartitions, for which the cut-anticommutation relations are indicated by dashed lines, two possible distributions of values ‘1’s are depicted in the respective columns. The weights of the operators are denoted by, e.g.,  $v_{1103}$ .

is shown in Fig. 1 c) and d), the cut-anticommutativity relations are such that each operator cut-anticommutes with two operators. Thus, by considering all seven bipartitions, all six operators behave similarly, suggesting equal weights for all operators, i.e. we introduce  $\alpha$  with  $\alpha = v_{0311} = v_{3011} = \dots = v_{0033}$ . Without loss of generality, we can set  $\alpha = 1$ . Because four ‘1’s can be distributed for the bipartition  $AB|CD$ , our chosen weights result in  $G_{AB|CD}^{(1)} = G_{AB|CD}^{(2)} = 4$ . In conclusion, we find for the anticommutativity graph  $G_0 = 6\alpha = 6$  and by maximizing over all bipartitions  $G = 4\alpha = 4$ , leading to the criterion Eq. (8) in the main text.

### B. Dicke state

Criteria for the symmetric four-qubit Dicke state with two excitations  $|D_4^{(2)}\rangle \propto |1100\rangle + |1010\rangle + |1001\rangle + |0110\rangle + |0101\rangle + |0110\rangle$  can also be derived easily. Measurement of  $\mathcal{M}_{1111}$  enables to deduce the expectation values of the operators  $\sigma_{0011}$ ,  $\sigma_{0101}$ ,  $\sigma_{0110}$ ,  $\sigma_{1001}$ ,  $\sigma_{1010}$ ,  $\sigma_{1100}$ , and  $\sigma_{1111}$ . Adding the measurement of  $\mathcal{M}_{2222}$  results in a set of commuting operators also containing cut-anticommuting operators for every cut. The possibility of one-versus-three qubit separability for the Dicke state is eliminated by a violation of

$$\mathcal{W}_{\{1:3\}}^{D_4^{(2)}} = \frac{1}{2} (T_{1111}^2 + T_{2222}^2) \leq_{\text{SEP}} \frac{1}{2}. \quad (4)$$

Since only  $T_{1111}^2$  and  $T_{2222}^2$  are 1 for  $|D_4^{(2)}\rangle$  while all further expectation values of operators deduced by those

two are less, one cannot enhance the criterion by averaging over additional values as in, e.g., Eq. (6) of the main text. To rule out separability along cut, say,  $AB|CD$  the inequality

$$\mathcal{W}_{AB|CD}^{D_4^{(2)}} = \frac{1}{2} \left[ \frac{1}{4} (T_{1010}^2 + T_{1001}^2 + T_{0110}^2 + T_{0101}^2) + T_{2222}^2 \right] \leq_{\text{SEP}} \frac{1}{2} \quad (5)$$

has to be violated. Note that even the ideal Dicke state will score only  $\mathcal{W}_{AB|CD}^{D_4^{(2)}} = 13/18 \approx 0.72$ . By permutations of the indices in Eq. (5) one obtains the criteria  $\mathcal{W}_{AC|BD}^{D_4^{(2)}}$  and  $\mathcal{W}_{AD|BC}^{D_4^{(2)}}$ . The combined witness for the Dicke state reads

$$\mathcal{W}^{D_4^{(2)}} = \frac{1}{10} \left[ 2 (T_{1111}^2 + T_{2222}^2) + T_{1100}^2 + T_{0011}^2 + T_{0101}^2 + T_{1010}^2 + T_{1001}^2 + T_{0110}^2 + T_{2200}^2 + T_{0022}^2 + T_{0202}^2 + T_{2020}^2 + T_{2002}^2 + T_{0220}^2 \right] \leq_{\text{SEP}} \frac{4}{5}, \quad (6)$$

which is difficult to violate experimentally, as even the ideal Dicke state scores only 14/15. Since our data cannot show a significant proof of genuine fourpartite entanglement, all possible biseparations have to be ruled out.

### C. Singlet state

For the four-qubit singlet state  $|\Psi_4\rangle \propto |0011\rangle + |1100\rangle - 1/2 (|0110\rangle + |1001\rangle + |0101\rangle + |1010\rangle)$ , the set of correlations for the operators deduced from the measurement settings of  $\mathcal{M}_{1111}$  and  $\mathcal{M}_{2222}$  are similar as for  $|D_4^{(2)}\rangle$ . Thus, the criteria  $\mathcal{W}_{\{1:3\}}^{D_4^{(2)}}$  and  $\mathcal{W}_{AB|CD}^{D_4^{(2)}}$  also apply here while the criteria for the bipartitions  $AC|BD$  and  $AD|BC$  are slightly modified since  $T_{0011}$  and  $T_{1100}$  reach a value of only 1/3 for the singlet state and are therefore left out. The combined witness for the singlet state reads

$$\mathcal{W}^{\Psi_4} = \frac{1}{10} \left[ 4T_{1111}^2 + 2T_{3333}^2 + T_{3003}^2 + T_{0330}^2 + T_{3030}^2 + T_{0303}^2 \right] \leq_{\text{SEP}} \frac{3}{5}. \quad (7)$$

The ideal state  $|\Psi_4\rangle$  scores 7/9. The experimentally prepared state could be proven to be genuinely fourpartite entangled with high significance.

### D. W state

Because correlations of  $|W_4\rangle \propto |0001\rangle + |0010\rangle + |0100\rangle + |1000\rangle$  are (besides of  $T_{3333} = -1$ ) at most only  $\pm 1/2$  and thus too weak for a robust combined criterion,

we again have to find criteria for the different bipartitions in order to build sensitive indicators. Entanglement along the cut  $A|BCD$  can be detected by violation of

$$\mathcal{W}_{A|BCD}^{W_4} = \frac{1}{2} \left[ T_{3333}^2 + \frac{1}{3} (T_{1001}^2 + T_{1010}^2 + T_{1100}^2) \right] \leq_{\text{SEP}_{A|BCD}} \frac{1}{2}. \quad (8)$$

Criteria for the other one-versus-three-separations are obtained by permuting the parties in  $\mathcal{W}_{A|BCD}^{W_4}$ .  $AB|CD$ -separability can be ruled out by the criterion

$$\mathcal{W}_{AB|CD}^{W_4} = \frac{1}{2} \left[ T_{3333}^2 + \frac{1}{4} (T_{0101}^2 + T_{0110}^2 + T_{1001}^2 + T_{1010}^2) \right] \leq_{\text{SEP}_{AB|CD}} \frac{1}{2}, \quad (9)$$

whose permutations lead to the criteria to eliminate separability along  $AC|BD$  and  $AD|BC$ .

## SM 2: SETUP AND MEASUREMENT

The general idea of the experimental setup was already explained in the main text of this letter. Here, we want to focus on the details of both the spontaneous parametric down conversion (SPDC) source and the interferometers which allow to perform a complete tomographic analysis of the prepared states.

The experiment starts with the generation of pairs of polarization entangled photons in the state

$$\frac{1}{\sqrt{2}} (|HH\rangle + e^{i\phi}|VV\rangle) \quad (10)$$

as obtained from the process of spontaneous parametric down conversion (SPDC) [19]. The SPDC source consists of a pair of crossed type I cut  $\beta$ -Barium-Borate (BBO) crystals that are pumped by a continuous wave laser diode at a central wavelength of 402 nm, with approximately 60 mW of pump power, and linear polarization of 45°. The phase  $\phi$  between the emitted photons can be set by means of an Yttrium Vanadate crystal (YVO<sub>4</sub>) in front of the BBO crystals, see Fig. 2 of the main text. An additional half waveplate set at an angle  $\theta$  enables to rotate the polarization of the second photon to any linear polarization, leading to the state

$$\frac{1}{\sqrt{2}} (|H\rangle(\cos(2\theta)|H\rangle + \sin(2\theta)|V\rangle) + e^{i\phi}|V\rangle(\sin(2\theta)|H\rangle - \cos(2\theta)|V\rangle)) \quad (11)$$

as mentioned in the main text.

The emitted photons are spectrally filtered by interference filters with a bandwidth of 5 nm. Spatial filtering is achieved by coupling the pairs into two single mode fibers

that are connected to one of the input ports of each of the two interferometers.

In principle, a Mach-Zehnder type interferometer as given in the main text suits the purpose to analyze the phase between the two spatial modes  $a$  and  $b$ . However, in terms of phase stability, a Sagnac configuration is preferable to a Mach-Zehnder type interferometer if one wants to avoid using an active stabilization scheme, as is the case here. Therefore, in our experiments we resorted

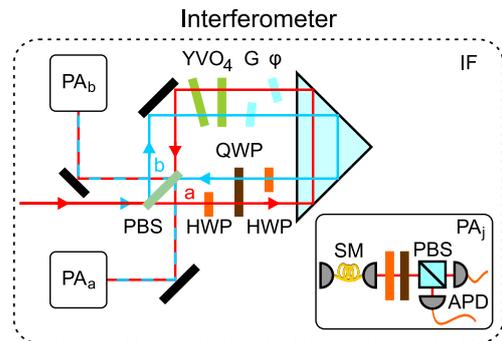


FIG. 2. For state analysis, it is necessary to be able to characterize the phase  $\phi$  between the modes  $a$  and  $b$ . In principle, a Mach-Zehnder interferometer can be used to measure the (relative) phase between the modes  $a$  and  $b$ , but in terms of stability a Sagnac configuration is preferable. Therefore, for the experimental realization of the interferometer, we resorted to a Sagnac loop, where instead of two polarizing beam splitters (PBS) a single one that is hit by the photons twice is sufficient. Half (HWP) and quarter waveplates (QWP) are used for state analysis and enable tomographic analysis of the experimentally prepared state. A pair of YVO<sub>4</sub> crystals is used for the compensation of unwanted additional phase shifts resulting from the total internal reflection at the prism. Two thin glass plates (G) of which one is motorized is used to balance the interferometer arms and to set arbitrary relative phases  $\phi$  between the two arms.

to using an unstabilized Sagnac interferometer as shown in Fig. 2. Although all optical components were mounted carefully to avoid birefringence as induced by mechanical stress, still an unwanted polarization dependent phase shift due to the total internal reflection at the prism remains (Goos-Hänchen effect). This phase shift would for example rotate diagonally polarized light to elliptically polarized light. In order to compensate this phase shift, two YVO<sub>4</sub> crystals with their optical axis crossed were utilized (a zero-order configuration). A motorized thin piece of glass ( $\approx 120 \mu\text{m}$  thick) was applied to set the phase difference between the two interferometer paths  $a$  and  $b$  to any wanted value. The interferometer could be balanced by a second piece of glass plate of the same thickness. Please note that the second glass plate was aligned such that the relative phase between the modes is compensated, i.e., the transformation of the input polarizing beam splitter was  $|H\rangle \rightarrow |Ha\rangle$  and  $|V\rangle \rightarrow |Vb\rangle$

for photons entering the interferometer. The wave plates inside the interferometer are required to analyze the polarization degree of freedom. In order to keep the setup as compact as possible, the quarter waveplate covers both spatial modes. The polarization analysis in both outputs of the interferometer enables to measure the path degree of freedom, i.e., allowing to distinguish between e.g.  $|a+b\rangle$  and  $|a-b\rangle$ . The task is now to find angle settings for the waveplates inside and outside of the interferometer such that a tomographically complete set of projection measurements is obtained. Let us therefore review the (unitary) transformations that are induced by the respective waveplates. For the half waveplate the transformation is

$$U_{\text{HWP}}(\theta) = \begin{pmatrix} \cos(2\theta) & \sin(2\theta) \\ \sin(2\theta) & -\cos(2\theta) \end{pmatrix} \quad (12)$$

and correspondingly for the quarter waveplate

$$U_{\text{QWP}}(\theta) = \begin{pmatrix} \cos(\theta)^2 - i \sin(\theta)^2 & (1+i) \cos(\theta) \sin(\theta) \\ (1+i) \cos(\theta) \sin(\theta) & -i \cos(\theta)^2 + \sin(\theta)^2 \end{pmatrix} \quad (13)$$

with  $|H\rangle = (1,0)^T$  and  $|V\rangle = (0,1)^T$ . As common to most multiqubit experiments, we choose to measure in the eigenbases of all combinations of local Pauli bases. In order to make the general procedure for finding the angle settings more illustrative, let us discuss the  $\mathcal{M}_{31}$  basis as an example. Measuring in this basis means that projections onto its eigenvectors have to be performed, i.e., on  $|H(a+b)\rangle$ ,  $|H(a-b)\rangle$ ,  $|V(a+b)\rangle$  and  $|V(a-b)\rangle$ .

If one wants to project onto  $|H(a+b)\rangle$  for example, the waveplates inside the interferometer have to transform the state just behind the polarizing input beam splitter such that  $|Ha\rangle \rightarrow e^{i\xi_a}|Ha\rangle$  and  $|Hb\rangle \rightarrow e^{i\xi_b}|Vb\rangle$  with respective phases  $\xi_a$  and  $\xi_b$ . One possible choice would be

$$|Ha\rangle \rightarrow U_{\text{QWP}}(0) U_{\text{HWP}}(0) |Ha\rangle = |Ha\rangle, \quad (14)$$

$$|Hb\rangle \rightarrow U_{\text{QWP}}(0) U_{\text{HWP}}\left(-\frac{\pi}{4}\right) |Hb\rangle = i|Vb\rangle, \quad (15)$$

where the identity operation acting on the spatial mode is omitted. Then, the detection of a right circular polarized photon by the polarization analysis  $\text{PA}_a$  corresponds to a successful projection on  $|H(a+b)\rangle$ . On the other hand a left circular polarized photon in  $\text{PA}_a$  would correspond to  $|H(a-b)\rangle$ . Consequently, the polarization analysis  $\text{PA}_b$  in the other output of the interferometer allows for projection on  $|V(a+b)\rangle$  and  $|V(a-b)\rangle$ . Please note that one has to trigger on coincidence counts then, i.e., one photon from each interferometer has to be detected. For four qubits, i.e. two interferometers, this scheme then yields  $3^4 = 81$  different measurement settings, where in each setting  $2^4 = 16$  projection measurements are performed. In our experiment we used fiber coupled single photon counting modules (SPCM from Perkin Elmer) that were connected to a coincidence electronic with a coincidence window of 10 ns. All in all  $81 \times 16 = 1296$  different projectors were measured and a tomographically overcomplete set of data is obtained, which is processed with the method described in [20]. The angles for all the measurement settings can be seen in Tab. II. It has to be noted that the given angles are not the only possible choice to obtain a tomographically (over-)complete set of projectors.

TABLE II. The angles of the waveplates of the interferometer ( $\text{HWP}_{\text{IF},1}$ ,  $\text{QWP}_{\text{IF}}$ ,  $\text{HWP}_{\text{IF},2}$ ) and of the two polarization analyses ( $\text{HWP}_A$ ,  $\text{QWP}_A$ ;  $\text{HWP}_B$ ,  $\text{QWP}_B$ ) to perform the given projections. ‘HA’, ‘VA’, ‘HB’, ‘VB’ denote the detectors for the transmitted (‘H’) and reflected light (‘V’) of the PBS of the polarization analysis in the output modes A and B, respectively. For example, an event of the detector ‘HA’ while measuring in the basis  $\mathcal{M}_1 \otimes \mathcal{M}_1$  corresponds to a successful projection onto the state  $|P(a+b)\rangle$ . Please note that all angles are referenced with respect to mode  $a$  which means that for calculating the transformations induced for light in mode  $b$  a minus sign has to be added.

Basis	Interferometer			Polarisation analysis				Projectors			
	$\text{HWP}_{\text{IF},1}$	$\text{QWP}_{\text{IF}}$	$\text{HWP}_{\text{IF},2}$	$\text{HWP}_A$	$\text{QWP}_A$	$\text{HWP}_B$	$\text{QWP}_B$	HA	VA	HB	VB
$\mathcal{M}_1 \otimes \mathcal{M}_1$ ( $\mathcal{M}_{11}$ )	$\frac{\pi}{8}$	0	$\frac{\pi}{8}$	0	$\frac{\pi}{4}$	0	$-\frac{\pi}{4}$	P(a+b)	P(a-b)	M(a+b)	M(a-b)
$\mathcal{M}_1 \otimes \mathcal{M}_2$ ( $\mathcal{M}_{12}$ )	$\frac{\pi}{8}$	0	$\frac{\pi}{8}$	$-\frac{\pi}{8}$	0	$-\frac{\pi}{8}$	0	P(a+ib)	P(a-ib)	M(a+ib)	M(a-ib)
$\mathcal{M}_1 \otimes \mathcal{M}_3$ ( $\mathcal{M}_{13}$ )	$\frac{\pi}{8}$	0	$\frac{\pi}{8}$	0	0	$\frac{\pi}{4}$	0	Pa	Pb	Ma	Mb
$\mathcal{M}_2 \otimes \mathcal{M}_1$ ( $\mathcal{M}_{21}$ )	0	$\frac{\pi}{4}$	0	0	$-\frac{\pi}{4}$	0	$\frac{\pi}{4}$	R(a+b)	R(a-b)	L(a+b)	L(a-b)
$\mathcal{M}_2 \otimes \mathcal{M}_2$ ( $\mathcal{M}_{22}$ )	0	$\frac{\pi}{4}$	0	$\frac{\pi}{8}$	0	$\frac{\pi}{8}$	0	R(a+ib)	R(a-ib)	L(a+ib)	L(a-ib)
$\mathcal{M}_2 \otimes \mathcal{M}_3$ ( $\mathcal{M}_{23}$ )	0	$\frac{\pi}{4}$	0	0	0	$\frac{\pi}{4}$	0	Ra	Rb	La	Lb
$\mathcal{M}_3 \otimes \mathcal{M}_1$ ( $\mathcal{M}_{31}$ )	0	0	$\frac{\pi}{4}$	0	$\frac{\pi}{4}$	0	$-\frac{\pi}{4}$	H(a+b)	H(a-b)	V(a+b)	V(a-b)
$\mathcal{M}_3 \otimes \mathcal{M}_2$ ( $\mathcal{M}_{32}$ )	0	0	$\frac{\pi}{4}$	$-\frac{\pi}{8}$	0	$-\frac{\pi}{8}$	0	H(a+ib)	H(a-ib)	V(a+ib)	V(a-ib)
$\mathcal{M}_3 \otimes \mathcal{M}_3$ ( $\mathcal{M}_{33}$ )	0	0	$\frac{\pi}{4}$	0	0	$\frac{\pi}{4}$	0	Ha	Hb	Va	Vb

#### SM 4: RESULTS

We prepared and characterized 13 states belonging to the states given by  $|\Psi(\theta, \phi)\rangle$  in the main text, including the GHZ and Cluster state. For all states, we carried out full quantum state tomography. Fig. 3 shows the experimental density matrices of the GHZ state, the cluster state, and for another GHZ-type state  $|\text{GHZ}'\rangle =$

$(|0011\rangle - |1100\rangle)/\sqrt{2}$ . From the density matrices, the fidelity with the theoretically expected states could be inferred, see Tab. IV. The fidelities of the prepared states compared with the respective target state were above 95.8% in all cases. The correlation values used for the witnesses are listed in Tab. III for the GHZ and Cluster state.

TABLE III. Experimentally determined correlation values for the GHZ and Cluster state, which are used to calculate the values of the two measurement witnesses. For the GHZ state, the measurement settings  $\mathcal{M}_{1221}$  and  $\mathcal{M}_{3333}$  were used and for the Cluster state  $\mathcal{M}_{1133}$  and  $\mathcal{M}_{3311}$ .

GHZ		Cluster	
$T_{3333}$	$0.982 \pm 0.003$	$T_{3300}$	$0.987 \pm 0.002$
$T_{3300}$	$0.993 \pm 0.002$	$T_{3011}$	$0.986 \pm 0.003$
$T_{0033}$	$0.988 \pm 0.002$	$T_{0311}$	$0.974 \pm 0.003$
$T_{3003}$	$0.963 \pm 0.004$	$T_{1130}$	$-0.945 \pm 0.006$
$T_{0330}$	$0.969 \pm 0.004$	$T_{1103}$	$-0.934 \pm 0.006$
$T_{3030}$	$0.972 \pm 0.004$	$T_{0033}$	$0.989 \pm 0.002$
$T_{0303}$	$0.960 \pm 0.005$		
$T_{1221}$	$-0.925 \pm 0.006$		

For all prepared states, at least one of the two combined witnesses  $\mathcal{W}^{\text{GHZ}}$  and  $\mathcal{W}^{\text{C}_4}$ , which both could be determined from two measurement settings only, lies above the respective threshold (7/11 for the GHZ criterion or

2/3 for the cluster criterion). Therefore, genuine four-partite entanglement could be proven for all considered states.

TABLE IV. Characterization of 13 states given by  $|\Psi(\theta, \phi)\rangle$  in the main text with  $\phi = \pi$ . The fidelities with the respective target states were determined from the experimental density matrices as obtained via quantum state tomography. The values for the entanglement criteria  $\mathcal{W}^{\text{GHZ}}$  and  $\mathcal{W}^{c_4}$  as presented in the main text, however, were inferred from two measurement settings only. For all prepared states, genuine four-partite entanglement can be proved by at least one of the two criteria. Successful entanglement detection of the respective criterion is indicated by bold letters.

Name	$\theta$	Fidelity $\mathcal{F}$	$\mathcal{W}^{\text{GHZ}} \leq_{\text{BISEP}} \frac{7}{11}$	$\mathcal{W}^{c_4} \leq_{\text{BISEP}} \frac{2}{3}$
GHZ	0	$0.958 \pm 0.004$	<b><math>0.916 \pm 0.005</math></b>	$0.333 \pm 0.002$
	$\frac{\pi}{48}$	$0.959 \pm 0.004$	<b><math>0.894 \pm 0.006</math></b>	$0.387 \pm 0.005$
	$\frac{\pi}{24}$	$0.958 \pm 0.004$	<b><math>0.828 \pm 0.007</math></b>	$0.509 \pm 0.007$
	$\frac{\pi}{16}$	$0.965 \pm 0.003$	<b><math>0.740 \pm 0.007</math></b>	<b><math>0.685 \pm 0.007</math></b>
	$\frac{\pi}{12}$	$0.963 \pm 0.003$	<b><math>0.644 \pm 0.006</math></b>	<b><math>0.835 \pm 0.006</math></b>
	$\frac{5\pi}{48}$	$0.963 \pm 0.003$	$0.603 \pm 0.004$	<b><math>0.918 \pm 0.005</math></b>
Cluster	$\frac{\pi}{8}$	$0.962 \pm 0.003$	$0.590 \pm 0.004$	<b><math>0.940 \pm 0.004</math></b>
	$\frac{7\pi}{48}$	$0.959 \pm 0.004$	$0.608 \pm 0.006$	<b><math>0.886 \pm 0.007</math></b>
	$\frac{\pi}{6}$	$0.959 \pm 0.004$	<b><math>0.679 \pm 0.008</math></b>	<b><math>0.761 \pm 0.008</math></b>
	$\frac{3\pi}{16}$	$0.958 \pm 0.003$	<b><math>0.746 \pm 0.007</math></b>	$0.640 \pm 0.008$
	$\frac{5\pi}{24}$	$0.960 \pm 0.004$	<b><math>0.820 \pm 0.007</math></b>	$0.497 \pm 0.007$
	$\frac{11\pi}{48}$	$0.963 \pm 0.004$	<b><math>0.890 \pm 0.006</math></b>	$0.375 \pm 0.004$
GHZ'	$\frac{\pi}{4}$	$0.967 \pm 0.004$	<b><math>0.927 \pm 0.005</math></b>	$0.330 \pm 0.002$

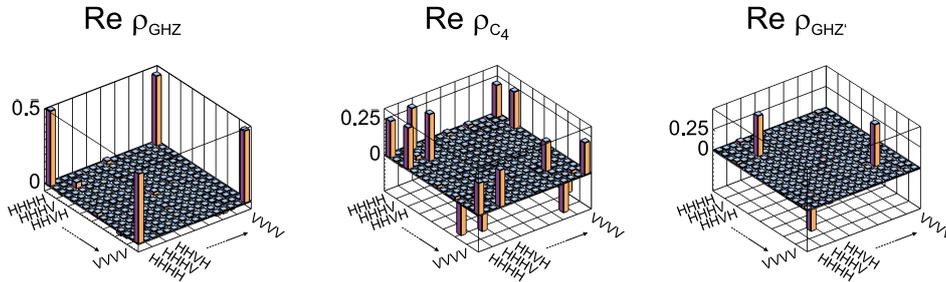


FIG. 3. Real parts of the experimental density matrices of (a) the GHZ state ( $\theta = 0$ ,  $\phi = \pi$ ), (b) the cluster state ( $\theta = \pi/8$ ,  $\phi = \pi$ ), and (c) the GHZ' state ( $\theta = \pi/4$ ,  $\phi = \pi$ ).

- [19] P. Trojek, Ph.D. thesis, Ludwig-Maximilians-Universität München (2007); D. F. V. James, P. G. Kwiat, W. J. Munro, and A. G. White, Phys. Rev. A **64**, 052312 (2001).  
[20] N. Kiesel, Ph.D. thesis, Ludwig-Maximilians-Universität München (2007).

### 4.2.6 Application for Higher Dimensional Systems

For systems with  $n$  qubits, the correlation tensor formulation as given in Eq. (2.16) has proven to be a powerful description. Unfortunately, however, this description can not be used for multiqutrit systems. Here, a nonlinear entanglement witness similar to the one presented in [P3], reprinted in Sec. 4.2.5, is derived for the general class of multiqutrit graph states, i.e., systems with  $n$  three-dimensional subsystems, which interacted via generalized controlled-phase gates.

The entanglement criterion for multiqubit systems as described in Eq. (4.10), and more generally in Eq. (4.11), serves as inspiration for the multiqutrit case. Motivated by these inequalities, entanglement is most easily detected for states, which have several correlation values large in modulus, ideally of modulus 1. Hence, graph states are an ideal testbed for these entanglement certifiers.

Qubit measurements using Pauli matrices  $\sigma_0$ ,  $\sigma_x$ ,  $\sigma_y$ , and  $\sigma_z$  have several neat properties. One of those is the fact that the expectation value, say,  $\langle \sigma_x \rangle$ , can be used to deduce the probability distribution of both possible outcomes [P4]. Furthermore, their eigenbases are mutually unbiased, leading to independent elements of the correlation tensor, which cannot be directly generalized for higher dimensional systems. Instead of using the Pauli matrices for parametrization of the quantum state  $\rho$ , one can now use a set of 8 Heisenberg-Weyl matrices  $\{h_i\}$  as they provide mutually unbiased bases [P4]. Subsequently, instead of complementarity relations between expectation values of Pauli matrices [189], one obtains complementarity relations between the expectation values of Heisenberg-Weyl matrices. For example, the four qutrit GHZ state defined as

$$|\text{GHZ}_{3,4}\rangle = \frac{1}{\sqrt{3}} \sum_{j=0}^2 |jjjj\rangle \quad (4.16)$$

exhibits a correlation when all four parties perform measurements in basis

$$h_2 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix} \quad (4.17)$$

as well as if two observers measure in  $h_1$  and two in  $h_5 = h_1^\dagger$  with

$$h_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \exp(i\frac{2\pi}{3}) & 0 \\ 0 & 0 & \exp(i\frac{4\pi}{3}) \end{pmatrix}. \quad (4.18)$$

Separability along a specific cut, for example for the bipartition  $AB|CD$  can be excluded by the criterion

$$|\langle h_1 h_5 h_1 h_5 \rangle|^2 + |\langle h_2 h_2 h_2 h_2 \rangle|^2 \leq_{\text{SEP}_{AB|CD}} 1, \quad (4.19)$$

for which the four qutrit GHZ state obtains a value of 2.

In the following section, which is a reprint of [P4], entanglement criteria are derived and explained for the examples of the above mentioned four qutrit GHZ state as well as the four qutrit cluster state. Furthermore, the correlation tensor parametrization of multiqutrit states based upon above mentioned Heisenberg-Weyl matrices is introduced, allowing to construct complementarity relations. An open question for further research is still the generalization for multiqutrit systems with qudits of dimensions  $d > 3$ .

This work is based on the same initial idea from Marcin Wieśniak as [P3]. Due to my refinement of the method in aforementioned publication, he could extend our method together with Krzysztof Rosołek to the case of multiqutrit systems. The first manuscript was written by Marcin Wieśniak and Krzysztof Rosołek, and later edited also by myself.

## Quadratic Entanglement Criteria for Qutrits

Krzysztof Rosolek,<sup>1</sup> Marcin Wieśniak,<sup>2</sup> and Lukas Knips<sup>3,4</sup>

<sup>1</sup>*Institute of Theoretical Physics and Astrophysics,  
Faculty of Mathematics, Physics, and Informatics,  
University of Gdańsk, ul. Wita Swosza 57, 80-308 Gdańsk, Poland*

<sup>2</sup>*Institute of Informatics, Faculty of Mathematics, Physics, and Informatics, University of Gdańsk,  
ul. Wita Swosza 57, 80-952 Gdańsk, Poland*

<sup>3</sup>*Max-Planck-Institut für Quantenoptik, Hans-Kopfermann-Strasse 1, D-85748 Garching, G, Germany*

<sup>4</sup>*Department für Physik, Ludwig-Maximilians-Universität, D-80797 München, Germany*

The problem of detecting non-classical correlations of states of many qudits is incomparably more involved than in a case of qubits. The reason is that for qubits we have a convenient description of the system by the means of the well-studied correlation tensor. Simply, the complete information about the state can be encoded in mean values of dichotomic measurements. We demonstrate that for three-dimensional quantum subsystems we are able to formulate nonlinear entanglement criteria of the state with existing formalisms. We also point out where the idea for constructing these criteria fails for higher-dimensional systems, which poses well-defined open questions.

### I. INTRODUCTION

Quantum correlations are known to be capable of outperforming classical ones. While separable states can be perfectly correlated in one way at a time, entangled ones may reveal perfect correlations, say, whenever the same quantity is measured by two observers. This observation has led to a serious debate about the most fundamental aspects of nature. First, Einstein, Podolsky, and Rosen [1] have asked if quantum mechanics can be supplemented with additional, hidden parameters, and later it was answered that if it was indeed so, these parameters would need to go beyond certain reasonable requirements, such as locality [2], or noncontextuality [3].

The Bell theorem [2] has consequences of not only philosophical nature, but has also found applications in certain communication tasks. In particular, having a Bell inequality violated by a quantum state is equivalent to an advantage in a distributed computing [4]. Specifically, if protocol users share an entangled state, they can achieve a higher probability of locally getting the correct value of a certain function than when they are allowed only to communicate classically. The role of the Bell theorem has been also pointed out in the context of, e.g., cryptography [5].

Therefore, schemes of entanglement detection have gathered a lot of attention for both fundamental and practical reasons. The task is very simple for pure states, which practically never occur in a real life. However, for mixed states it is still an open question. One method is to apply a positive, but not a completely positive map to one of subsystems [6, 7]. This should drive an entangled state out of the set of physically admissible density operators. By the Jamiolkowski-Choi isomorphism [8], we can equivalently use an entanglement witness, a composite observable taking negative mean values only for entangled states. In this manner, we can certify all forms of entanglement, but we do not know all the non-completely positive maps. In order to make entanglement detection schemes more efficient, nonlinear criteria were

introduced. They appeared also in particular context of necessary conditions on states to violate Bell inequalities [9–11]. A state can violate Werner-Wolf-Weinfurter-Żukowski-Brukner inequalities only if (but not necessarily if) certain of its squared elements of the correlation tensor add up to more than 1. A similar condition appeared in the context of so-called geometrical inequalities [12], which treat correlations of the system as multidimensional vector not belonging to a convex set of local realistic models. This approach resulted in geometrical entanglement criteria [13], which are highly versatile, and quadratic ones, particularly easy to construct [14–16].

Up to date, these methods turn out to be successful mainly for collections of qubits, as their states are conveniently described by the means of the correlation tensor. The deficit of Bell inequalities and entanglement criteria for higher-dimensional constituents of quantum systems follow also from our inability to generalize this tool. Pauli matrices, the foundation of this achievement, have many interesting properties, each contributing to the success. They are Hermitian, unitary, traceless, for individual subsystems their measurements are complete (except for the unit matrix), meaning that the individual mean values contain the full information about the statistics of outcomes, and they have unbiased bases as their eigenbases. In contrast, one of the straight-forward generalizations, the Gell-Mann matrices, do not satisfy any commutativity relations. If we used them to create a correlation tensor, its elements would not be independent from one another.

In this contribution we show that the notions known for the formalism of the tensor product for multiqubit states can be straight-forwardly stretched to qutrits, when we associate complex root of infinity to local measurement outcomes. In particular, this generalized tensor product is a subject to linear and quadratic bounds. Basing on these bounds, we can derive quadratic (and geometrical) entanglement criteria. For higher-dimensional systems, this is still an open challenge.

## II. FORMALISM OF MANY-QUBIT STATES

As we have already mentioned, the success of describing and analyzing the states of many qubits is due to the particularly convenient representation through a correlation tensor. Its elements are mean values of tensor products of Pauli matrices,  $T_i = \langle \sigma_i \rangle$ ,  $\sigma_i = \sigma_{i_1}^{[1]} \otimes \sigma_{i_2}^{[2]} \otimes \dots$ ,  $\bar{i} = i_1, i_2, \dots, i_N$ , and

$$\begin{aligned} \sigma_0 &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \\ \sigma_2 &= \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \end{aligned} \quad (1)$$

$$(2)$$

Operators  $\sigma_{\bar{i}}$  form an orthonormal basis,  $(\sigma_{\bar{i}}, \sigma_{\bar{j}}) = \text{tr} \sigma_{\bar{i}} \sigma_{\bar{j}} = 2^N \delta_{\bar{i}, \bar{j}}$ . This orthogonality can have three different reasons. When either  $\sigma_{\bar{i}}$  or  $\sigma_{\bar{j}}$  is the unit matrix, the other operator is traceless. When  $\sigma_{\bar{i}}$  and  $\sigma_{\bar{j}}$  commute, but differ from each other and neither of them is the unit matrix, their eigenvalues are distributed in such a way that their product adds up to zero. Finally, when they do not commute, they anticommute and their eigenbases can be chosen to be unbiased, i.e., the scalar product between any vector from one basis and any one from the other is constant in modulo. For a given state  $\rho$ , let the correlation tensor be a set of averages  $\{T_i\} = \{\text{tr} \rho \sigma_i\}$ . Naturally,  $T_{00\dots 0} = 1$ , but also for a single qubit we have the pronounced complementarity relation [17],

$$\sum_{i=1}^3 \langle \sigma_i \rangle^2 \leq 1. \quad (3)$$

This relation can be straight-forwardly generalized to any set of mutually anticommuting operators (where  $Z$  is some set of multiindex values),

$$\begin{aligned} \{\sigma_{\bar{i}}, \sigma_{\bar{j}}\}_{\bar{i}, \bar{j} \in Z} &\propto \delta_{\bar{i}, \bar{j}} \Rightarrow \\ \sum_{\bar{i} \in Z} T_{\bar{i}}^2 &\leq 1. \end{aligned} \quad (4)$$

Notice that operators  $\sigma_{\bar{i}}$  and  $\sigma_{\bar{j}}$  anticommute iff superindices differ on odd number of positions, excluding those, where one superindex has "0". In Ref. [14] this property was further generalized to cut-anticommutativity. Namely, consider two operators,  $\sigma_1 = \sigma_1^{[A]} \otimes \sigma_1^{[B]}$  and  $\sigma_2 = \sigma_2^{[A]} \otimes \sigma_2^{[B]}$ . We say that they anticommute with respect to cut  $A|B$  if they anticommute on either of the subsystem. Consequently,

$$\begin{aligned} \langle \sigma_1, \sigma_2 \rangle_{A|B} = 0 &\Rightarrow \\ \langle \sigma_1 \rangle^2 + \langle \sigma_2 \rangle^2 &\leq 1 \end{aligned} \quad (5)$$

for states, which are factorizable (and, by convexity, thus separable) with respect to the cut. This lead in Ref. [14] to constructing quadratic entanglement criteria based solely on anticommutativity properties of operators. The main goal of this contribution is to show that the formalism for qutrits can also be used for this purpose.

## III. CORRELATION TENSOR FORMALISM FOR MANY QUTRITS

We are now looking for a description of a qutrit, in which each measurement gives us a complete information about the probability distribution of three outcomes. To remove any dependencies, we expect the measurements on individual qutrits used for establishing the correlation tensor to be have mutually unbiased bases (MUBs) as their eigenbases. Lastly, since we want to formulate the complementarity relation similar to Eq. (5), so we expect the eigenvalues to be of modulo 1. A family satisfying these requirements for three-dimensional subsystems are the Heisenberg-Weyl matrices. They are given as

$$\begin{aligned} h_0 &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \\ h_1 &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & \omega & 0 \\ 0 & 0 & \omega^2 \end{pmatrix}, h_2 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}, \\ h_3 &= \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & \omega \\ \omega^2 & 0 & 0 \end{pmatrix}, h_4 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & \omega^2 \\ \omega & 0 & 0 \end{pmatrix}, \\ h_5 &= h_1^\dagger, h_6 = h_2^\dagger, \\ h_7 &= h_3^\dagger, h_8 = h_4^\dagger \end{aligned} \quad (6)$$

( $\omega = \exp(2\pi i/3)$ ).

First, let us show that this representation of a state is complete, that is, the data can be used for state tomography. As given in Ref. [18], a state can be given as

$$\rho = -1 + \sum_{m=1}^4 \sum_{k=0}^2 p(m, k) |mk\rangle \langle mk|, \quad (7)$$

where  $m$  enumerates the mutually unbiased basis, the eigenbasis of  $h_m$ ,  $|mk\rangle$  is the  $k$ th state of this basis and  $p(m, k) = \langle mk | \rho | mk \rangle$ . Now, consider the following quantity:

$$T_m = \text{tr} \rho h_m^\dagger. \quad (8)$$

For simplicity, let us represent complex numbers and operators as vectors, i.e.,  $\vec{a} = (\text{Re} a, \text{Im} a)$  and  $\vec{\sigma} = 1/2(o + o^\dagger, -i(o - o^\dagger))$ . Furthermore, let us denote  $1, \omega, \omega^2$  as  $\vec{v}_0 = (1, 0)$ ,  $\vec{v}_1 = (-\frac{1}{2}, \frac{\sqrt{3}}{2})$ ,  $\vec{v}_2 = (-\frac{1}{2}, -\frac{\sqrt{3}}{2})$ . Notice that this defines a new scalar product, which leads to

$$\begin{aligned} \vec{T}_m \cdot \vec{\sigma}_m &= \left( \sum_{k=0}^2 p(m, k) \vec{v}_k \right) \cdot \left( \sum_{l=0}^2 \vec{v}_l |ml\rangle \langle ml| \right) \\ &= -\frac{1}{2} + \frac{3}{2} p(m, k) |mk\rangle \langle mk|, \\ &\sum_{k=0}^2 p(m, k) |mk\rangle \langle mk| \end{aligned}$$

$$\begin{aligned} &= \frac{2\vec{T}_m \vec{o}_m + 1}{3}, \\ \rho &= \sum_m \frac{2}{3} \vec{T}_m \vec{o}_m - \frac{1}{3}. \end{aligned} \quad (9)$$

The last equation can be plugged in to Eq. (7). When the usual tensor product is used, this formula is extended by replacing products of probabilities with joint probabilities,  $p(k, m)p(l, n) \rightarrow p(k, m, l, n) = \langle k, m | \otimes \langle l, n | \rho | k, m \rangle \otimes | l, n \rangle$ .

Let us now consider the complementarity relations between tensor products of Heisenberg-Weyl operators. For certain noncommuting groups of operators,  $\{o_{\vec{j}}\}_{\vec{j}}$ , we shall have

$$\sum_{\vec{j}} |\langle o_{\vec{j}} \rangle|^2 \leq 1, \quad (10)$$

the equivalent of which was one of the key ingredients of Ref. [14] for qubits. Therein, this complementarity directly follows from the anticommutativity relations between the various Pauli matrix tensor products. Here, the situation is not as simple. The argument cannot go through directly as Heisenberg-Weyl tensor product operators do not anticommute. Still, we find some forms of complementarity between these operators. For an individual qutrit we shall have

$$\begin{aligned} 1 &\leq \text{tr} \rho^2 \\ &= \sum_{i,j=1}^3 |\rho_{ij}|^2 \\ &= \frac{1}{9} \sum_{i,j=0}^2 \left| \left\langle \sum_{k=0}^2 \omega^i h_2^j h_1^{jk} \right\rangle \right|^2 \\ &= \frac{1}{3} \sum_{i,j=0}^2 |\langle h_1^i h_2^j \rangle|^2, \\ 3 &\leq 1 + 2 \sum_{(i,j) \in \{(1,0), (0,1), (1,1), (1,2)\}} |\langle h_1^i h_2^j \rangle|^2, \end{aligned} \quad (11)$$

where the transition between the third and the fourth line comes from the Parseval's theorem for the Fourier transform.

Now, we are ready to consider the complementarity for many-qutrit operators. Here our possibilities are quite limited. One would expect that as long as tensor products do not commute, the sum of squared moduli of their averages for any state would not exceed 1. This is false, however. We have found 792 distinguished sets of seven mutually non-commuting two-qutrit operators,  $\{o_i\}_{i=1}^7$ , and found that for all of them there exist states, for which  $\sum_{i=1}^7 |\langle o_i \rangle|^2 = \frac{5}{4}$ . For the complete set of two-tensor products of Heisenberg-Weyl operators, from the semi-positivity of the state one can show that

$$\sum_{i,j=0}^8 |\langle h_i \otimes h_j \rangle|^2 \leq 9. \quad (12)$$

Nevertheless, we can easily argue for the complementarity of a smaller set. In particular consider a pair of operators,  $o_1$  and  $o_2$ , which do not commute with each other. By diagonalization one of them and a proper choice of phases of the new basis states we can bring to the  $3 \times 3$  block-diagonal form, where each of the blocks takes form

$$\begin{aligned} [o_1]_{\text{block}} &= h_1, \\ [o_2]_{\text{block}} &\propto h_2 \end{aligned} \quad (13)$$

and the complementarity follows directly from Eqs. (11). In addition, one may have two more operators, the blocks of which correspond to  $h_3$  ( $h_7$ ) and  $h_4$  ( $h_8$ ), up to global phases, extending the complementarity principle from two general to four specific operators. Notice that the operation diagonalizing  $o_1$  does not need to be local, so this complementarity is not of a strictly local nature.

We can now transplant the rest of ingredients from Ref. [14] to this consideration. Obviously, if we have mutually commuting operators, it suffices to choose a common eigenstate of all of them, to have all the mean values equal to 1. Also, we can use the proof from the reference that in quadratic entanglement criteria, mixing states cannot improve the situation.

Another fact we need for the construction is that for product states  $\rho = \rho_1^{[A]} \otimes \rho_2^{[B]}$  and a multiqutrit operator in form  $\vec{O} = o_1^{[A]} \otimes o_2^{[B]}$ , where  $[A]$  and  $[B]$  are subsystems, we have

$$|\langle \vec{O} \rangle|^2 = |\langle o_1^{[A]} \rangle|^2 |\langle o_2^{[B]} \rangle|^2, \quad (14)$$

which, again follows directly from the correspondence between the two-dimensional vector eigenvalues and the complex root-of-unity eigenvectors. However, this relation fails for  $d > 3$ , when we replace the complex roots of unity as eigenvalues with  $(d-1)$ -vectors  $\vec{v}_{d,i}$  satisfying relation

$$\vec{v}_{d,i} \cdot \vec{v}_{d,j} = \frac{d\delta_{i,j} - 1}{d-1}. \quad (15)$$

Thus our method is applicable only for a collection of qutrits.

#### IV. EXAMPLES

Consider the four-qutrit GHZ state, which in the computational basis ( $h_1|i\rangle = \omega^i|i\rangle$ ) has form

$$|\text{GHZ}_{3,4}\rangle = \frac{1}{\sqrt{3}} \sum_{j=0}^2 |jjj\rangle. \quad (16)$$

Perfect correlations of this state include (hereafter, we omit the tensor product signs)

$$\langle h_2 h_2 h_2 h_2 \rangle = \langle \Pi(h_1 h_5 h_1 h_5) \rangle = 1, \quad (17)$$

where  $\Pi(abcd)$  denotes an arbitrary permutation of  $a, b, c, d$  in terms of the tensor product. Hence we can

use the criterion

$$|\langle h_1 h_5 h_1 h_5 \rangle|^2 + |\langle h_2 h_2 h_2 h_2 \rangle|^2 \leq_{\text{SEP}} 1 \leq 2 \text{ (for } |GHZ_{3,4}\rangle) \quad (18)$$

to exclude separability with respect to bipartitions  $AB|CD$  and  $AD|BC$ , while the criterion

$$|\langle h_1 h_1 h_5 h_5 \rangle|^2 + |\langle h_1 h_5 h_1 h_5 \rangle|^2 \leq_{\text{SEP}} 1 \leq 2 \text{ (for } |GHZ_{3,4}\rangle) \quad (19)$$

can be used to exclude separability between subsystems  $AC$  and  $BD$ . Additionally, both of these criteria are sensitive to all one-versus-three cuts. Thus, a simultaneous violation of both of these inequalities certifies true multipartite entanglement of the tested state (in principle, different from the GHZ state).

The next example is the four-qutrit cluster state,

$$|C_{3,4}\rangle = \frac{1}{3} \sum_{i,j=0}^2 \omega^{ij} |ijij\rangle, \quad (20)$$

for which we can utilize correlations

$$\begin{aligned} \langle h_0 h_2 h_5 h_2 \rangle &= \langle h_2 h_0 h_2 h_5 \rangle \\ &= \langle h_5 h_2 h_0 h_2 \rangle = \langle h_2 h_5 h_2 h_0 \rangle = 1, \end{aligned} \quad (21)$$

Which gives us the following criterion for true four-qutrit entanglement:

$$\begin{aligned} &\frac{1}{2} (|\langle h_0 h_2 h_5 h_2 \rangle|^2 + |\langle h_2 h_0 h_2 h_5 \rangle|^2) + \\ &\frac{1}{2} (|\langle h_5 h_2 h_0 h_2 \rangle|^2 + |\langle h_2 h_5 h_2 h_0 \rangle|^2) > 1. \end{aligned} \quad (22)$$

Again, these four correlations can be, in principle established together, and while measuring in local MUBs, it again takes only two series of measurements to establish all four of them.

To demonstrate the usefulness and convenience of our method, let us consider four-qutrit graph states in general. Imagine a collection of four qutrits, each initialized in state  $\frac{1}{\sqrt{3}}(|0\rangle + |1\rangle + |2\rangle)$ . Now we take a graph, which connects four vertices. There are two such graphs with three edges (a path and a three-arm star), two with four (a loop and a triangle with a leg), one with five (a loop with one diagonal) and the complete graph has six edges. The graphs are presented in Fig. 1. If two qutrits are connected by an edge on the graph, we entangle them by applying a generalization of the control- $Z$  operation,

$$Ch_1 = \text{Diag}(1, 1, 1, 1, \omega, \omega^2, 1, \omega^2, \omega). \quad (23)$$

Each four-qutrit graph state has a total of 80 perfect correlations for nontrivial tensor products of the  $h$  operators. Lists of these correlations have been made, and from them we choose triples of operators, which satisfy the following conditions: (i) their mean value for the reference state has the absolute value equal to 1, (ii) for every bipartite cut, at least one pair does not cut-commute, (iii) they can be established in two measurement series.

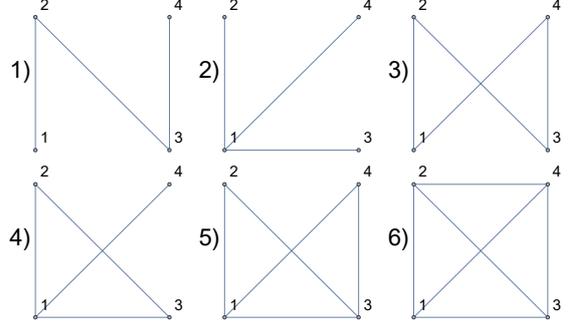


FIG. 1. Layouts of all four qutrit graph states. Qutrits are represented by vertices, while edges symbolize the application of the generalized control- $Z$  operation of Eq. (23).

We came to the conclusion that the true four-partite entanglement is certified if

$$\begin{aligned} &\text{for Graph 1:} \\ &|\langle h_3 h_8 h_4 h_7 \rangle|^2 + |\langle h_6 h_0 h_2 h_5 \rangle|^2 > 1 \\ &\wedge |\langle h_3 h_8 h_4 h_7 \rangle|^2 + |\langle h_0 h_5 h_2 h_5 \rangle|^2 > 1, \\ &\text{for Graph 2:} \\ &|\langle h_2 h_5 h_5 h_5 \rangle|^2 + |\langle h_1 h_6 h_6 h_4 \rangle|^2 > 1 \\ &\wedge |\langle h_2 h_5 h_5 h_5 \rangle|^2 + |\langle h_5 h_6 h_6 h_0 \rangle|^2 > 1, \\ &\text{for Graph 3:} \\ &|\langle h_3 h_3 h_3 h_3 \rangle|^2 + |\langle h_1 h_2 h_1 h_2 \rangle|^2 > 1 \\ &\wedge |\langle h_3 h_3 h_3 h_3 \rangle|^2 + |\langle h_1 h_0 h_1 h_6 \rangle|^2 > 1, \\ &\text{for Graph 4:} \\ &|\langle h_2 h_5 h_5 h_5 \rangle|^2 + |\langle h_4 h_3 h_3 h_7 \rangle|^2 > 1 \\ &\wedge |\langle h_2 h_5 h_5 h_5 \rangle|^2 + |\langle h_8 h_3 h_0 h_3 \rangle|^2 > 1, \\ &\text{for Graph 5:} \\ &|\langle h_4 h_2 h_6 h_2 \rangle|^2 + |\langle h_0 h_3 h_7 h_1 \rangle|^2 > 1 \\ &\wedge |\langle h_4 h_2 h_6 h_2 \rangle|^2 + |\langle h_3 h_7 h_0 h_5 \rangle|^2 > 1, \\ &\text{for Graph 6:} \\ &|\langle h_2 h_8 h_8 h_8 \rangle|^2 + |\langle h_0 h_3 h_3 h_3 \rangle|^2 > 1 \\ &\wedge |\langle h_2 h_8 h_8 h_8 \rangle|^2 + |\langle h_3 h_3 h_3 h_0 \rangle|^2 > 1. \end{aligned} \quad (24)$$

Notice that not all of these correlations are equal to 1, but since the criteria are quadratic, this is no concern.

## V. CONCLUSIONS

We have shown a generalization of a derivation of graph-based quadratic entanglement criteria, known already for qubits, to qutrits. It was achieved by using Heisenberg-Weyl operators, which can be seen vector-valued observables. While the obtained criteria can be applied to a relatively small set of states, namely those with very strong correlations, they are easy to derive,

compared to most other methods. One does not need to optimize over the whole set of product states, but simply find some pairs of correlations, that we expect to be simultaneously high. This was well demonstrated in case of four-qutrit graph states.

There are some differences between the derivation presented in Ref. [14] and the above. Therein, we enjoyed the complementarity relation for an arbitrarily large set of cut-anticommuting operators. For qutrits, we have found counterexamples. The complementarity principle holds in general for pairs of (cut-)noncommuting observables, and for more only in special cases. One still can, however, construct criteria such as those in Ref. [15], involving only two terms each. For a given term, we take as many pairs as necessary to exclude separability of the state along all cuts.

Interestingly, we were not able to push the reasoning even further, to dimensions of subsystems higher than 3. There are few obstacles in generalizing the proofs. One

is that the proof of the complementarity relation (Eq. (10)) explicitly refers to the Heisenberg-Weyl formalism, which consists of the shift operators. The other difficulty is that for vector eigenvalues, we were unable to derive a dependence between the length of the mean value of a joint observable and the lengths for local operators.

## VI. ACKNOWLEDGEMENTS

The work is subsidized from funds for science for years 2012-2015 approved for international co-financed project BRISQ2 by Polish Ministry of Science and Higher Education (MNiSW). This work is a part of BRISQ2 Project co-financed by EU. MW is supported by NCN Grant No. 2012/05/E/ST2/02352. KR is supported by NCN Grant No. 2015/19/B/ST2/01999. L.K. acknowledges support by the international PhD programme ExQM (Eli te Network of Bavaria).

- 
- [1] A. Einstein, B. Podolsky, N. Rosen, “Can Quantum-Mechanical Description of Physical Reality be Considered Complete?”, *Physical Review* **47**, 777 (1935).
  - [2] J. S. Bell, “On the Einstein-Podolsky-Rosen paradox”, *Physics* **1**, 195 (1964).
  - [3] S. Kochen and E.P. Specker, “The problem of hidden variables in quantum mechanics”, *Journal of Mathematics and Mechanics*. **17**, 59 (1967).
  - [4] Č. Brukner, M. Żukowski, J.-W. Pan, and A. Zeilinger, “Bells Inequalities and Quantum Communication Complexity”, *Phys. Rev. Lett.* **92**, 127901 (2004).
  - [5] A. K. Ekert, “Quantum cryptography based on Bells theorem”, *Phys. Rev. Lett.* **67**, 661 (1991).
  - [6] A. Peres, “Separability Criterion for Density Matrices”, *Phys. Rev. Lett.* **77**, 1413 (1996).
  - [7] M. Horodecki, P. Horodecki, and R. Horodecki, “Separability of Mixed States: Necessary and Sufficient Conditions”, *Phys. Lett. A* **223**, 1 (1996).
  - [8] M.-D. Choi, “Completely positive maps on complex matrices”, *Linear Alg. and Its Appl.* **10**, 285 (1975); A. Jamiolkowski, “Linear transformations which preserve trace and semidefiniteness of operators”, *Rep. Math. Phys.* **3**, N4, 275 (1972).
  - [9] R. F. Werner and M. M. Wolf, “All-multipartite Bell-correlation inequalities for two dichotomic observables per site”, *Phys. Rev. A* **64**, 032112 (2001).
  - [10] H. Weinfurter and M. Żukowski, “Four-photon entanglement from down-conversion”, *Phys. Rev. A* **64** 010102(R) (2001).
  - [11] M. Żukowski and Č. Brukner, “Bells Theorem for General N-Qubit States”, *Phys. Rev. Lett.* **88**, 210401 (2002).
  - [12] K. Nagata, W. Laskowski, M. Wieśniak, and M. Żukowski, “Rotational Invariance as an Additional Constraint on Local Realism”, *Phys. Rev. Lett.* **93**, 230403 (2004).
  - [13] P. Badziąg, Č. Brukner, W. Laskowski, T. Paterek, and M. Żukowski, “Experimentally Friendly Geometrical Criteria for Entanglement”, *Phys. Rev. Lett.* **100**, 140403 (2008).
  - [14] M. Wieśniak and K. Maruyama, “Package of facts and theorems for efficiently generating entanglement criteria for many qubits”, *Phys. Rev. A* **85**, 062315 (2012).
  - [15] L. Knips, C. Schwemmer, N. Klein, M. Wieśniak, and H. Weinfurter, “Multipartite entanglement detection with minimal effort”, arXiv:1412.5881.
  - [16] R. Weinar, W. Laskowski, M. Pawłowski, “Activation of entanglement in teleportation”, *J. Phys. A* **46**, 435301 (2013).
  - [17] G. Jaeger, A. Shimony, and L. Vaidman, “Two interferometric complementarities”, *Phys. Rev. A* **51**, 54 (1995).
  - [18] W. K. Wootters and B. D. Fields “Optimal state-determination by mutually unbiased measurements”, *Ann. Phys. (N.Y.)* **191**, 363 (1989).
  - [19] I. D. Ivanović, “Geometrical description of quantal state determination”, *J. Phys. A* **14**, 3241 (1981).
  - [20] J.-L. Chen, Ch.-F. Wu, L. C. Kwek, D. Kaszlikowski, M. Żukowski, and C. H. Oh, “Multicomponent Bell inequality and its violation for continuous-variable systems”, *Phys. Rev. A* **71**, 032107 (2005).

### 4.3 Entanglement and Random Measurements

While the previously introduced schemes provide versatile tools to detect entanglement, there are some circumstances where they cannot be applied at all. Those criteria as well as the more basic approaches described in Sec. 2.2.2 require fixed and known local reference frames. If instead, in a bipartite scenario, Alice and Bob are allowed to perform several measurements, but do not have control of or either knowledge about the current measurement setting, usual criteria fail.

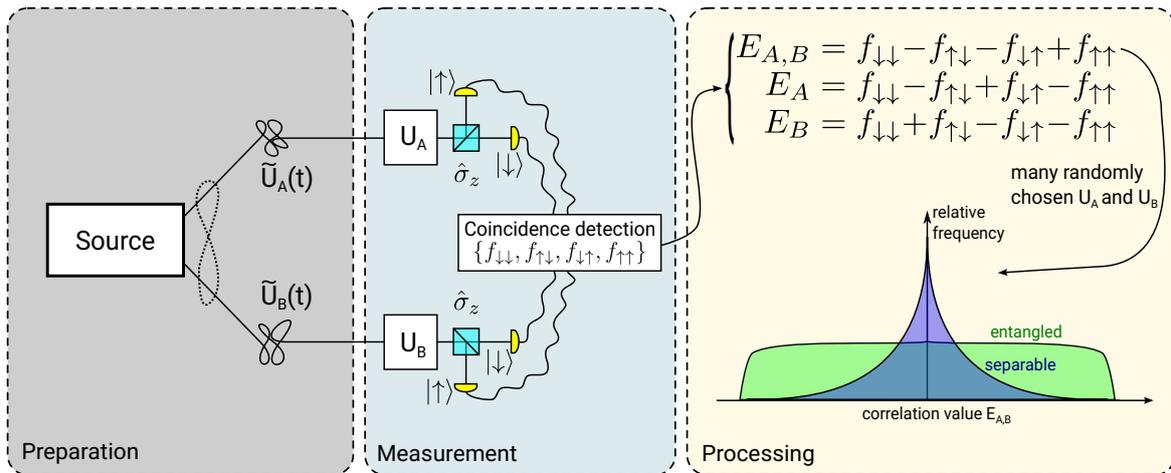


FIGURE 4.3: Scheme for random measurements, shown here for two-qubit states. An unknown quantum state is produced by a source and affected by (in general fluctuating) local unitary transformations. By performing local measurements in a large set of randomly chosen bases, one can eventually obtain information about the quantum state. Most obvious, the distribution of correlation values for a highly entangled two-qubit state (schematically shown in green) strongly deviates from the distribution of a product state (blue).

This scenario is schematically shown in Fig. 4.3. A source produces a bipartite quantum state, e.g., pairs of photons, which might exhibit entanglement between the two photons. Due to noisy quantum channels - or even a noisy source in the first place -, local unitary transformations  $\tilde{U}_A(t)$  and  $\tilde{U}_B(t)$  can (locally) change the reference frame. Afterwards, the two photons are distributed to the observers Alice and Bob, allowing each of them to apply a (known) local unitary rotation, perform measurements and record coincidence clicks between their outcomes. It is important to note that the timescale of the noise in the channels, i.e., of  $\tilde{U}_A(t)$  and  $\tilde{U}_B(t)$ , is slow enough that the expectation value in the (unknown) basis can be measured, but possibly too fast for conducting measurements in several bases. Thus, it is impossible to find the current local reference frames and perform measurements therein.

In general, the uncontrollable and unknown unitaries  $\tilde{U}_A(t)$  and  $\tilde{U}_B(t)$  of the noisy channels will be biased, i.e., they might lead to more frequent measurements along some specific directions than others. However, if *additionally* Alice and Bob apply well defined

unitary transformations  $U_A$  and  $U_B$ , respectively, which are drawn from a set of random unitary matrices distributed uniformly according to the Haar measure, also the products of the local unitaries,  $U_A \cdot \tilde{U}_A(t)$  as well as  $U_B \cdot \tilde{U}_B(t)$ , are Haar randomly distributed. When expectation values measured in many Haar randomly distributed measurement bases, one will observe that states which cannot be transformed into each other using local unitary transformations lead to different distributions of expectation values. As sketched in Fig. 4.3, the distribution of expectation values will be rather flat for a bipartite entangled state, while a pure product state shows a concentration to small expectation values [192], as the product state shows a large correlation only along a very particular measurement direction, whereas most measurement results are uncorrelated. In short, it is more likely to find a large correlation in an entangled state than in a product state.

In this section, the technique of genuine multipartite entanglement detection based on distributions of expectation values will be discussed. First, a general criterion based on the knowledge of the *full* distribution together with all *marginal* distributions will be derived. Afterwards, this criterion will be applied to experimental data.

Although some criteria for detecting genuine multipartite entanglement in a similar scenario exist [193, 194], those methods require some additional knowledge. For example, the method proposed in Ref. [194] requires measurements in two bases with a known relation to each other, which does not apply in the aforementioned scenario, where each measurement basis is unknown and chosen independently of previous measurements. Other approaches such as [190, 195] are limited to, e.g., the detection of entanglement and do not reveal genuine multipartite entanglement.

### 4.3.1 Distribution of Expectation Values

In a scenario without fixed or with even unknown reference frames, still meaningful measurements can be conducted. Although one does not obtain meaningful information by a single measurement, the distribution of measurement outcomes contains a surprising large extent of information. By considering the distribution of correlation measurement outcomes between all different combinations of observers, the structure of entanglement may be revealed in detail.

In the two qubit scenario, the measurement scheme can be described as following. An unknown, but fixed state described by a density matrix  $\varrho$  is prepared. Two observers, Alice and Bob, perform local measurements of the locally randomly rotated state in some fixed

bases, e.g.,  $\sigma_z$ , to obtain

$$E_{A,B}^{(i)} = \text{tr} \left[ (\sigma_z \otimes \sigma_z) \left( U_A^{(i)} \otimes U_B^{(i)} \right) \varrho \left( U_A^{(i)} \otimes U_B^{(i)} \right)^\dagger \right], \quad (4.20a)$$

$$E_A^{(i)} = \text{tr} \left[ (\sigma_z \otimes \sigma_0) \left( U_A^{(i)} \otimes U_B^{(i)} \right) \varrho \left( U_A^{(i)} \otimes U_B^{(i)} \right)^\dagger \right], \quad (4.20b)$$

$$E_B^{(i)} = \text{tr} \left[ (\sigma_0 \otimes \sigma_z) \left( U_A^{(i)} \otimes U_B^{(i)} \right) \varrho \left( U_A^{(i)} \otimes U_B^{(i)} \right)^\dagger \right], \quad (4.20c)$$

$$E_0^{(i)} = \text{tr} \left[ (\sigma_0 \otimes \sigma_0) \left( U_A^{(i)} \otimes U_B^{(i)} \right) \varrho \left( U_A^{(i)} \otimes U_B^{(i)} \right)^\dagger \right], \quad (4.20d)$$

where, due to normalization,  $E_0^{(i)} \equiv 1$  holds. For each randomly rotated state, indexed by  $(i)$ , Alice and Bob can record this set of expectation values. After having measured in several locally randomly rotated bases, where the local unitary matrices are uniformly distributed according to the Haar measure [196], the observers can perform a statistical analysis of the obtained expectation values. The probability distribution  $g(E_{A,B})$  for observation of a specific correlation value  $E_{A,B}$  is then described by

$$g(E_{A,B}) = \frac{\int dU_A dU_B \delta \left[ E_{A,B} - \text{tr} \left( (\sigma_z \otimes \sigma_z) \left( U_A \otimes U_B \right) \varrho \left( U_A \otimes U_B \right)^\dagger \right) \right]}{\int dU_A dU_B}, \quad (4.21)$$

where  $\delta(x)$  denotes Dirac's delta distribution with  $\int dx g(x) \delta(x - x_0) = g(x_0)$ . The probability distributions of the marginal correlations  $E_A$  and  $E_B$  are obtained accordingly.

For example, the maximally entangled two qubit Bell states, e.g.,  $\varrho_{\text{Bell}} = |\phi^{(+)}\rangle\langle\phi^{(+)}|$ , will result in a flat distribution for  $E_{A,B}$ , i.e., the probability for observation is identical for all possible values,

$$g_{\text{Bell}}(E_{A,B}) = \begin{cases} \frac{1}{2} & E_{A,B} \in [-1, 1], \\ 0 & \text{otherwise.} \end{cases} \quad (4.22)$$

Intuitively, Alice can freely choose any measurement direction. If Bob chooses to measure parallel to Alice's direction, they will observe a strong correlation or anticorrelation, i.e., a result large in modulus. Hence, the measured correlation depends solely on the projection of Bob's measurement direction onto Alice's direction. For uniform sampling of vectors on the surface of a unit sphere, the distribution of the projection of those samples onto a fixed axis is uniform, leading to Eq. (4.22). In contrast, a two qubit product state, e.g.,  $\varrho_{\text{product}} = |00\rangle\langle 00|$ , results in a probability distribution for  $E_{A,B}$  described by

$$g_{\text{product}}(E_{A,B}) = -\frac{1}{2} \log(|E_{A,B}|). \quad (4.23)$$

Compared to above's scenario, not only the projection of Bob's measurement direction onto Alice's one is relevant, but a large correlation is only observed if also Alice measured

along the direction of state preparation. The maximally mixed two qubit state  $\rho_1 = \mathbb{1}/2$  leads to

$$g_1(E_{A,B}) = \delta(E_{A,B}), \quad (4.24)$$

as the outcomes of Alice and Bob are uncorrelated in all directions. Some example distributions are shown in Fig. 4.4.

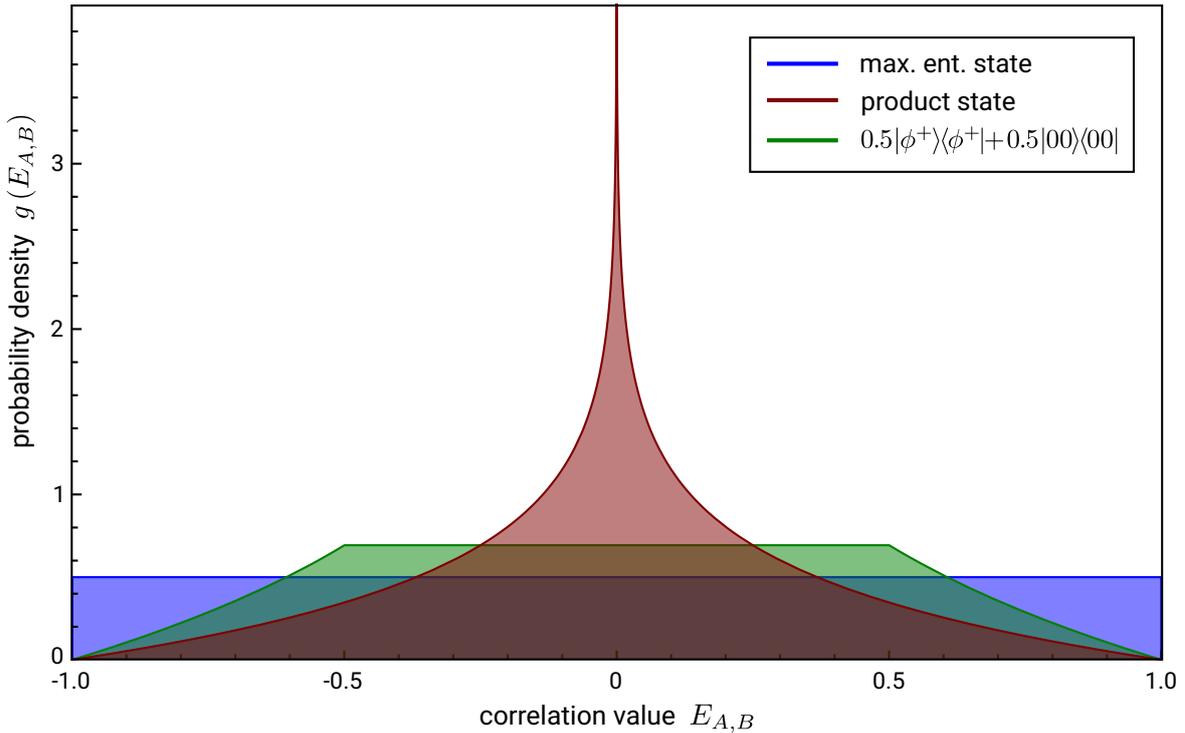


FIGURE 4.4: The probability density for correlation values in the random measurement scenario. For measuring a Bell state (blue) in random bases, all correlation values are equally likely, see Eq. (4.22). For a product state (red), e.g.,  $|00\rangle$ , it is much more likely to measure a correlation with smaller modulus than with a larger modulus. The probability density follows a logarithmic distribution, see Eq. (4.23). Equally mixing those two states (green) results in a distribution which is flat within some bounds and subsequently decaying logarithmically.

To obtain these distributions in the general  $n$ -qubit setting,  $2^n$  different expectation values have to be recorded instead of only the 4 (one of which is trivial) given in Eqs. (4.20).

#### 4.3.1.1 Obtaining the Purity

The probability distribution of Eq. (4.21) and the probability distributions of the marginal correlations allow to obtain quantities invariant under local unitary transformations. One of the most essential quantities is the purity, which will also be needed to certify genuine

multipartite entanglement. The purity of the  $n$ -qubit state  $\varrho$  is given by

$$\mathcal{P}(\varrho) = \text{tr}(\varrho^2) = \frac{1}{2^n} \sum_{\mu_1, \dots, \mu_n} T_{\mu_1, \dots, \mu_n}^2, \quad (4.25)$$

with  $\mu_i \in \{0, 1, 2, 3\}$ ,  $i \in \{1, \dots, n\}$ . When separating the summation index 0, which indicates that the result of the respective observer is not taken into account, from the rest, the summation becomes

$$\mathcal{P}(\varrho) = \frac{1}{2^n} \left[ T_{0, \dots, 0}^2 + \sum_{j_1} T_{j_1, 0, \dots, 0}^2 + \dots + \sum_{j_2, \dots, j_n} T_{0, j_2, \dots, j_n}^2 + \dots + \sum_{j_1, j_2, \dots, j_n} T_{j_1, j_2, \dots, j_n}^2 \right] \quad (4.26)$$

with  $j_i \in \{1, 2, 3\}$ ,  $i \in \{1, \dots, n\}$ .  $T_{0, \dots, 0}^2 \equiv 1$  is given by normalization. The second contribution  $\sum_{j_1, 0, \dots, 0} T_{j_1, 0, \dots, 0}^2$  can be obtained from the probability distribution  $g(E_{O_1, 0, \dots, 0})$  of correlation values  $E_{O_1}$  involving only the first party,

$$\underbrace{\frac{1}{3} \sum_{j_1} T_{j_1, 0, \dots, 0}^2}_{\mathbb{E}[T_{j_1, 0, \dots, 0}^2]} = \underbrace{\int_{-1}^1 dE_{O_1} g(E_{O_1}) E_{O_1}^2}_{\mathbb{E}[E_{O_1}^2]}. \quad (4.27)$$

$$\begin{aligned} \mathcal{P}(\varrho) = & \frac{1}{2^n} \left[ 1 + 3 \left( \mathbb{E}[E_{O_1, 0, \dots, 0}^2] + \dots + \mathbb{E}[E_{0, 0, \dots, O_n}^2] \right) \right. \\ & + 3^2 \left( \mathbb{E}[E_{O_1, O_2, 0, \dots, 0}^2] + \dots + \mathbb{E}[E_{0, \dots, 0, O_{n-1}, O_n}^2] \right) \\ & \left. + \dots + 3^n \mathbb{E}[E_{O_1, \dots, O_n}^2] \right], \end{aligned} \quad (4.28)$$

which is the mean value of the sum of all  $k$ -point correlation functions weighted with  $3^k$ . This expression solely depends on the measured distributions of correlations and the number of qubits.

### 4.3.2 Certifying Genuine Multipartite Entanglement

Following the idea of [195], entanglement can be directly detected. If the second moment of the  $n$ -partite correlation function with

$$m_{\mathcal{S}} \equiv \int_{-1}^1 dE_{\mathcal{S}} g(E_{\mathcal{S}}) E_{\mathcal{S}}^2, \quad (4.29)$$

where  $\mathcal{S} \equiv \{1, 2, \dots, n\}$  denotes the set of observers, exceeds the threshold of a product state,  $m_{\mathcal{S}}^{\text{prod}} \leq \frac{1}{3^n}$ , entanglement is detected, i.e.,

$$m_{\mathcal{S}} > \frac{1}{3^n} \Rightarrow \varrho \text{ is entangled.} \quad (4.30)$$

All pure two qubit states  $\cos \frac{\theta}{2}|00\rangle + e^{i\varphi} \sin \frac{\theta}{2}|11\rangle$  (for properly chosen basis states  $|00\rangle$  and  $|11\rangle$ ) with  $\theta \in (0, \pi)$  are entangled. For those states, one obtains

$$m_{A,B} = \frac{1}{3^2} [1 + 2 \sin(\theta)], \quad (4.31)$$

which indeed reveals entanglement for  $\theta \in (0, \pi)$ . In other words, the state cannot be separable along the bipartition between the first and the second qubit as soon as the joint measurement results of both observers cannot be explained by the product of the respective local measurement results. Hence, the state is entangled.

More generally, entanglement across a specific partition  $\mathcal{A}|\mathcal{B}$  where  $\mathcal{A}$  and  $\mathcal{B}$  denote (possibly) multiple indices given by disjoint proper subsets of  $\mathcal{A} \cup \mathcal{B} = \mathcal{S} \equiv \{1, 2, \dots, n\}$  can be checked equally well. A state which is product across the partition  $\mathcal{A}|\mathcal{B}$ ,  $\varrho_{\mathcal{A}\mathcal{B}}^{\mathcal{A}|\mathcal{B}} = \varrho_{\mathcal{A}}^{\mathcal{A}|\mathcal{B}} \otimes \varrho_{\mathcal{B}}^{\mathcal{A}|\mathcal{B}}$ , fulfills

$$m_{\mathcal{A}\mathcal{B}}^{\mathcal{A}|\mathcal{B}} = m_{\mathcal{A}}^{\mathcal{A}|\mathcal{B}} m_{\mathcal{B}}^{\mathcal{A}|\mathcal{B}}. \quad (4.32)$$

A violation of Eq. (4.32) hence indicates entanglement spanned across the  $\mathcal{A}|\mathcal{B}$ -partition. However, if the state contains  $n$  qubits with  $n > 2$ , this separability test cannot directly be used to certify genuine  $n$ -partite entanglement as will be shown by a counterexample.

Detection of genuine multipartite entanglement for  $n > 2$  parties, where the notion of *genuine* multipartite entanglement becomes meaningful, is more difficult. Consider the state

$$\varrho_{\text{bisep}} = \frac{1}{2} (|0\rangle\langle 0| \otimes |\psi^-\rangle\langle\psi^-| + |\psi^-\rangle\langle\psi^-| \otimes |0\rangle\langle 0|), \quad (4.33)$$

which is the mixture of a state carrying entanglement between  $A$  and  $B$ , separable from  $C$  with a state entangled between  $B$  and  $C$ , separable from  $A$ . By construction, this state is biseparable as it is the incoherent mixture of two biseparable states, see Sec. 2.2.2.1. Yet, for all possible biseparations,  $A|BC$ ,  $B|AC$ , and  $C|AB$ , the measurement results will differ from the results of a joint measurement involving all three observers since at least one of the mixed pure states is entangled along the respective partition. Hence, the state  $\varrho_{\text{bisep}}$  is a counterexample to the naive approach to certify genuine multipartite entanglement by ruling out separability along every bipartition.

Nevertheless, we have found a means to detect genuine multipartite entanglement in the scenario of noisy channels. In order to detect genuine multipartite entanglement (GME), consider the function

$$f(\varrho) = m_{\mathcal{S}} - \frac{1}{2} \sum_{\substack{\{i_1, \dots, i_n\}=0 \\ 0 < i_1 + \dots + i_n < n}}^1 m_{i_1, i_2, \dots, i_n} m_{1-i_1, 1-i_2, \dots, 1-i_n}, \quad (4.34)$$

which can be understood as a measure *by how much* the state cannot be written as a bi-product. The labels  $i_1, i_2, \dots, i_n$  in the notation  $m_{i_1, i_2, \dots, i_n}$  are used to indicate if the outcome of the corresponding qubit is considered ( $i_j = 1$ ) in the measurement or not ( $i_j = 0$ ). Hence,  $m_{1,1,\dots,1}$  denotes the second moment of a measurement involving all observers

and is therefore identical to  $m_S$ , whereas, e.g.,  $m_{1,1,0,\dots,0}$  denotes the second moment of a measurement involving only the first two observers. As an example, the product state  $|\psi_1\rangle = |0\rangle \otimes |\psi^-\rangle$  factorizes, leading to  $m_{1,2,3} = m_1 m_{2,3}$  and hence to  $f(|\psi_1\rangle\langle\psi_1|) = 0$ . The state  $|\psi_2\rangle = |0\rangle \otimes |0\rangle \otimes |0\rangle$  factorizes even long multiple bipartitions,  $m_{1,2,3} = m_1 m_{2,3} = m_2 m_{1,3} = m_3 m_{1,2}$ , therefore  $f(|\psi_2\rangle\langle\psi_2|)$  even becomes negative.

If the experimentally observed state  $\varrho_{\text{exp}}$  results in a larger value of this function than all biseparable states with the same purity, i.e.,

$$f(\varrho_{\text{exp}}) > f_{\max}^{\mathcal{P}(\varrho_{\text{bs}})} \equiv \max_{\substack{\varrho_{\text{bs}} \\ \text{with} \\ \mathcal{P}(\varrho_{\text{bs}}) = \mathcal{P}(\varrho_{\text{exp}})}} f(\varrho_{\text{bs}}) \Rightarrow \varrho \text{ is gen. mult. entangled.} \quad (4.35)$$

Though, obtaining the maximum over all biseparable states is difficult. However, by simulating a large set of random biseparable states for  $n = 2, 3, 4$  and examine the purity dependent values for  $f(\varrho_{\text{bs}})$ , the maximization can be carried out numerically. For the cases of  $n = 3$  and  $n = 4$  qubits, the results of the simulation of biseparable states are shown in Fig. 4.5 and Fig. 4.6, respectively. All biseparable states are found below some parabola, which can be used as the critical threshold.

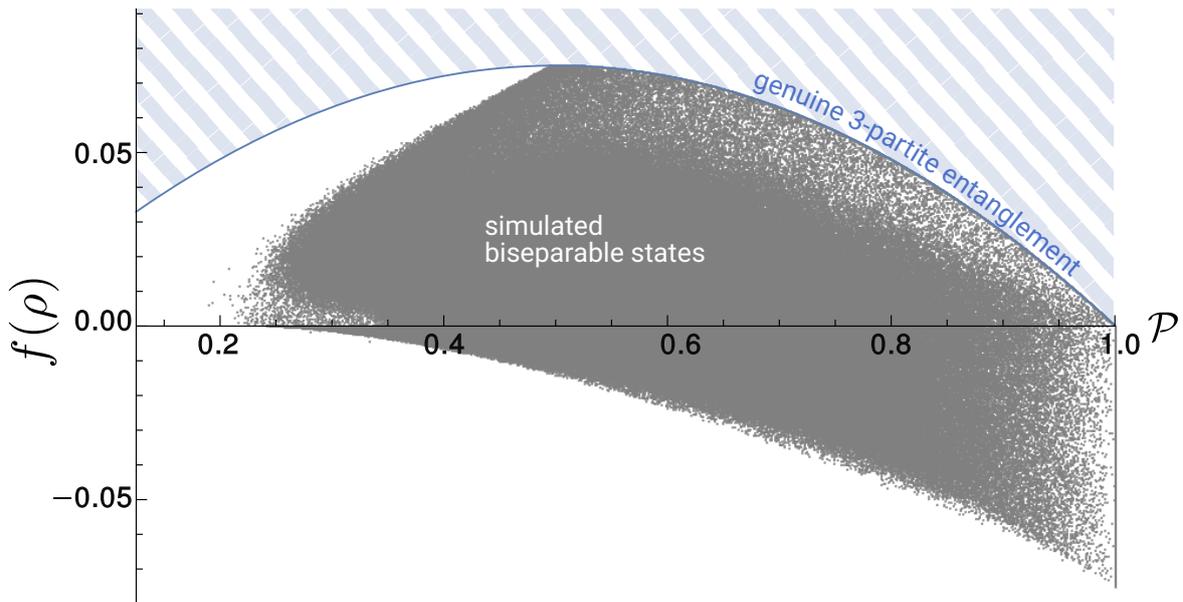


FIGURE 4.5: The function  $f(\varrho)$  for about  $3.3 \times 10^6$  randomly obtained biseparable tripartite states is shown. Those states are found below the parabola  $f(\varrho) = \frac{8}{27}(1 - \mathcal{P})\mathcal{P}$ , where  $\mathcal{P}$  denotes the purity. The parabola gives a tight bound for  $\mathcal{P} \geq \frac{1}{2}$ . As the simulation was focused on states with rather high purity, it is an open question if the parabola is tight for states with lower purity or if a tighter bound can be found. The straight vertical line at  $\mathcal{P} = 1$  is due to (pure) product states, which are additionally simulated.

For four qubits, the function  $f(\varrho)$  in dependence of the purity  $\mathcal{P}(\varrho)$  is shown in Fig. 4.6. Due to the increasing complexity of four-qubit simulations, the method to obtain bisepa-

rable states is tuned such that the borders of the area can be assessed, leading to areas of lower density.

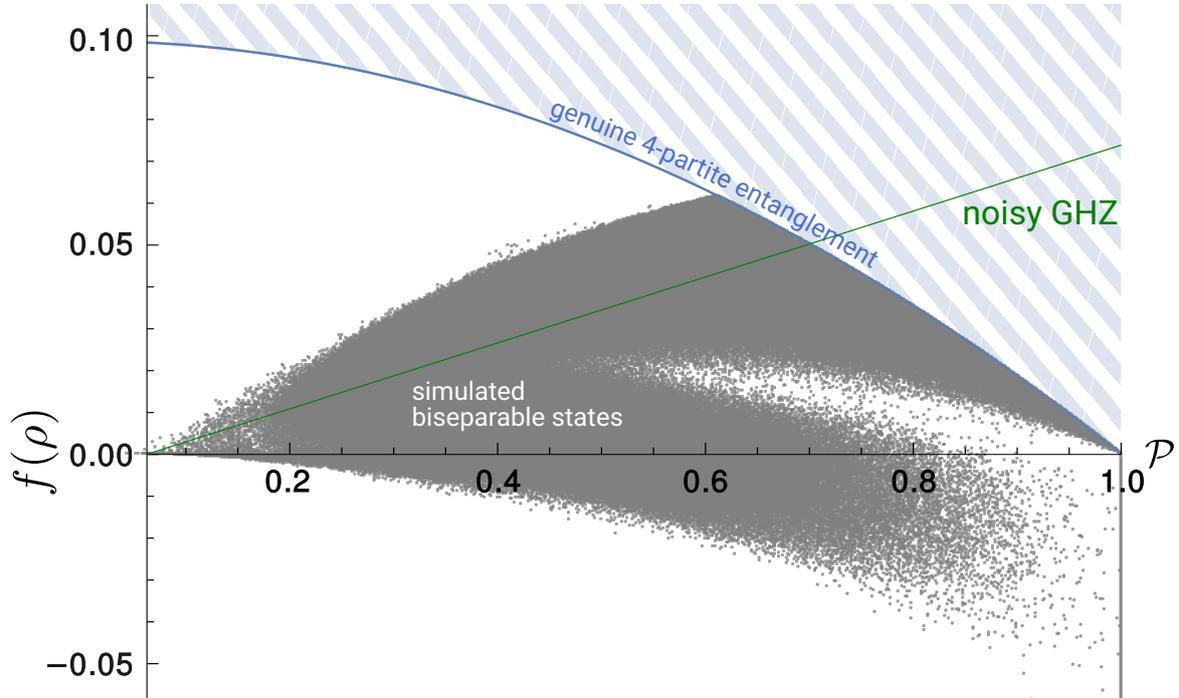


FIGURE 4.6: All about  $5.5 \times 10^6$  randomly obtained biseparable four-partite states are found below the parabola  $f(\varrho) = \frac{8}{81}(1 - \mathcal{P}^2)$ . This bound allows to prove genuine four-partite entanglement for a four-partite Greenberger-Horne-Zeilinger state mixed with white noise, see Eq. (4.38), for purities above approximate 0.7. This is indicated by the green line, which is derived in Eq. (4.41).

Together with an analogous simulation for  $n = 2$ , one eventually finds bounds for biseparable states,

$$f(\varrho_{\text{bs}}) < f_{\text{max},n}^{\mathcal{P}(\varrho_{\text{bs}})} \leq \begin{cases} \frac{4}{9}(1 - \mathcal{P})\mathcal{P} & \text{for } n = 2, \\ \frac{8}{27}(1 - \mathcal{P})\mathcal{P} & \text{for } n = 3, \\ \frac{8}{81}(1 - \mathcal{P}^2) & \text{for } n = 4 \end{cases} \quad (4.36)$$

with  $\mathcal{P} \equiv \mathcal{P}(\varrho_{\text{bs}})$ . Note that both  $f(\varrho_{\text{exp}})$  and  $\mathcal{P}$  only depend on the (second) moments of the correlation values, allowing to test for

$$f(\varrho_{\text{exp}}) - f_{\text{max}}^{\mathcal{P}(\varrho_{\text{exp}})} > 0 \quad (4.37)$$

solely by using the distributions of correlation values. Finding bounds - and ideally proving their optimality - for  $n > 4$  remains unsolved.

Finding a state outside the given bound, i.e., above the parabola, indicates genuine  $n$ -partite entanglement. The tolerance against white noise can easily be assessed. For the

four-qubit Greenberger-Horne-Zeilinger state, assume a noise model of

$$\varrho_{\text{GHZ},p} = p|\text{GHZ}_4\rangle\langle\text{GHZ}_4| + (1-p)\varrho_{\text{mms}}, \quad (4.38)$$

where  $\varrho_{\text{mms}}$  denotes the maximally mixed state. Then, the purity is given by

$$\mathcal{P}(p) = \frac{1}{16} + \frac{15p^2}{16}, \quad (4.39)$$

from which one can obtain  $p = \sqrt{(16\mathcal{P} - 1)/15}$ . With the function detecting entanglement

$$f(\varrho) = p^2 \left( \frac{1}{9} - 3 \cdot \frac{1}{9} \cdot \frac{1}{9} \right) = \frac{2}{27}p^2, \quad (4.40)$$

one finally obtains

$$f(\varrho) = \frac{2}{27} \frac{16\mathcal{P} - 1}{15}. \quad (4.41)$$

This function, which is shown as a straight green line in Fig. 4.6, intersects the parabola of Eq. (4.36) at around  $\mathcal{P}_{\text{crit}} \approx 0.7$ . Hence, in order to still reveal genuine four-partite entanglement for a noisy GHZ state, white noise can be tolerated as long as  $p > p_{\text{crit}} \approx 0.82$ . With this robustness against white noise, the criterion is applicable for practical purposes. In the next section, this method will be applied to a set of experimentally prepared states including the Greenberger-Horne-Zeilinger state, the cluster state as well as some separable states.

In the following preprint of the manuscript ‘‘Multipartite Entanglement Analysis From Random Correlations’’ [P5], the scenario of performing measurements with noisy channels is introduced and discussed in detail. Four-qubit GHZ and cluster states are prepared as well as a tri-separable state and a bi-separable state with partial entanglement and analyzed in this scenario. Using the numerical simulations which had been introduced above, entanglement identifier are used to deduce the entanglement structures of those states. It is remarkable how much information can still be obtained even when lacking any knowledge about the conducted measurements.

This manuscript is based on initial ideas both of Jasmin Meinecke and former co-workers and of Wiesław Laskowski and Tomasz Paterek who already published related work for detecting entanglement, albeit not genuine multipartite entanglement, with a single random measurement. After discussing with Jasmin Meinecke, Jan Dziewior and Harald Weinfurter, I developed means to experimentally realize the random measurements, which were later refined by Jasmin Meinecke, Jan Dziewior and myself. Jan Dziewior, Jasmin Meinecke and me conducted the experimental measurements based on the experimental setup constructed for aforementioned publication [P3]. I performed the numerical simulations for biseparable states and conceived, based on discussions with all co-authors, a means to detect genuine multipartite entanglement based on random measurements. My simulations have been verified by Waldemar Kłobus. Together with Jan Dziewior, I derived some correction for finite sample sizes. The data evaluation has been shared between Jasmin Meinecke and myself. The manuscript, which is still in preparation, was written and edited by all authors.

## Multipartite Entanglement Analysis From Random Correlations

Lukas Knips,<sup>1,2,3</sup> Jan Dziewior,<sup>1,2,3</sup> Waldemar Kłobus,<sup>4</sup> Wiesław Laskowski,<sup>4,5</sup> Tomasz Paterek,<sup>6,7</sup> Peter J. Shadbolt,<sup>8</sup> Harald Weinfurter,<sup>1,2,3</sup> and Jasmin D. A. Meinecke<sup>1,2,3</sup>

<sup>1</sup>*Max-Planck-Institut für Quantenoptik, Hans-Kopfermann-Straße 1, 85748 Garching, Germany*

<sup>2</sup>*Department für Physik, Ludwig-Maximilians-Universität, Schellingstraße 4, 80799 München, Germany*

<sup>3</sup>*Munich Center for Quantum Science and Technology (MCQST), Schellingstraße 4, 80799 München, Germany*

<sup>4</sup>*Institute of Theoretical Physics and Astrophysics, Faculty of Mathematics,*

*Physics and Informatics, University of Gdańsk, 80-308 Gdańsk, Poland*

<sup>5</sup>*International Centre for Theory of Quantum Technologies, University of Gdańsk, 80-308 Gdańsk, Poland*

<sup>6</sup>*School of Physical and Mathematical Sciences, Nanyang Technological University, 637371 Singapore*

<sup>7</sup>*MajuLab, International Joint Research Unit UMI 3654,*

*CNRS, Université Côte d'Azur, Sorbonne Université,*

*National University of Singapore, Nanyang Technological University, Singapore.*

<sup>8</sup>*Department of Physics, Imperial College London, London SW7 2AZ, United Kingdom*

Quantum entanglement is usually revealed via a well aligned, carefully chosen set of measurements. Yet, under a number of experimental conditions, for example in communication within multiparty quantum networks, noise along the channels or fluctuating orientations of reference frames may ruin the quality of the distributed states. Here we show that even for strong fluctuations one can still gain detailed information about the state and its entanglement using random measurements. Correlations between all or subsets of the measurement outcomes and especially their distributions provide information about the entanglement structure of a state and also enable to witness genuine multipartite entanglement. As long as the rate of entanglement generation is sufficiently high, this method overcomes any type and strength of localized unitary noise.

### I. INTRODUCTION

One of the most striking features of quantum entanglement is the existence of correlated measurement outcomes between spatially separated particles, which exceed expectations based on classical physics. These correlations are typically observed with carefully aligned local measurements. They get distorted if a common reference frame is lacking and especially in the presence of noise along the channels distributing the entangled particles. In practice, for many channels the instabilities are often irremovable: optical fibers rotate polarization, changing phases affect a path degree of freedom, atmospheric turbulence acts on the modes of orbital angular momentum, magnetic field fluctuations influence trapped ions, etc. Common sense tells that this renders the distributed quantum state useless and unrecognizable.

Here we provide a method for entanglement detection and analysis that is insensitive to local rotations and thus overcomes these difficulties. It requires neither reference frames nor alignment nor calibration of measuring devices. Still, it can both witness as well as classify multipartite entanglement in the presence of local unitary noise. The key to overcome the lack of control and knowledge regarding each single measurement is to harness uniform sampling of the entirety of all measurements. Especially without any prior knowledge about the state, the conceptually simple method of random sampling proves highly beneficial for entanglement detection and state analysis.

Previous work on entanglement detection relaxing the requirement of fully aligned reference frames first considered the absence of a shared reference frame, but still

required the ability to choose or at least to repeat local measurement settings from a given set in order to detect, for example, the violation of a Bell inequality [1–6], or for tomographic reconstruction [7]. Under the same constraints, also adaptive methods for entanglement detection have been developed [8, 9]. In the absence of any reference frames Bell violations can be measured with some probability [10, 11] and entanglement can be detected by evaluating the second moment of the distribution of correlations obtained by measuring random observables on each subsystem [12–17]. Furthermore, it has been shown recently that higher-order moments of this distribution allow discrimination of very specific types of multipartite entanglement [18]. While these methods analyze full correlations, a recent experiment used second moments of subsets to deduce entanglement in systems of more than ten particles [19].

Our analysis method considers not only a specific moment of the distributions of full correlations, but all probability distributions of full as well as of marginal correlations taking into account their interdependencies. We show for specific pure states that this provides a detailed picture of the type of state and its entanglement structure. Furthermore, we derive general witnesses of genuine multipartite entanglement for pure and mixed states. We experimentally measure such distributions of various multiqubit states using reference frame free random measurements and show the applicability of the extended analysis methods. These methods are robust as they do not depend on the local unitary noise as long as the rate of generated entangled states is high enough to estimate the correlations for a momentarily constant noise.

## II. SCENARIO

Consider a source producing copies of an unknown  $n$ -qubit state  $\varrho$ , which is transmitted through unstable quantum channels to  $n$  local observers (Fig. 1). During the  $j$ -th transmission the state  $\varrho$  is transformed by  $n$  random local unitary operators  $U_i^{(j)}$  with  $i = 1, 2, \dots, n$  according to

$$\varrho \rightarrow \varrho^{(j)} = U_1^{(j)} \otimes \dots \otimes U_n^{(j)} \varrho U_1^{(j)\dagger} \otimes \dots \otimes U_n^{(j)\dagger}. \quad (1)$$

Additionally, each of the  $n$  observers is free to choose an arbitrary measurement setting  $\sigma_i^{(j)}$  to measure her qubit. If for each transmitted copy of  $\varrho$  the transformations  $U_i^{(j)}$  change significantly, all information about the state is lost. However, in a very common scenario encountered by experimenters the unitary noise has a timescale which is sufficiently slow to obtain at least a few copies of  $\varrho$  which have been affected by essentially the same noise, i.e., by the same set of local transformations  $U_i^{(j)}$ . In this case the transformations are still much too fast to apply standard techniques of state analysis [20], yet, it becomes possible to use the few equally transformed states to reliably record correlations

$$\begin{aligned} E_{1\dots n}^{(j)} &= \text{tr} \left( \sigma_1^{(j)} \otimes \sigma_2^{(j)} \otimes \dots \otimes \sigma_n^{(j)} \varrho^{(j)} \right) \\ &= \text{tr} \left( \tilde{\sigma}_1^{(j)} \otimes \tilde{\sigma}_2^{(j)} \otimes \dots \otimes \tilde{\sigma}_n^{(j)} \varrho \right), \end{aligned} \quad (2)$$

where each observer is keeping her local observable  $\sigma_i^{(j)}$  constant in the timescale of constant noise, which results in the effective random observable  $\tilde{\sigma}_i^{(j)} \equiv U_i^{(j)\dagger} \sigma_i U_i^{(j)}$ . Note that here and below the index  $j$  refers to a set of transmitted states which have all been affected by the same noise transformations and measured using the same settings.

We refer to  $E_{1\dots n}^{(j)}$  as “full correlation” or  $n$ -partite correlation because it involves measurement outcomes of all  $n$  observers. Besides full correlations, also “marginal correlations” can be measured, which are computed from

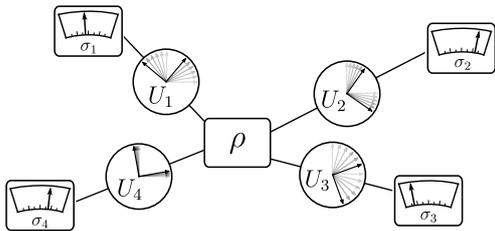


FIG. 1. Quantum communication over noisy channels. A source produces an entangled state of, say, four qubits. Each of them propagates through a noisy channel resulting in an unknown unitary transformation. When choosing local observables  $\sigma_i$  uniformly at random, the statistics of correlations reveal detailed information on multipartite entanglement, independently of the noise in the channels or of the lack of shared reference frames.

the outcomes of a subset of observers. For example, the marginal correlation of all observers but the first one is

$$E_{2\dots n}^{(j)} = \text{tr} \left( \mathbb{1} \otimes \tilde{\sigma}_2^{(j)} \otimes \dots \otimes \tilde{\sigma}_n^{(j)} \varrho \right). \quad (3)$$

The essential ingredient in our approach is to sample all local measurement directions  $\tilde{\sigma}_i^{(j)}$ , for each observer randomly according to a Haar uniform distribution. This removes any dependence of the obtained information on the actual structure or time dependence of the various  $U_i^{(j)}$  and thus overcomes any bias in the random noise.

In our experiment we prepare four different four-qubit states using entangled photon pairs, where we encode two qubits in the polarization degree of freedom and two qubits in the path degree of freedom. To comprehensively demonstrate the informational content of distributions of random correlations, we consider four quantum states belonging to different entanglement classes, in particular a tri-separable, a bi-separable and two genuinely multipartite entangled states, namely a Greenberger-Horne-Zeilinger (GHZ) state and a cluster state,

$$|\psi_{\text{trisep}}\rangle \propto (|00\rangle + |11\rangle) \otimes |0\rangle \otimes |0\rangle, \quad (4a)$$

$$|\psi_{\text{bisep}}\rangle \propto (|00\rangle + |11\rangle) \otimes (\sin \varphi |00\rangle + \cos \varphi |11\rangle), \quad (4b)$$

$$|\text{GHZ}\rangle \propto (|0000\rangle + |1111\rangle), \quad (4c)$$

$$|\mathcal{C}_4\rangle \propto (|0000\rangle + |0011\rangle - |1100\rangle + |1111\rangle). \quad (4d)$$

We utilize the full experimental control over the choice of measurement settings to emulate the local unitary transformations due to noisy channels and the Haar random choices of measurement settings. Details of the setup can be found in [21]. It should be noted that while we clearly can deduce how characteristics of the state are reflected in the form of the distribution the other direction of deduction is in general much more difficult.

## III. ANALYZING ENTANGLEMENT STRUCTURES

In the following we study distributions of random correlations for these four states, see Fig. 2. It is helpful to recall that for some particular states the distributions are known analytically. A pure product state of  $n$  qubits results in a distribution proportional to  $-(\ln|E|)^{n-1}$  [16, 17], which becomes uniform for  $n = 1$ , and a maximally entangled state of two qubits gives rise to a flat distribution [22, 23]. In addition to this established knowledge, we use new criteria to show that the experimental data not only provide information about the amount of entanglement in the full state, but also give insight into how the entanglement is shared among the parties, allowing to reconstruct the whole multipartite entanglement structure. An important finding arises from the fact that for arbitrary product states of the subsystems  $A$  and  $B$  any full correlation value  $E_{AB}$  is the product of the corresponding marginal values with

$$\varrho_{AB} = \varrho_A \otimes \varrho_B \Rightarrow E_{AB} = E_A E_B. \quad (5)$$

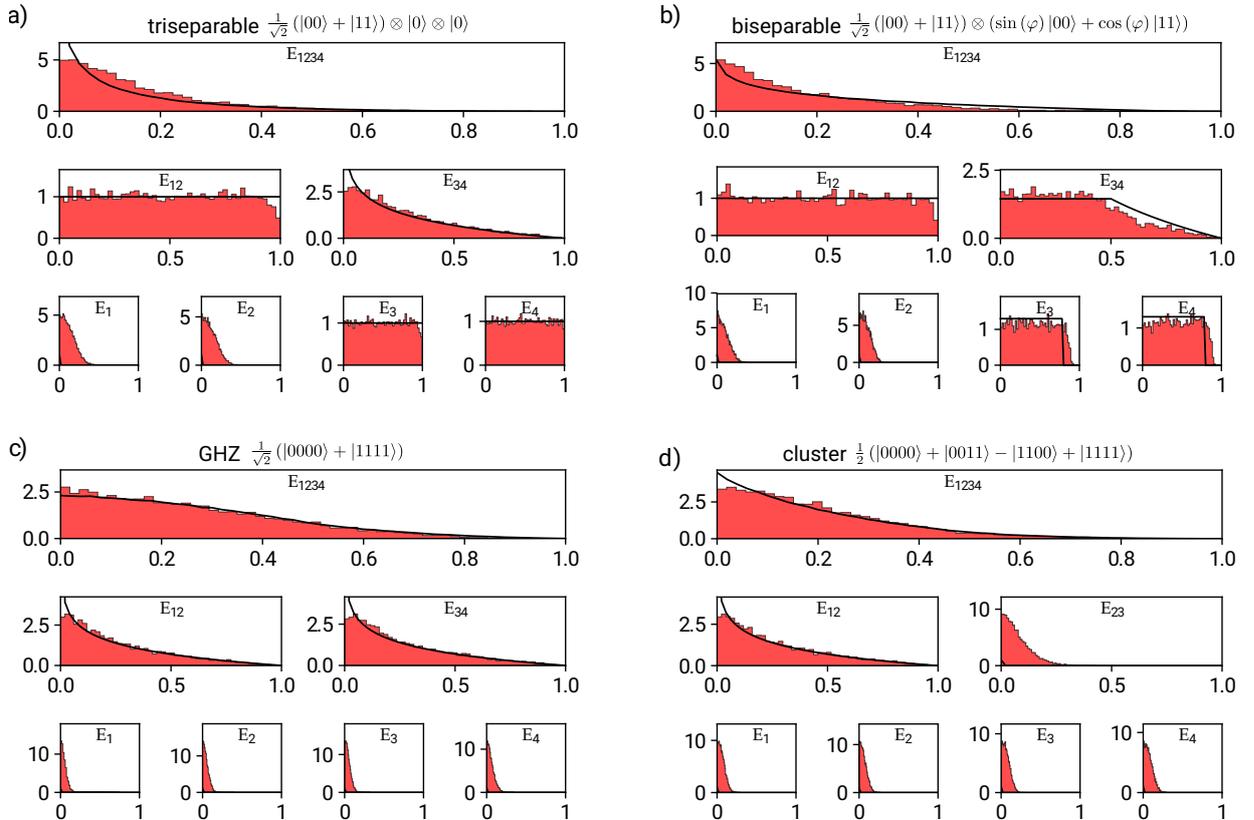


FIG. 2. Experimental distributions of correlations for four typical states. For each state we plot the distribution of the modulus of the measured full correlation  $E_{1234}$  together with two of the six two-qubit marginal distributions and all four single-qubit marginals. For this visualization, we measured each state along  $10^4$  different settings (in panel b only 6000 settings; we choose  $\varphi \approx 0.2$ ). The histograms are derived from raw measured data corrected for detection efficiencies. Solid lines represent theoretical curves for ideal states. Deviations of the measured data from the ideal distributions are due to finite statistics and finite fidelity of the state preparation.

This relation between single expectation values implies that the correlation distribution of parties  $AB$  is a so-called product distribution of measurement results obtained on  $A$  and  $B$ . Whenever this is not the case we can infer that, if the state is pure, it is entangled across the partition  $AB$ . Here, we first apply this criterion to product states, but later we will also generalize it for arbitrary mixed states.

Consider first the triseparable state in Fig. 2a. The bipartite distribution  $E_{34}$ , i.e., the distribution of the multiplication of outcomes for qubits 3 and 4, shows a logarithmic decay, which indicates a pure product state over these two parties. The bipartite distribution  $E_{12}$  is uniform as it is characteristic for maximally entangled two-qubit states. The single qubit marginals confirm this observation.  $E_3$  and  $E_4$  are almost uniform (pure states), whereas  $E_1$  and  $E_2$  correspond to the maximally mixed state. Ideally, the correlation function for the maximally mixed state is equal to zero and results in a delta peak around 0. Finite statistics causes a broadening of this

theoretical distribution and leads to the observed Gaussian shape. Several of the distributions are product distributions. For example we can verify that the full distribution  $E_{1234}$  is the product distribution of multiplied results obtained on qubits 12 and on qubits 34, and that  $E_{34}$  is the product distribution of the results on qubit 3 and on qubit 4. This is compatible with the state being separable across these partitions. On the other hand, clearly the distribution  $E_{12}$  is not a product one for the outcomes on qubit 1 and on qubit 2, which indicates the presence of entanglement.

The distributions for the biseparable state (4b) are shown in Fig. 2b. As expected, the bipartite marginal  $E_{12}$  is the same as for the triseparable state. The same also holds for the respective single qubit marginals of  $E_1$  and  $E_2$ . In the bipartite distribution of  $E_{34}$ , however, one can nicely observe the signature of a pure state intermediate between a maximally entangled and a product state, as tuned by the parameter  $\varphi$ . For  $\varphi \approx 0.2$ , the bipartite distribution of  $E_{34}$  is almost uniform until

approximately 0.5 and decays logarithmically for larger values. Equally, the respective single qubit marginals also show an intermediate behavior between a uniform distribution (pure state) until approximately 0.8 and vanishing (white noise) for values above. Both the distributions  $E_{12}$  and  $E_{34}$  do not correspond to the product distributions from the constituent systems which implies entanglement across these partitions of the pure state.

The maximally entangled GHZ state (Fig. 2c) and the cluster state (Fig. 2d) are not distinguishable on the level of the four respective single qubit marginals. Also certain bipartite marginals are the same, e.g., when tracing out qubits 3 and 4. However, while for the permutationally invariant GHZ state all marginal distributions for the same number of qubits must be the same, a significantly different distribution (corresponding to the maximally mixed state) can be obtained for the cluster state, when tracing out for example qubits 1 and 4, i.e., for  $E_{23}$ . Finally, the cluster and GHZ state can be distinguished also via their distributions of the full correlations. From the plotted distributions for these two states only the distribution  $E_{23}$  of the cluster state is (trivially) the product distribution for the results on qubits 2 and 3 (the same holds also for  $E_{13}$ ,  $E_{14}$ , and  $E_{24}$ ). All other distributions are not the product distributions and thus reveal entanglement.

While our data reflect the theoretical predictions based on Eqs. (4a-d) well, there are systematic differences which can be traced back chiefly to a broadening of the distributions due to finite statistics [24]. We used approximately 475 counts per estimated expectation value for the GHZ state, giving rise to the broadening of a normal distribution with standard deviation on the order of  $1/\sqrt{475} \approx 0.046$ . Accounting for these systematics is vital for the application of our quantitative analysis below and is explained in Appendix A.

#### IV. WITNESSING ENTANGLEMENT

To quantitatively analyze the experimentally obtained distributions, we focus on their statistical moments. The  $k$ -th moment of the distribution of the full correlation is defined as

$$m_{1\dots n}^{(k)} \equiv \int_{SU(2)^n} dU_1 \dots dU_n \operatorname{tr} (U^\dagger \sigma_z^{\otimes n} U \varrho)^k, \quad (6)$$

with  $U \equiv U_1 \otimes \dots \otimes U_n$  and where integration over  $SU(2)$  is equivalent to sampling measurement directions uniformly from the single qubit Bloch spheres. We will show in the following how to deduce the amount of purity and the presence of genuine multipartite entanglement using only the second moments of our measured correlation distributions. We denote the second moment simply by  $m_{1\dots n} \equiv m_{1\dots n}^{(2)}$ .

One of the most elementary properties of a quantum

state is its purity. For  $n$  qubits it is defined by

$$\mathcal{P}(\varrho) \equiv \operatorname{tr} (\varrho^2) = \frac{1}{2^n} \sum_{\mathcal{A} \in \mathbb{P}(\mathcal{S})} 3^{|\mathcal{A}|} m_{\mathcal{A}} \quad (7)$$

where  $\mathbb{P}(\mathcal{S})$  is the set of all subsets of  $\mathcal{S} = \{1, \dots, n\}$  and  $|\mathcal{A}|$  denotes the cardinality of the set  $\mathcal{A}$ . Clearly, purity is accessible in the experiment with random measurements and forms the basis of our methods for detecting multipartite entanglement. Note that in the case of a single qubit, the purity parameterizes the spectrum of the density matrix and hence any function of the quantum state which is invariant under local unitary transformations.

Let us consider the simplest case of pure two-qubit states. The second moments of any product state satisfy  $m_{12} = m_1 m_2$ . In consequence, the observation of  $m_{12} > m_1 m_2$  indicates entanglement for pure states. This reasoning cannot be easily extended to general states, since this inequality can also be satisfied for incoherent mixtures of product states. However, we have found a purity dependent tightening of the inequality such that any  $m_{12}$  above a certain purity dependent threshold must be due to quantum entanglement. In Appendix B we derive the following entanglement witness condition:

$$\mathcal{M}_2 \equiv m_{12} - m_1 m_2 \leq \begin{cases} 4(1 - \mathcal{P})\mathcal{P}/9 & \text{for } \mathcal{P} \geq \frac{1}{2}, \\ (4\mathcal{P} - 1)/9 & \text{for } \mathcal{P} < \frac{1}{2}. \end{cases} \quad (8)$$

It holds for all separable states of two qubits with purity  $\mathcal{P} \equiv \mathcal{P}(\varrho)$ . The bound is tight and achieved, e.g., by the state  $p|00\rangle\langle 00| + (1-p)|11\rangle\langle 11|$ . This powerful criterion can be generalized to the detection of genuine multipartite entanglement.

A state is genuinely multipartite entangled if it cannot be represented as a mixture of product states across any bipartition (such a mixture is called a bi-separable state). The left-hand side of Eq. (8) generalizes for an  $n$ -qubit state to

$$\mathcal{M}_n \equiv m_{\mathcal{S}} - \frac{1}{2} \sum_{\mathcal{A} \in \{\mathbb{P}(\mathcal{S}) \setminus \{S \cup \emptyset\}\}} m_{\mathcal{A}} m_{\mathcal{S} \setminus \mathcal{A}}, \quad (9)$$

where the factor of 1/2 resolves the issue of the double counting in the sum.

By numerical simulations, we find that the following condition holds for three-qubit bi-separable states

$$\mathcal{M}_3 = m_{123} - m_1 m_{23} - m_2 m_{13} - m_3 m_{12} \leq \frac{8}{27}(1 - \mathcal{P})\mathcal{P}. \quad (10)$$

We have verified this inequality by extensive numerical search described in Appendix C. The bound is tight for  $\mathcal{P} \geq \frac{1}{2}$  and is achieved by, e.g., the state  $p|\phi^+\rangle\langle \phi^+| \otimes |0\rangle\langle 0| + (1-p)|\phi^-\rangle\langle \phi^-| \otimes |1\rangle\langle 1|$  with the Bell states  $|\phi^\pm\rangle = \frac{1}{\sqrt{2}}(|00\rangle \pm |11\rangle)$ .

The bounds of the last two inequalities give hope for a simple dependence on the number of qubits. Unfortunately, already for four qubits we found by a numerical

study that the inequality satisfied by bi-separable states has a different functional dependence on the purity [25],

$$\mathcal{M}_4 \leq \frac{8}{81}(1 - \mathcal{P}^2). \quad (11)$$

This bound is also tight for  $\mathcal{P} \geq \frac{5}{8}$  and achieved by, e.g., the state  $p|\phi^+\rangle_{12}\langle\phi^+| \otimes |\phi^+\rangle_{34}\langle\phi^+| + (1-p)|\phi^+\rangle_{13}\langle\phi^+| \otimes |\phi^+\rangle_{24}\langle\phi^+|$ .

Any violation of inequality (10) or (11) indicates genuine multipartite entanglement between three or four qubits respectively. We emphasize that these criteria require only the second moments of the observed distributions.

Application of the conditions of Eqs. (8), (10) and (11) to experimental data (Fig. 3) indeed enables detection of genuine  $n$ -partite entanglement for various subsets of particles. For the cluster and the GHZ state, genuine 4-partite entanglement is revealed with Eq. (11) using  $\mathcal{M}_4 \approx 0.0330 > 0.0076$  and  $\mathcal{M}_4 \approx 0.0311 > 0.0099$ , respectively. The bi- and tri-separable states do not violate their respective bound. Investigating the entanglement properties for their marginal states, one can now prove the entanglement for the 12-marginal and the 34-marginal of the biseparable state as well as the 12-marginal of the tri-separable state. It is therefore possible to conclude that the biseparable state contains contributions of at least  $\varrho_{12} \otimes \varrho_{34}$ , with entanglement between 1 and 2 and between 3 and 4, and the tri-separable state contains  $\varrho_{12} \otimes \varrho_3 \otimes \varrho_4$ . Note that the state could also contain genuine 4-partite entanglement, which was not revealed by  $\mathcal{M}_4$ .

## V. CONCLUSION

This work introduces a scheme to detect genuine multipartite entanglement and reveal its detailed structure in the absence of any reference frames and even for strongly fluctuating channels. Key to this method is to subject a multipartite quantum system to randomly chosen local measurements and to analyze full and marginal correlations between all local results using second moments of respective correlation distributions. Haar random sampling removes any bias of the noise and, provided that the generation rate of multiqubit states is higher than the rate of fluctuations along the channel, neither the strength nor any characteristics of the noise matter. The power of our procedure is demonstrated here by reconstructing the entanglement structure of various, experimentally prepared photonic four-qubit states. From this, many more interesting questions arise, e.g., whether it is possible to - up to suitable transformations - tomographically reconstruct quantum states or characterize quantum processes in our scenario of fully randomized local measurement directions.

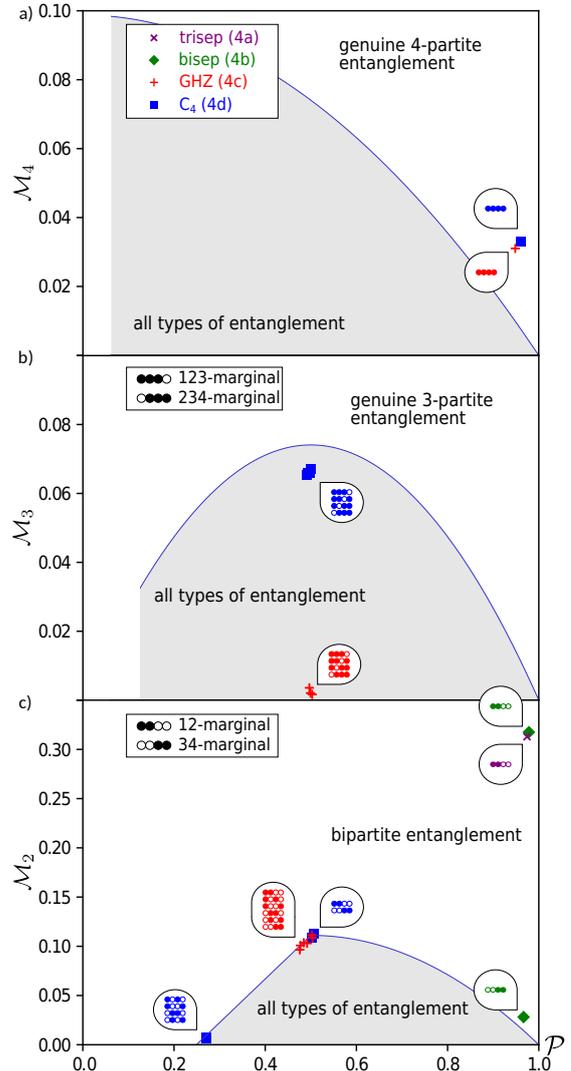


FIG. 3. Analyzing the entanglement structure using  $\mathcal{M}_i$ : (a)  $\mathcal{M}_4$  of the GHZ state (4c) (red plus) and the cluster state (4d) (blue square) are violating the bound for biseparable states (11), clearly indicating genuine 4-partite entanglement. The negative values for  $\mathcal{M}_4$  of the tri- and the biseparable states are not shown. (b) Evaluation of  $\mathcal{M}_3$  for tripartite marginals for these states does not indicate any genuine tripartite entanglement as expected, as no point is found above the threshold given in Eq. (10). The filled and non-filled circles indicate the type of marginals giving rise to different values of  $\mathcal{M}_3$ . (c)  $\mathcal{M}_2$  is shown for all bipartite marginals. The four-qubit biseparable state (4b) (green diamond) and the four-qubit tri-separable state (4a) (purple cross) have two and one marginals, respectively, which themselves are shown to be two-qubit entangled. The shaded regions contain all types of quantum states, irrespective of their entanglement properties. All error bars are smaller than the markers.

### ACKNOWLEDGMENTS

We thank Otfried Gühne, Felix Huber and Nikolai Wyderka for fruitful discussions. This research was supported by the DFG (Germany) and NCN (Poland) within the joint funding initiative “Beethoven 2” (2016/23/G/ST2/04273, 381445721), and by the DFG under Germany’s Excellence Strategy EXC-2111 390814868. We acknowledge the Singapore Ministry of Education Academic Research Fund Tier 2, Project No. MOE2015-T2-2-034. W.L. acknowledges partial support from the Foundation for Polish Science (IRAP project IC-TQT, Contract No. 2018/MAB/5, cofinanced by EU via Smart Growth Operational Programme). JD and LK acknowledge support from the PhD programs IMPRS-QST and ExQM, respectively.

#### Appendix A: Finite sample size correction

In our experiment, two different types of statistical effects have to be taken into account. On one hand, for obtaining the distributions as in Fig. 2, a finite number  $N_s$  of measurement settings ( $N_s = 10\,000$  in our case) is used. This leads to an uncertainty in estimating the second moments  $m_A \equiv m_A^{(2)}$ . This statistical error can be approximated by

$$\left(\Delta m_A^{(2)}\right)^2 = \frac{1}{N_s} \left[ m_A^{(4)} - \frac{N_s - 3}{N_s - 1} \left(m_A^{(2)}\right)^2 \right], \quad (\text{A1})$$

which describes the variance of the sample variance.

On the other hand, each correlation  $E_A^{(j)} \equiv E$  is obtained by performing  $N_c$  measurements in the same setting. Due to this finite sample size, for each expectation value in general we do not obtain the ideal result  $E_R$ , but measure a value  $E_M$  at random from a conditional probability distribution  $p(E_M|E_R)$ , approximately given by the Gaussian

$$p(E_M|E_R) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(E_M - E_R)^2}{2\sigma^2}\right) \quad (\text{A2})$$

centered around  $E_R$  with  $\sigma = \sqrt{1 - E_R^2}/\sqrt{N_c}$ , see, e.g., [24].

This statistical deviation leads to an overestimation of  $m_A$ . We mitigate this systematic inaccuracy by taking into account the well known statistical effect from Eq. (A2). Employing Bayesian methods, we are able to obtain  $p(E_R|E_M)$  from  $p(E_M|E_R)$  allowing to calculate  $m_A$  with reduced bias as

$$\begin{aligned} m_A &= \int_{-1}^1 dE_R p(E_R) E_R^2 \\ &= \int_{-1}^1 dE_R \int_{-1}^1 dE_M p(E_R|E_M) p(E_M) E_R^2. \end{aligned} \quad (\text{A3})$$

Bayes’ theorem provides  $p(E_R|E_M)$  as

$$\begin{aligned} p(E_R|E_M) &= \frac{p(E_M|E_R)\tilde{p}(E_R)}{p(E_M)} \\ &= \frac{p(E_M|E_R)\tilde{p}(E_R)}{\int_{-1}^1 dE'_R p(E_M|E'_R)\tilde{p}(E'_R)}, \end{aligned} \quad (\text{A4})$$

where  $\tilde{p}(E_R)$  represents the prior assumption about the unknown distribution  $p(E_R)$ . For our evaluation we use the measured distribution  $p(E_M)$  as the prior guess about  $p(E_R)$  and obtain an updated distribution according to the statistical analysis above. This distribution is used to evaluate the moments.

#### Appendix B: Two-qubit condition

Here we prove Eq. (8) of the main text. The problem is to maximize the value of  $\mathcal{M}_2 = m_{12} - m_1 m_2$  over separable states of two qubits with fixed purity  $\mathcal{P}$ . Any two-qubit state admits a decomposition

$$\varrho = \frac{1}{4} \sum_{\mu,\nu=0}^3 T_{\mu\nu} \sigma_\mu \otimes \sigma_\nu, \quad (\text{B1})$$

where  $T_{\mu\nu} = \text{tr}(\varrho \sigma_\mu \otimes \sigma_\nu)$ . In order to simplify numerical factors, we note that the second moments satisfy [16, 17]:

$$m_{12} = \frac{1}{9} \sum_{j,k=1}^3 T_{jk}^2 \equiv \frac{1}{9} \bar{m}_{12}, \quad (\text{B2})$$

$$m_1 = \frac{1}{3} \sum_{j=1}^3 T_{j0}^2 \equiv \frac{1}{3} \bar{m}_1, \quad (\text{B3})$$

$$m_2 = \frac{1}{3} \sum_{k=1}^3 T_{0k}^2 \equiv \frac{1}{3} \bar{m}_2. \quad (\text{B4})$$

The problem is therefore to maximize  $\bar{m}_{12} - \bar{m}_1 \bar{m}_2$  (and then multiply the result by  $\frac{1}{9}$ ). Using the definition of the purity results in

$$\bar{m}_{12} = 4\mathcal{P} - 1 - \bar{m}_1 - \bar{m}_2. \quad (\text{B5})$$

For the figure of merit we obtain

$$9\mathcal{M}_2 = 4\mathcal{P} - 1 - (\bar{m}_1 + \bar{m}_2 + \bar{m}_1 \bar{m}_2). \quad (\text{B6})$$

Since the purity is fixed we solve the following optimization

$$\text{minimize: } \bar{m}_1 + \bar{m}_2 + \bar{m}_1 \bar{m}_2, \quad (\text{B7})$$

$$\text{under condition: } \bar{m}_1 + \bar{m}_2 = 4\mathcal{P} - 1 - \bar{m}_{12}. \quad (\text{B8})$$

By the Lagrange multiplier method one finds that the minimum is achieved for  $\bar{m}_1 = \bar{m}_2 \equiv m$ . Since the function to be minimized,  $2m + m^2$ , is increasing with  $m$ , the minimum is achieved for the smallest  $m$  compatible with

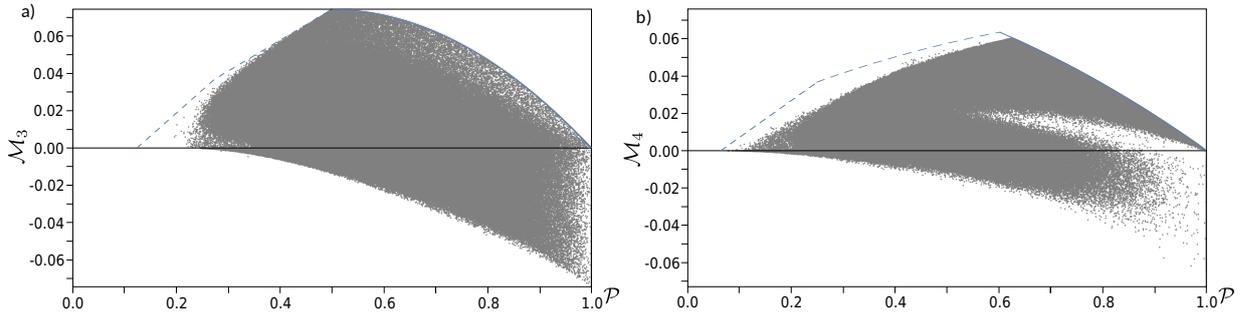


FIG. 4. Numerical evidence supports our witnesses of genuine tripartite and four-partite entanglement. We sampled more than  $10^6$  biseparable states from various (also random) families. The numerical boundary for biseparable states is plotted with a solid line, whereas the numerical boundary that holds for all quantum states (boundary of physicality) is plotted as a dashed line. (a) The biseparable states of 3 qubits are confined to the region below the boundary given by Eq. (C1). (b) The biseparable states of 4 qubits are confined to the region below the boundary given by  $\frac{8}{81}(1 - \mathcal{P}^2)$ .

the condition  $2m = 4\mathcal{P} - 1 - \bar{m}_{12}$ . All two-qubit separable states satisfy  $\bar{m}_{12} \leq 1$  [16], and hence the minimum is for  $m = \max(0, 2\mathcal{P} - 1)$ . There are therefore two cases. For  $\mathcal{P} \leq \frac{1}{2}$ , the figure of merit equals  $(4\mathcal{P} - 1)/9$  obtained by putting  $\bar{m}_1 = \bar{m}_2 = 0$  in Eq. (B6). For  $\mathcal{P} > \frac{1}{2}$  the figure of merit reads  $4\mathcal{P}(1 - \mathcal{P})/9$  obtained by putting  $\bar{m}_1 = \bar{m}_2 = 2\mathcal{P} - 1$  in Eq. (B6). Fig. 3c presents both of these bounds.

### Appendix C: Numerical simulations

Here we give numerical evidence for the bounds of Eqs. (10) and (11) of the main text.

We performed sampling of more than  $10^6$  biseparable

states and always found the bounds satisfied. Fig. 4 illustrates the results of numerical simulation.

For the case of three qubits we find the following improved boundary for small values of  $\mathcal{P}$ :

$$\mathcal{M}_3 \leq \begin{cases} (8\mathcal{P} - 1)/27 & \text{for } \mathcal{P} \in [\frac{1}{8}, \frac{1}{4}], \\ 4\mathcal{P}/27 & \text{for } \mathcal{P} \in (\frac{1}{4}, \frac{1}{2}], \\ 8(1 - \mathcal{P})\mathcal{P}/27 & \text{for } \mathcal{P} > \frac{1}{2}, \end{cases} \quad (\text{C1})$$

while the improved boundary for four qubits reads

$$\mathcal{M}_4 \leq \begin{cases} (16\mathcal{P} - 1)/81 & \text{for } \mathcal{P} \in [\frac{1}{16}, \frac{1}{4}], \\ 2(-8\mathcal{P}^2 + 16\mathcal{P} + 1)/243 & \text{for } \mathcal{P} \in (\frac{1}{4}, \mathcal{P}_0], \\ 8(1 - \mathcal{P}^2)/81 & \text{for } \mathcal{P} > \mathcal{P}_0, \end{cases} \quad (\text{C2})$$

where  $\mathcal{P}_0 = \frac{-4+3\sqrt{3}}{2} \approx 0.60$ .

- 
- [1] J. J. Wallman, Y.-C. Liang, and S. D. Bartlett, Generating nonclassical correlations without fully aligning measurements, *Phys. Rev. A* **83**, 022110 (2011).
  - [2] J. J. Wallman and S. D. Bartlett, Observers can always generate nonlocal correlations without aligning measurements by covering all their bases, *Phys. Rev. A* **85**, 024101 (2012).
  - [3] M. S. Palsson, J. J. Wallman, A. J. Bennet, and G. J. Pryde, Experimentally demonstrating reference-frame-independent violations of bell inequalities, *Phys. Rev. A* **86**, 032322 (2012).
  - [4] C. F. Senel, T. Lawson, M. Kaplan, D. Markham, and E. Diamanti, Demonstrating genuine multipartite entanglement and nonseparability without shared reference frames, *Phys. Rev. A* **91**, 052118 (2015).
  - [5] Y.-C. Liang, N. Harrigan, S. D. Bartlett, and T. Rudolph, Nonclassical correlations from randomly chosen local measurements, *Phys. Rev. Lett.* **104**, 050401 (2010).
  - [6] P. Shadbolt, T. Vertesi, Y.-C. Liang, C. Branciard, N. Brunner, and J. L. O'Brien, Guaranteed violation of a bell inequality without aligned reference frames or calibrated devices, *Sci. Rep.* **2**, 470 (2012).
  - [7] A. Elben, B. Vermersch, C. F. Roos, and P. Zoller, Statistical correlations between locally randomized measurements: A toolbox for probing entanglement in many-body quantum states, *Phys. Rev. A* **99**, 052323 (2019).
  - [8] W. Laskowski, D. Richart, C. Schwemmer, T. Paterek, and H. Weinfurter, Experimental schmidt decomposition and state independent entanglement detection, *Phys. Rev. Lett.* **108**, 240501 (2012).
  - [9] W. Laskowski, C. Schwemmer, D. Richart, L. Knips, T. Paterek, and H. Weinfurter, Optimized state-independent entanglement detection based on a geometrical threshold criterion, *Phys. Rev. A* **88**, 022327 (2013).
  - [10] A. de Rosier, J. Gruca, F. Parisio, T. Vertesi, and W. Laskowski, Multipartite nonlocality and random measurements, *Phys. Rev. A* **96**, 012101 (2017).
  - [11] A. Fonseca, A. de Rosier, T. Vertesi, W. Laskowski, and F. Parisio, Survey on the bell nonlocality of a pair of entangled qudits, *Phys. Rev. A* **98**, 042105 (2018).

- [12] A. S. M. Hassan and P. S. Joag, Separability criterion for multipartite quantum states based on the bloch representation of density matrices, *Quant. Inf. Comp.* **8**, 773 (2007).
- [13] A. S. M. Hassan and P. S. Joag, Experimentally accessible geometric measure for entanglement in n-qubit pure states, *Phys. Rev. A* **77**, 062334 (2008).
- [14] A. S. M. Hassan and P. S. Joag, Geometric measure for entanglement in n-qudit pure states, *Phys. Rev. A* **80**, 042302 (2009).
- [15] T. Lawson, A. Pappa, B. Bourdoncle, I. Kerenidis, D. Markham, and E. Diamanti, Reliable experimental quantification of bipartite entanglement without reference frames, *Phys. Rev. A* **90**, 042336 (2014).
- [16] M. C. Tran, B. Dakić, F. Arnault, W. Laskowski, and T. Paterek, Quantum entanglement from random measurements, *Phys. Rev. A* **92**, 050301(R) (2015).
- [17] M. C. Tran, B. Dakić, W. Laskowski, and T. Paterek, Correlations between outcomes of random measurements, *Phys. Rev. A* **94**, 042302 (2016).
- [18] A. Ketterer, N. Wyderka, and O. Gühne, Characterizing multipartite entanglement with moments of random correlations, *Phys. Rev. Lett.* **122**, 120505 (2019).
- [19] T. Brydges, A. Elben, P. Jurcevic, B. Vermersch, C. Maier, B. P. Lanyon, P. Zoller, R. Blatt, and C. F. Roos, Probing rényi entanglement entropy via randomized measurements, *Science* **364**, 260 (2019).
- [20] D. F. V. James, P. G. Kwiat, W. J. Munro, and A. G. White, Measurement of qubits, *Phys. Rev. A* **64**, 052312 (2001).
- [21] L. Knips, C. Schwemmer, N. Klein, M. Wieśniak, and H. Weinfurter, Multipartite entanglement detection with minimal effort, *Phys. Rev. Lett.* **117**, 210504 (2016).
- [22] J. D. A. Meinecke, *Quantum Correlations in Multi-Photon Quantum Walks*, Ph.D. thesis, University of Bristol (2014).
- [23] P. Shadbolt, *Complexity and Control in Quantum Photonics* (Springer International Publishing, 2016).
- [24] L. Knips, C. Schwemmer, N. Klein, J. Reuter, G. Tóth, and H. Weinfurter, How long does it take to obtain a physical density matrix?, *ArXiv e-prints* (2015), arXiv:1512.06866 [quant-ph].
- [25] A straightforward generalization from the previous bounds gives  $(2/3)^4(1 - \mathcal{P})\mathcal{P}$ . However, there exist bi-separable four-qubit states that violate this hypothetical bound.

## 4.4 Entanglement without Correlations

Although the connection between entanglement and correlations seems to be at the very heart of quantum mechanics, the usual definition of (quantum) correlations can lead to unexpected situations. In the following sections, the example of the so-called *no-correlation* state, a genuinely  $n$ -partite entangled state without any  $n$ -partite correlations [197] according to the definition of Eq. (2.17), will be introduced. Using this example, the concept of predictability of outcomes will be discussed, indicating that, although the  $n$ -partite correlations might vanish, the knowledge of the measurement outcomes of all but one party can increase the knowledge about the measurement outcome for the measurement of the last party. Subsequently, the role of classical and quantum correlations will be addressed including a discussion of reasonable postulates for a proper definition of quantum correlations. Finally, in Sec. 4.4.6, which is a reprint of [P6], a construction scheme for states with odd number of particles is given, which allows to obtain rank-2 states without correlations. A more comprehensive discussion including the role of the number of qubits and the rank of the state is given in [P7], which is reprinted in Sec. 4.4.7.

### 4.4.1 Entanglement and Correlations

Consider the tripartite  $|W_3\rangle$  state with

$$|W_3\rangle = \frac{1}{\sqrt{3}} (|001\rangle + |010\rangle + |100\rangle). \quad (4.42)$$

This state exhibits nonvanishing correlations in 19 basis combinations when all three observers or a subset of them jointly measure. These correlations are given in Tab. 4.2.

TABLE 4.2: Correlations of the state  $|W_3\rangle = \frac{1}{\sqrt{3}} (|001\rangle + |010\rangle + |100\rangle)$ .  $\Pi$  denotes permutations, e.g.,  $\Pi(\langle\sigma_0 \otimes \sigma_0 \otimes \sigma_z\rangle) = \{\langle\sigma_0 \otimes \sigma_0 \otimes \sigma_z\rangle, \langle\sigma_0 \otimes \sigma_z \otimes \sigma_0\rangle, \langle\sigma_z \otimes \sigma_0 \otimes \sigma_0\rangle\}$ . Except of the trivial correlation  $\langle\sigma_0 \otimes \sigma_0 \otimes \sigma_0\rangle = 1$ , all correlations which are not shown vanish.

$\Pi(\langle\sigma_0 \otimes \sigma_0 \otimes \sigma_z\rangle)$	$\Pi(\langle\sigma_0 \otimes \sigma_x \otimes \sigma_x\rangle)$	$\Pi(\langle\sigma_z \otimes \sigma_x \otimes \sigma_x\rangle)$
$\frac{1}{3}$	$\frac{2}{3}$	$\frac{2}{3}$
$\Pi(\langle\sigma_0 \otimes \sigma_z \otimes \sigma_z\rangle)$	$\Pi(\langle\sigma_0 \otimes \sigma_y \otimes \sigma_y\rangle)$	$\Pi(\langle\sigma_z \otimes \sigma_y \otimes \sigma_y\rangle)$
$-\frac{1}{3}$	$\frac{2}{3}$	$\frac{2}{3}$
$\langle\sigma_z \otimes \sigma_z \otimes \sigma_z\rangle$		
-1		

If one compares these correlations with the correlations of the state

$$|\overline{W}_3\rangle = \frac{1}{\sqrt{3}} (|110\rangle + |101\rangle + |011\rangle), \quad (4.43)$$

which are given in Tab. 4.3, one realizes that the full correlations, which include all three observers, as well as local correlations including only a single observer are exactly opposite, while bipartite correlations are identical.

TABLE 4.3: Nontrivial correlations of the state  $|\overline{W}_3\rangle = \frac{1}{\sqrt{3}}(|110\rangle + |101\rangle + |011\rangle)$ .

$\Pi(\langle\sigma_0 \otimes \sigma_0 \otimes \sigma_z\rangle)$	$\Pi(\langle\sigma_0 \otimes \sigma_x \otimes \sigma_x\rangle)$	$\Pi(\langle\sigma_z \otimes \sigma_x \otimes \sigma_x\rangle)$
$-\frac{1}{3}$	$\frac{2}{3}$	$\frac{2}{3}$
$\Pi(\langle\sigma_0 \otimes \sigma_z \otimes \sigma_z\rangle)$	$\Pi(\langle\sigma_0 \otimes \sigma_y \otimes \sigma_y\rangle)$	$\Pi(\langle\sigma_z \otimes \sigma_y \otimes \sigma_y\rangle)$
$-\frac{1}{3}$	$\frac{2}{3}$	$\frac{2}{3}$
$\langle\sigma_z \otimes \sigma_z \otimes \sigma_z\rangle$		
1		

As the correlation tensor of a mixed state is given by the convex combination of the correlation tensors of the states which are mixed,  $T_{\text{mixed}} = p_1 T_1 + p_2 T_2$  for  $\varrho_{\text{mixed}} = p_1 \varrho_1 + p_2 \varrho_2$ , the equally weighted mixture of  $|W_3\rangle$  and  $|\overline{W}_3\rangle$  results in a state with bipartite correlations identical to those of  $|W_3\rangle$  and of  $|\overline{W}_3\rangle$ , while local and tripartite correlations vanish, see Tab. 4.4. This state will be called the *no-correlation state*

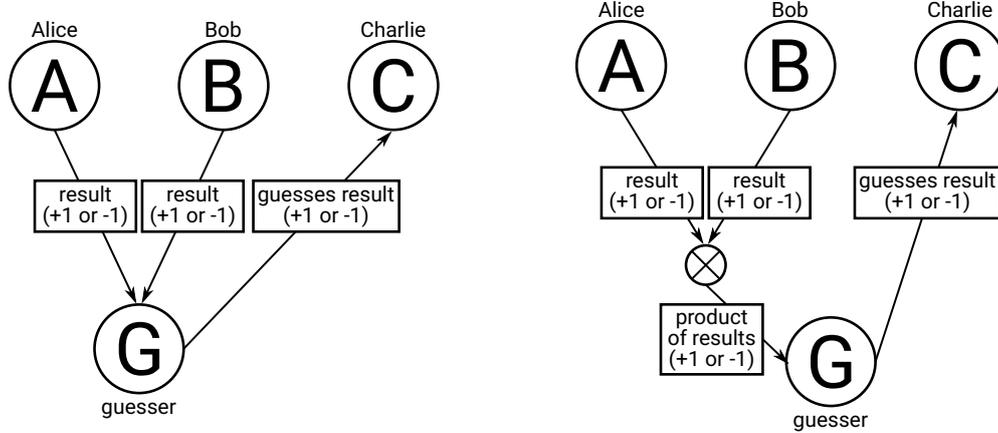
$$\varrho_{\text{nc}} = \frac{1}{2} (|W_3\rangle\langle W_3| + |\overline{W}_3\rangle\langle \overline{W}_3|). \quad (4.44)$$

TABLE 4.4: Nontrivial correlations of the state  $\varrho_{\text{nc}} \equiv \frac{1}{2} (|W_3\rangle\langle W_3| + |\overline{W}_3\rangle\langle \overline{W}_3|)$ . Local and tripartite correlation measurements are indistinguishable from the respective white noise,  $\frac{1}{2}$  and  $\frac{1}{8}$ , as all correlations which include one or three observers vanish. Obviously, this state is of rank 2.

$\Pi(\langle\sigma_0 \otimes \sigma_0 \otimes \sigma_z\rangle)$	$\Pi(\langle\sigma_0 \otimes \sigma_x \otimes \sigma_x\rangle)$	$\Pi(\langle\sigma_z \otimes \sigma_x \otimes \sigma_x\rangle)$
0	$\frac{2}{3}$	0
$\Pi(\langle\sigma_0 \otimes \sigma_z \otimes \sigma_z\rangle)$	$\Pi(\langle\sigma_0 \otimes \sigma_y \otimes \sigma_y\rangle)$	$\Pi(\langle\sigma_z \otimes \sigma_y \otimes \sigma_y\rangle)$
$-\frac{1}{3}$	$\frac{2}{3}$	0
$\langle\sigma_z \otimes \sigma_z \otimes \sigma_z\rangle$		
0		

Interestingly, as all tripartite correlations vanish, the correlations of this state are indistinguishable from white noise, when only considering measurements involving all three parties. However, as one can show, this state is still genuinely tripartite entangled [P6].

## 4.4.2 Correlations and Predictability



(A) Alice and Bob inform the guesser about their respective measurement result. Based on this knowledge, the guesser may guess Charlie's outcome.

(B) The guesser does not know the individual measurement results of A and B, but only the product of them. This corresponds to the definition of correlations.

FIGURE 4.7: Correlations and predictability in a tripartite scenario. The guesser wants to predict the measurement outcome of Charlie's measurement based on information he or she obtains from the results of Alice's and Bob's measurement given the knowledge of the experimentally prepared quantum state.

Consider the scenario that some external party, which may be called the *guesser*, wants to predict the measurement outcome of the third observer (Charlie) given some information about the respective measurement outcomes of the first two observers (Alice and Bob) and the experimentally prepared quantum state, see Fig. 4.7. For example, assume that the state is  $\varrho_{\text{nc}} = \frac{1}{2} (|W_3\rangle\langle W_3| + |\overline{W}_3\rangle\langle \overline{W}_3|)$  with the correlations given in Tab. 4.4 as first proposed in [197].

In the first scenario, Alice and Bob tell the guesser their respective measurement result +1 or -1, see Fig. 4.7a.  $\varrho_{\text{nc}}$  written in the basis of eigenstates of  $\sigma_z^{\otimes 3}$  reads

$$\varrho_{\text{nc}} = \frac{1}{6} [ (|001\rangle + |010\rangle + |100\rangle) (\langle 001| + \langle 010| + \langle 100|) + \quad (4.45)$$

$$(|110\rangle + |101\rangle + |011\rangle) (\langle 110| + \langle 101| + \langle 011|) ]. \quad (4.46)$$

If the state is also measured in the basis of the eigenstates of  $\sigma_z^{\otimes 3}$ , only the populations, i.e., the diagonal terms, are relevant, while cross terms such as  $|001\rangle\langle 010|$  can be neglected. For measurements along the respective  $\sigma_z$  direction, there are 4 different combinations of results for Alice and Bob ( $-1 \wedge -1$ ,  $-1 \wedge 1$ ,  $1 \wedge -1$ , and  $1 \wedge 1$ ) in the notation  $\sigma_z|0\rangle = |0\rangle$  (outcome 1) and  $\sigma_z|1\rangle = -|1\rangle$  (outcome -1). If Alice and Bob obtain different results, the

guesser cannot infer anything about the results of Charlie's measurement since the terms  $|010\rangle\langle 010|$  and  $|011\rangle\langle 011|$  (as well as  $|100\rangle\langle 100|$  and  $|101\rangle\langle 101|$ ) appear with the same probability. However, if Alice and Bob both observe  $-1$  ( $+1$ ), the guesser can conclude with certainty that Charlie's result will be  $+1$  ( $-1$ ). Thus, in two of six cases, Charlie's measurement outcome can be predicted with certainty, while in the remaining four cases, the prediction of Charlie's outcome corresponds to a coin flip. Hence, the probability for successfully predicting his outcome is  $\frac{1}{6} (1 + 1 + \frac{1}{2} + \frac{1}{2} + \frac{1}{2} + \frac{1}{2}) = \frac{2}{3}$ .

In contrast, the usual definition of a *correlation* does not account for single observer's outcomes, but only for the parity of the product of all outcomes. This is resembled in the scenario shown in 4.7b. The guesser cannot access the individual outcomes, but only their product. Then, the two cases  $-1 \wedge -1$  and  $1 \wedge 1$  are equivalent for the guesser (product 1) and no information about Charlie's result is obtained. Therefore, for the state  $\varrho_{nc}$ , no inference (guessing better than tossing coins) is possible. Charlie's result can still be  $+1$  or  $-1$  and can hence not be predicted. Thus, the usual definition of a correlation does not incorporate all information which might be accessible due to the single observers' measurements.

### 4.4.3 Postulates for Correlations

In [198] and a series of subsequent papers, the distinction between classical, quantum and total correlations has been discussed [199]. For bipartite systems, the quantum mutual information [200] provides a quantification for the total correlations between the subsystems  $\varrho_A = \text{tr}_B(\varrho_{AB})$  and  $\varrho_B = \text{tr}_A(\varrho_{AB})$  in a total state  $\varrho_{AB}$ . It is defined as

$$\mathcal{I}(\varrho) = \mathcal{S}(\varrho_{AB}) - \mathcal{S}(\varrho_A) - \mathcal{S}(\varrho_B) \quad (4.47)$$

with the von Neumann entropy  $\mathcal{S}(\varrho) = -\text{tr}(\varrho \log_2 \varrho)$  [201] as the quantum analog of the Shannon entropy [202]. The classical correlation can be obtained by a generalization of the classical mutual information based on the entropy maximizing over POVM [198].

In order to find a good criterion for genuine multipartite correlation, the authors of Ref. [199] phrased three postulates:

1. Adding a party to a state without  $n$ -partite correlations may not result in  $n+1$ -partite correlations.
2. Application of trace non-increasing local operators, which includes local unitary transformation, but also postselection, may not generate  $n$ -partite correlation from a state without  $n$ -partite correlations.
3. Splitting a subsystem of a state without  $n$ -partite correlations into two and creating a  $n+1$ -st subsystem may not create  $n+1$ -partite correlations.

The cumulants [203] and the correlations as defined in Eq. (2.17) do not satisfy those requirements [199], see also [P6]. This manifest discrepancy is best illustrated by the fact that for the no-correlation state all *full correlations* defined this way vanish, while genuine

tripartite entanglement persists. It is still an open question how to define multipartite quantum correlations in a consistent, operationally meaningful way [P6] and [P7].

#### 4.4.4 Even and Odd Number of Qubits

One possibility to generally obtain an  $n$ -qubit state without full correlations according to the standard definition of, e.g., Eq. (2.17), is by creating an equal mixture between a first initial state  $|\psi\rangle$  (or in general a mixed state  $\varrho$ ) and the corresponding antistate  $|\bar{\psi}\rangle$  (or  $\bar{\varrho}$ ). The antistate is a state which has opposite  $n$ -partite correlations, which thus leads for equal mixture of the two states to a state with vanishing  $n$ -partite correlations. One possibility to obtain such an antistate is by applying the *universal-not gate*  $N = \sigma_z \sigma_x K$  on every qubit, where  $K$  denotes conjugation in the computational basis, see [P6] and [P7] as well as the reprints thereof in Secs. 4.4.6 and 4.4.7.

However, this procedure does not work in general for pure states with an even number of qubits as can be shown by a counterexample. Nevertheless, also for an even number of qubits such *no-correlation states* exist, as is shown in [P7], see Sec. 4.4.7. These states are more difficult to find and need to have a rank larger than two. In the aforementioned publication, a constructive example of a four-qubit state of rank four with vanishing correlation functions is given as well as the result of a numerical procedure resulting in a four-qubit state with rank three. The constructive procedure can be generalized to any even number of qubits strictly larger than two. All two-qubit states with vanishing bipartite correlation functions cannot be entangled, thus, no entangled bipartite no-correlation state exists. All of the states given in [P6] and [P7] do not show full correlations in any basis combination and still exhibit genuine  $n$ -partite entanglement. As shown in [204], similar states do also exist for higher dimensions.

#### 4.4.5 Obtaining a State without Correlations

For the experimental implementation of the rank-2 no-correlation state  $\varrho_{\text{nc}}$ , the states  $|W_3\rangle$  and  $|\bar{W}_3\rangle$  have to be mixed equally. In general, the experimental production of mixed states is rather difficult in a coherent way (of course, there could be the possibility to randomly create either one or the other state). However, for this specific state, one can first produce a pure Dicke state  $|D_4^{(2)}\rangle = \frac{1}{\sqrt{6}}(|0011\rangle + |0101\rangle + \dots + |1100\rangle)$ . Projecting one of the qubits onto  $|1\rangle$  (and tracing over this qubit) results in the state  $|W_3\rangle$ ,

$$\text{tr}_4 \left[ \mathbb{1} \otimes \mathbb{1} \otimes \mathbb{1} \otimes |1\rangle\langle 1| |D_4^{(2)}\rangle\langle D_4^{(2)}| \right] = |W_3\rangle\langle W_3|. \quad (4.48)$$

Analogously, one can obtain  $|\bar{W}_3\rangle$ ,

$$\text{tr}_4 \left[ \mathbb{1} \otimes \mathbb{1} \otimes \mathbb{1} \otimes |0\rangle\langle 0| |D_4^{(2)}\rangle\langle D_4^{(2)}| \right] = |\bar{W}_3\rangle\langle \bar{W}_3|. \quad (4.49)$$

Thus, in order to obtain  $\varrho_{\text{nc}}$ , it is sufficient to trace out one qubit,

$$\text{tr}_4 \left[ |D_4^{(2)}\rangle\langle D_4^{(2)}| \right] = \varrho_{\text{nc}}. \quad (4.50)$$

The experimental preparation of the Dicke state is described in the following section, which is a reprint of [P6]. For more details on the state preparation, see also [132, 205].

These publications are based on ideas of Tomasz Paterek and Wiesław Laskowski together with Harald Weinfurter. Harald Weinfurter, Christian Schwemmer and me discussed and decided for experimental measurements of the three-qubit no-correlation state. Christian Schwemmer and me performed the experimental measurements and evaluated the obtained data. I theoretically analyzed a class of states in regards of its correlation content and derived some error model together with Christian Schwemmer in close collaboration with the other authors. Furthermore, I performed statistical analysis of the experimental results. I also conducted a numerical search for other no-correlation states, including some with higher rank and an even number of qubits. A similar simulation was central to the second publication, [P7]. The manuscript of [P6] was prepared by Christian Schwemmer, Wiesław Laskowski, Tomasz Paterek, Harald Weinfurter and myself, and edited by all authors. I co-edited the manuscript of [P7].

### Genuine Multipartite Entanglement without Multipartite Correlations

Christian Schwemmer,<sup>1,2</sup> Lukas Knips,<sup>1,2</sup> Minh Cong Tran,<sup>3</sup> Anna de Rosier,<sup>4</sup> Wiesław Laskowski,<sup>4</sup>  
Tomasz Paterek,<sup>3,5,6,\*</sup> and Harald Weinfurter<sup>1,2</sup>

<sup>1</sup>Max-Planck-Institut für Quantenoptik, Hans-Kopfermann-Straße 1, 85748 Garching, Germany

<sup>2</sup>Department für Physik, Ludwig-Maximilians-Universität, 80797 München, Germany

<sup>3</sup>School of Physical and Mathematical Sciences, Nanyang Technological University, Singapore, 637371 Singapore

<sup>4</sup>Institute of Theoretical Physics and Astrophysics, University of Gdańsk, PL-80-952 Gdańsk, Poland

<sup>5</sup>Center for Quantum Technologies, National University of Singapore, Singapore, 117543 Singapore

<sup>6</sup>MajuLab, CNRS-UNS-NUS-NTU International Joint Research Unit, Singapore, UMI 3654 Singapore

(Received 9 October 2014; published 6 May 2015)

Nonclassical correlations between measurement results make entanglement the essence of quantum physics and the main resource for quantum information applications. Surprisingly, there are  $n$ -particle states which do not exhibit  $n$ -partite correlations at all but still are genuinely  $n$ -partite entangled. We introduce a general construction principle for such states, implement them in a multiphoton experiment and analyze their properties in detail. Remarkably, even without multipartite correlations, these states do violate Bell inequalities showing that there is no classical, i.e., local realistic model describing their properties.

DOI: 10.1103/PhysRevLett.114.180501

PACS numbers: 03.67.Mn, 03.65.Ud, 42.50.Dv

Correlations between measurement results are the most prominent feature of entanglement. They made Einstein, Podolski, and Rosen [1] question the completeness of quantum mechanics and are nowadays the main ingredient for the many applications of quantum information like entanglement based quantum key distribution [2] or quantum teleportation [3].

Correlations enable us, e.g., when observing two maximally entangled qubits, to use a measurement result observed on the first system to infer exactly the measurement result on the second system. In this scenario, the two particle correlations are formally given by the expectation value of the product of the measurement results obtained by the two observers. Note, the single particle correlation, i.e., the expectation value of the results for one or the other particle are zero in this case. Consequently, we cannot predict anything about the individual results. When studying the entanglement between  $n$  particles, a natural extension is to consider  $n$ -partite correlations, i.e., the expectation value of the product of  $n$  measurement results. Such correlation functions are frequently used in classical statistics and signal analysis [4], moreover, in quantum information, almost all standard tools for analyzing multipartite systems like multiparty entanglement witnesses [5,6] and Bell inequalities [7,8] are based on the  $n$ -partite correlation functions.

Recently, Kaszlikowski *et al.* [9] pointed at a particular quantum state with vanishing multiparty correlations which, however, is genuinely multipartite entangled. This discovery, of course, prompted vivid discussions on a viable definition of classical and quantum correlations [10,11]. Still, the question remains what makes up such states with no full  $n$ -partite correlations and how nonclassical they can be, i.e., whether they are not only entangled but whether they also violate a Bell inequality.

Here, we generalize, highlight, and experimentally test such remarkable quantum states. We introduce a simple principle how to construct states without  $n$ -partite correlations for odd  $n$  and show that there are infinitely many such states which are genuinely  $n$ -partite entangled. We implement three and five qubit no-correlation states in a multiphoton experiment and demonstrate that these states do not exhibit  $n$ -partite correlations. Yet, due to the existence of correlations between a smaller number of particles, we observe genuine  $n$ -partite entanglement. Using our recently developed method to design  $n$ -partite Bell inequalities from lower order correlation functions only [12,13], we show that these states, despite not having full correlations, can violate Bell inequalities.

*Correlations.*—The quantum mechanical correlation function  $T_{j_1 \dots j_n}$  is defined as the expectation value of the product of the results of  $n$  observers

$$T_{j_1 \dots j_n} = \langle r_1 \dots r_n \rangle = \text{Tr}(\rho \sigma_{j_1} \otimes \dots \otimes \sigma_{j_n}), \quad (1)$$

where  $r_k$  is the outcome of the local measurement of the  $k$ th observer, parametrized by the Pauli operator  $\sigma_{j_k}$  with  $j_k \in \{x, y, z\}$ . Evidently, besides the  $n$ -partite correlations, for an  $n$ -partite state, one can also define  $l < n$  fold correlations  $T_{\mu_1 \dots \mu_l} = \text{Tr}(\rho \sigma_{\mu_1} \otimes \dots \otimes \sigma_{\mu_l})$  with  $\mu_i \in \{0, x, y, z\}$  and  $|\{\mu_i = 0\}| = n - l$ . Nonvanishing  $l$ -fold correlations indicate that we can infer (with higher probability of success than pure guessing) an  $l$ th measurement result from the *product* of the other  $(l - 1)$  results [see Supplemental Material [14]]. Only in the two particle scenario can we directly use the result from one measurement to infer the other result. For an  $n$ -qubit no-correlation state, the vanishing  $n$ -partite correlations do not imply vanishing correlations between a smaller number of observers, thus not necessarily destroying predictability. We will

see also in the experimentally implemented example that the various individual results still enable some possibility for inference, which is then largely due to bipartite correlations.

*Constructing no-correlation states.*—For any state  $|\psi\rangle$  with an odd number  $n$  of qubits, we can construct an “antistate”  $|\bar{\psi}\rangle$ , i.e., the state whose  $n$ -partite correlations are inverted with respect to the initial one. By evenly mixing these states

$$\rho_{\bar{\psi}}^{nc} = \frac{1}{2}|\psi\rangle\langle\psi| + \frac{1}{2}|\bar{\psi}\rangle\langle\bar{\psi}|, \quad (2)$$

we obtain a state  $\rho_{\bar{\psi}}^{nc}$  without  $n$ -partite correlations.

The antistate  $|\bar{\psi}\rangle$  of a state  $|\psi\rangle$  described in the computational basis by

$$|\psi\rangle = \sum_{k_1, \dots, k_n=0}^1 \alpha_{k_1, \dots, k_n} |k_1 \dots k_n\rangle, \quad (3)$$

with normalized coefficients  $\alpha_{k_1, \dots, k_n} \in \mathbb{C}$ , is given by

$$|\bar{\psi}\rangle \equiv \sum_{k_1, \dots, k_n=0}^1 (-1)^{k_1 + \dots + k_n} \alpha_{1-k_1, \dots, 1-k_n}^* |k_1 \dots k_n\rangle, \quad (4)$$

where the asterisk denotes complex conjugation. This state has inverted correlations with respect to those in  $|\psi\rangle$  for every odd number of observers, whereas all the correlation function values for an even number of observers remain unchanged.

$|\bar{\psi}\rangle$  is mathematically obtained from  $|\psi\rangle$  by applying local universal-not gates [24]. These gates introduce a minus sign to all local Pauli operators. Therefore, for odd  $n$ , the correlations of  $|\bar{\psi}\rangle$  have opposite sign to those of  $|\psi\rangle$ . Representing the universal-not gate by  $N = \sigma_z \sigma_x K$ , where  $K$  is the complex conjugation operating in the computational basis, i.e.,  $K(\alpha|0\rangle + \beta|1\rangle) = \alpha^*|0\rangle + \beta^*|1\rangle$ , indeed, we obtain  $N\sigma_x N^\dagger = -\sigma_x$ ,  $N\sigma_y N^\dagger = -\sigma_y$ , and  $N\sigma_z N^\dagger = -\sigma_z$ . Applying  $N$  to all the  $n$  subsystems, we find the anticipated result  $N \otimes \dots \otimes N |\psi\rangle = |\bar{\psi}\rangle$ .

Although  $N$  is antiunitary,  $|\bar{\psi}\rangle$  is always a proper physical state and can be obtained by some global transformation of  $|\psi\rangle$ . In general,  $N$  can be approximated [25], but if all the coefficients  $\alpha_{k_1, \dots, k_n}$  are real, complex conjugation can be omitted and no-correlation states can be generated by local operations.

This construction principle can be generalized to mixed states using  $\bar{\rho} = N^{\otimes n} \rho (N^{\otimes n})^\dagger$ , which changes every pure state in the spectral form to the respective antistate. Evenly mixing  $\rho$  and  $\bar{\rho}$  therefore produces a state with no  $l$ -party correlations for all odd  $l$ .

One may then wonder whether the principle of Eq. (2) can also be applied to construct a no-correlation state for every state with an even number of qubits. The answer is negative as shown by the following counterexample. Consider the Greenberger-Horne-Zeilinger state of an even number of qubits  $|\psi\rangle = (1/\sqrt{2})(|0\dots 0\rangle + |1\dots 1\rangle)$ . It has

nonvanishing  $T_{z\dots z}$ ,  $2^{n-1}$  multipartite correlations in the  $xy$  plane, and also,  $2^{n-1} - 1$  correlations between a smaller number of subsystems, all equal to  $\pm 1$ . However, for a state with inverted correlations between all  $n$  parties (making no assumptions about the correlations between smaller numbers of observers), the fidelity relative to the GHZ state, given by  $\frac{1}{2^n} \sum_{\mu_1, \dots, \mu_n=0}^3 T_{\mu_1 \dots \mu_n}^{\text{GHZ}} T_{\mu_1 \dots \mu_n}^{\text{anti}}$ , is negative because more than half of the correlations are opposite. Hence, this state is unphysical and there is no such “antistate”. In fact, so far we were unable to find an antistate to *any* genuinely multiqubit entangled state of even  $n$ .

*Entanglement without correlations: infinite family.*—Consider a three-qubit system in the pure state

$$|\phi\rangle = \sin\beta \cos\alpha |001\rangle + \sin\beta \sin\alpha |010\rangle + \cos\beta |100\rangle, \quad (5)$$

where  $\alpha, \beta \in (0, \pi/2)$  (which includes the state  $|W\rangle$  with  $\alpha = \pi/3$  and  $\beta = \cos^{-1}(1/\sqrt{3})$ ). Together with any local unitary transformation thereof, this defines a three dimensional subspace of genuinely tripartite entangled states within the eight dimensional space of three qubit states. To show that all the respective no-correlation states  $\rho_{\phi}^{nc}$  are genuinely entangled, we use a criterion similar to the one in [6], i.e.,

$$\max_{T^{\text{bi-prod}}} (T, T^{\text{bi-prod}}) < (T, T^{\text{exp}}) \Rightarrow \rho^{\text{exp}} \text{ is not biseparable}, \quad (6)$$

where maximization is over all biproduct pure states and  $(U, V) \equiv \sum_{\mu, \nu, \eta=0}^3 U_{\mu\nu\eta} V_{\mu\nu\eta}$  denotes the inner product in the vector space of correlation tensors. Condition [Eq. (6)] can be interpreted as an entanglement witness  $\mathcal{W} = \alpha \mathbb{1} - \rho_{\phi}^{nc}$ , where  $\alpha = L/8$  and  $L = \max_{T^{\text{bi-prod}}} (T, T^{\text{bi-product}})$  is the left-hand side of Eq. (6). In the ideal case of preparing  $\rho^{\text{exp}}$  perfectly,  $T^{\text{exp}} = T$ , the right-hand side of our criterion equals four for all the states of the family, and thus, the expectation value of the witness is given by  $\text{Tr}(\mathcal{W}\rho_{\phi}^{nc}) = (L - 4)/8$ .

A simple argument for  $\rho_{\phi}^{nc}$  being genuinely tripartite entangled can be obtained from the observation that  $|\phi\rangle$  and  $|\bar{\phi}\rangle$  span a two-dimensional subspace of the three qubit Hilbert space [9]. As none of the states  $|\Phi\rangle = a|\phi\rangle + b|\bar{\phi}\rangle$  is a biproduct (for the proof see Supplemental Material [14]), states in their convex hull do not intersect with the subspace of biseparable states and thus all its states, including  $\rho_{\phi}^{nc}$  are genuinely tripartite entangled. To evaluate the entanglement in the experiment, we calculated  $L$  for all states of Eq. (5). We obtain  $L_{|\phi\rangle} < 4$  in general, with  $L_{|W\rangle} = 10/3$ . Similar techniques were used to analyze five-qubit systems.

*Quantum correlations without classical correlations?*—The cumulants and correlations were initially proposed as a measure of genuinely multipart nonclassicality in Ref. [26]. Kaszlikowski *et al.* [9], however, showed that such a quantification is not sufficient as the state  $\rho_W^{nc}$  has

vanishing cumulants, yet contains genuinely multiparty entanglement. They suggested that the vanishing cumulants or standard correlation functions [Eq. (1)] indicate the lack of genuine multiparty “classical” correlations. This initiated a vivid discussion on a proper definition and measure of genuine multipartite “classical” and quantum correlations. Bennett *et al.* proposed a set of axioms for measures of genuine multipartite correlations [11]. They showed that the correlation function [Eq. (1)] does not fulfill all the requirements, but also still strive for computable measures that satisfy these axioms [15,27]. An information-theoretic definition of multipartite correlations was given by Giorgi *et al.* [15]. Their measure combines the entropy of all sizes of subsystems. Applying their definitions to  $\rho_W^{nc}$ , we obtain genuine classical tripartite correlations of 0.813 bit and genuine quantum tripartite correlations of 0.439 bit resulting in total genuine tripartite correlations of 1.252 bit (see Supplemental Material [14] for calculations for all  $\rho_\phi^{nc}$ ). While this approach does assign classical correlations in the context of Giorgi *et al.* [15] to  $\rho_W^{nc}$ , it does not fulfill all requirements of [11] either.

*Experiment.*—The three photon state  $|W\rangle$  can be observed either using a multiphoton interferometer setup [28] or by suitably projecting the fourth photon of a 4-photon symmetric Dicke state [29]. The latter scheme has the advantage that it also offers the option to prepare the states  $|\bar{W}\rangle$  and  $\rho_W^{nc}$ . The states  $|W\rangle$  and  $|\bar{W}\rangle$  are particular representatives of the symmetric Dicke states, which are defined as

$$|D_n^{(e)}\rangle = \binom{n}{e}^{-1/2} \sum_i \mathcal{P}_i(|H^{\otimes(n-e)}\rangle \otimes |V^{\otimes e}\rangle), \quad (7)$$

where  $|H/V\rangle$  denotes horizontal (vertical) polarization and  $\mathcal{P}_i$  all distinct permutations, and with the three photon states  $|W\rangle = |D_3^{(1)}\rangle$  and  $|\bar{W}\rangle = |D_3^{(2)}\rangle$ . We observed four- and six-photon Dicke states using a pulsed collinear type II spontaneous parametric down conversion source together with a linear optical setup (see Fig. 1) [30,31]. The  $|D_n^{(e)}\rangle$  states were observed upon detection of one photon in each of the four or six spatial modes, respectively. We characterized the state  $|D_4^{(2)}\rangle$  by means of quantum state tomography, i.e., a polarization analysis in each mode, collecting for each setting 26 minutes of data at a rate of 70 events per minute. The fidelity of the experimental state  $|D_4^{(2)}\rangle^{\text{exp}}$  was directly determined from the observed frequencies together with Gaussian error propagation as  $0.920 \pm 0.005$ , which due to the high number of detected events [16] is compatible with the value  $0.917 \pm 0.002$  as obtained from a maximum likelihood (ML) reconstruction and nonparametric bootstrapping [14,20]. The high quality achieved here allowed a precise study of the respective states. The fidelities of the observed three qubit states with respect to their target states are  $0.939 \pm 0.011$  for  $|W\rangle^{\text{exp}}$ ,  $0.919 \pm 0.010$  for  $|\bar{W}\rangle^{\text{exp}}$ , and  $0.961 \pm 0.003$  for  $\rho_W^{nc,\text{exp}}$ . Analogously, starting with a six-photon Dicke state  $|D_6^{(3)}\rangle$

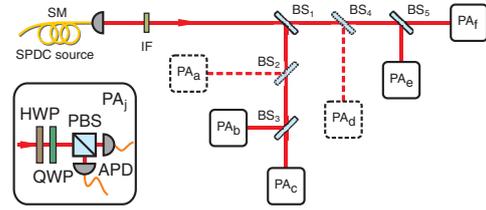


FIG. 1 (color online). Schematic of the linear optical setup used to observe symmetric Dicke states from which states with vanishing 3- and 5-partite correlations can be obtained. The photons are created by means of a cavity enhanced pulsed collinear type II spontaneous parametric down conversion source pumped at 390 nm [31]. Distributing the photons symmetrically into six modes by five beam splitters (BS) enables the observation of the state  $|D_6^{(3)}\rangle$ . Removing beam splitters BS<sub>2</sub> and BS<sub>4</sub> reduces the number of modes to four and thus the state  $|D_4^{(2)}\rangle$  is obtained. State analysis is enabled by sets of half wave (HWP) and quarter-wave plates (QWP) together with polarizing beam splitters (PBS) in each mode. The photons are measured by fiber-coupled single photon counting modules connected to a coincidence logic [30].

[32], we could also analyze the properties of the five photon state  $\rho_{D_5^{(2)}}^{nc,\text{exp}}$ . The five-qubit fidelity of  $\rho_{D_5^{(2)}}^{nc,\text{exp}}$  is determined via a ML reconstruction from fivefold coincidences to be  $0.911 \pm 0.004$  (for the detailed characterization see Supplemental Material [14]).

For the experimental analysis of the states, we start by determining  $T_{zzz}$  for the three states  $|W\rangle^{\text{exp}}$ ,  $|\bar{W}\rangle^{\text{exp}}$ , and  $\rho_W^{nc,\text{exp}}$ . As the first two have complementary structure of detection probabilities (with  $T_{zzz} = -0.914 \pm 0.034$  and  $T_{zzz} = 0.904 \pm 0.034$ , respectively), weighted mixing of these states leads to  $\rho_W^{nc,\text{exp}}$  with  $T_{zzz} = 0.022 \pm 0.023$ , i.e., a correlation value compatible with 0 (see Supplemental Material [14]). Figure 2 presents experimental data for all possible tripartite correlations of the observed states. Assuming a normal distribution centered at zero with a standard deviation given by our experimental errors, the observed correlations have a  $p$  value of 0.44 for the Anderson-Darling test, which shows that indeed one can adhere to the hypothesis of vanishing full correlations. Similarly, the five qubit state  $\rho_{D_5^{(2)}}^{nc,\text{exp}}$  exhibits strongly suppressed, almost vanishing correlations. For details on the five qubit state, please see Supplemental Material [14].

We want to emphasize that the vanishing tripartite correlations of  $\rho_W^{nc,\text{exp}}$  are no artifact of measuring in the Pauli bases. In fact, all states obtained via local unitary transformations do not exhibit any  $n$ -partite correlations. To illustrate this property, we considered correlation measurements in non-standard bases. As an example, we chose measurements in the  $zy$  plane  $\sigma_\theta = \cos\theta\sigma_z + \sin\theta\sigma_y$  with  $\theta \in [0, 2\pi]$  ( $\sigma_\phi = \cos\phi\sigma_y + \sin\phi\sigma_z$  with  $\phi \in [0, 2\pi]$ ) for the first (second) qubit resulting in the correlations  $T_{\theta_j z_j} = \text{Tr}(\rho\sigma_\theta \otimes \sigma_{j_2} \otimes \sigma_{j_3}) (T_{j_1\phi_j})$ . Indeed, as shown in

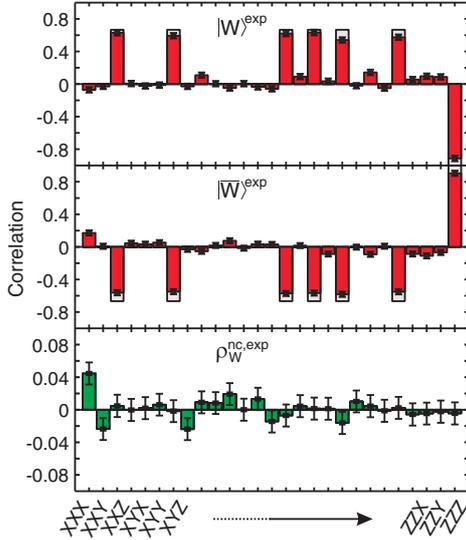


FIG. 2 (color online). Experimental tripartite correlations (red) for  $|W\rangle^{\text{exp}}$ ,  $|\bar{W}\rangle^{\text{exp}}$ , and (green)  $\rho_W^{\text{nc,exp}}$  in comparison to the theoretically expected values (gray). Note that the correlations of the state  $\rho_W^{\text{nc,exp}}$  are magnified by a factor of 10. The plot presents measured values of  $T_{j_1 j_2 j_3}$  for the observables listed below the plot. Obviously, the states  $|W\rangle^{\text{exp}}$  and  $|\bar{W}\rangle^{\text{exp}}$  have opposite tripartite correlations canceling each other when they are mixed.

Fig. 3,  $T_{\theta j_2 j_3}$  ( $T_{j_1 \phi j_3}$ ) vanishes independently of the choice of  $\theta$  ( $\phi$ ). In contrast, the bipartite correlations  $T_{\theta z_0}$  ( $T_{y \phi 0}$ ) between qubit 1 and 2 do not vanish at all and clearly depend on  $\theta$  ( $\phi$ ). By means of those even number correlations, one is still able to infer the result of another party from ones own result with probability  $2/3 > 1/2$ . For example, the values of  $T_{z z_0} = -1/3$  ( $T_{z_0 z} = -1/3$ ) indicate that knowing, e.g., result “0” for the first qubit, we can infer that the result will be “1” with  $p = 2/3$  on the second (third) qubit, etc.

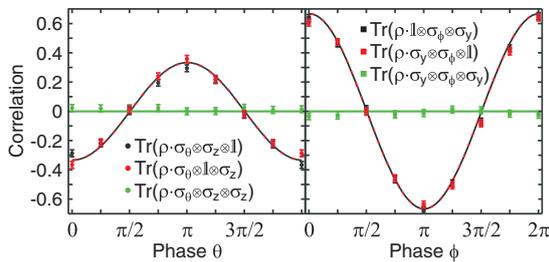


FIG. 3 (color online). Vanishing tripartite correlations for arbitrary measurements and non-vanishing bipartite correlations. Observable  $\sigma_\theta$  ( $\sigma_\phi$ ) was measured on the first (second) qubit and  $\sigma_z$  ( $\sigma_y$ ) measurements were performed on both other qubits (green curves) or one of them (red and black curves). The solid lines show the theoretically expected curves.

Although the three qubits are not tripartite correlated, the bipartite correlations shown above give rise to genuine tripartite entanglement. This can be tested for the experimental states employing Eq. (6). We observe

$$(T, T_W^{\text{nc,exp}}) = 3.858 \pm 0.079 > 3.33\bar{3},$$

$$(T, T_{D_5^{(2)}}^{\text{nc,exp}}) = 13.663 \pm 0.340 > 12.8,$$

both above the respective biseparable bound of  $10/3 = 3.33\bar{3}$  (12.8) by more than 6.6 (2.4) standard deviations, proving that in spite of vanishing full correlations the states are genuinely tripartite (five-partite) entangled [14].

The observed five-photon state has one more remarkable property [13]. For this state, every correlation between a fixed number of observers, i.e., bipartite correlations, tripartite correlations, etc. admits description with an explicit local hidden-variable model [8]. However, some of the models are different and thus cannot be combined in a single one. Using linear programming to find joint probability distributions reproducing quantum predictions [12], we obtain an optimal Bell inequality using only two- and four-partite correlations [13]. From the observed data, we evaluate the Bell parameter to be  $\mathcal{B} = 6.358 \pm 0.149$  which violates the local realistic bound of 6 by 2.4 standard deviations [33]. This violation confirms the nonclassicality [14] of this no-correlation state and also offers its applicability for quantum communication complexity tasks. Contrary to previous schemes, here, the communication problem can be solved in every instance already by only a subset of the communicating parties [35].

*Conclusions.*—We introduced a systematic way to define and to experimentally observe mixed multipartite states with no  $n$ -partite correlations for odd  $n$ , as measured by standard correlation functions. For the first time, we experimentally observed a state which allowed the violation of a Bell inequality without full correlations, thereby proving both the nonclassicality of no-correlation states as well as their applicability for quantum communication protocols. The remarkable properties of these states prompt intriguing questions. For example, what might be the dimensionality of these states or their respective subspaces, or whether we can even extend the subspace of states and antistates which give genuinely entangled no-correlation states? Moreover, can no-correlation states be used for quantum protocols beyond communication complexity, and, of course, whether these remarkable features can be cast into rigorous and easily calculable measures of genuine correlations satisfying natural postulates [11]?

We thank the EU-BMBF Project No. QUASAR and the EU Project No. QWAD and No. QOLAPS for supporting this work. T.P. acknowledges support by the National Foundation, the Ministry of Education of Singapore Grant No. RG98/13, Start-up Grant of the

Nanyang Technological University, and NCN Grant No. 2012/05/E/ST2/02352. C. S. and L. K. thank the Elite Network of Bavaria for support (Ph.D. Programs QCCC and ExQM).

- \*tomasz.paterek@ntu.edu.sg
- [1] A. Einstein, B. Podolsky, and N. Rosen, *Phys. Rev.* **47**, 777 (1935).
- [2] A. K. Ekert, *Phys. Rev. Lett.* **67**, 661 (1991).
- [3] C. H. Bennett, G. Brassard, C. Crépeau, R. Jozsa, A. Peres, and W. K. Wootters, *Phys. Rev. Lett.* **70**, 1895 (1993).
- [4] K. C. Chua, V. Chandran, U. R. Acharya, and C. M. Lim, *Medical Engineering and Physics* **32**, 679 (2010); J. M. Mendel, *Proc. IEEE* **79**, 278 (1991).
- [5] G. Tóth and O. Gühne, *Phys. Rev. A* **72**, 022340 (2005); B. Lücke, J. Peise, G. Vitagliano, J. Arlt, L. Santos, G. Tóth, and C. Klempt, *Phys. Rev. Lett.* **112**, 155304 (2014); J. I. de Vicente and M. Huber, *Phys. Rev. A* **84**, 062306 (2011); B. Jungnitsch, T. Moroder, and O. Gühne, *Phys. Rev. Lett.* **106**, 190502 (2011).
- [6] P. Badziąg, Č. Brukner, W. Laskowski, T. Paterek, and M. Żukowski, *Phys. Rev. Lett.* **100**, 140403 (2008); W. Laskowski, M. Markiewicz, T. Paterek, and M. Żukowski, *Phys. Rev. A* **84**, 062305 (2011).
- [7] J. F. Clauser, M. A. Horne, A. Shimony, and R. A. Holt, *Phys. Rev. Lett.* **23**, 880 (1969); N. D. Mermin, *Phys. Rev. D* **22**, 356 (1980); H. Weinfurter and M. Żukowski, *Phys. Rev. A* **64**, 010102 (2001); R. F. Werner and M. M. Wolf, *Phys. Rev. A* **64**, 032112 (2001); K. Nagata, W. Laskowski, and T. Paterek, *Phys. Rev. A* **74**, 062109 (2006); W. Laskowski, T. Paterek, M. Żukowski, and Č. Brukner, *Phys. Rev. Lett.* **93**, 200401 (2004); M. Ardehali, *Phys. Rev. A* **46**, 5375 (1992); A. V. Belinskii and D. N. Klyshko, *Phys. Usp.* **36**, 653 (1993); D. Collins and N. Gisin, *J. Phys. A* **37**, 1775 (2004); T. Vértesi, *Phys. Rev. A* **78**, 032112 (2008).
- [8] M. Żukowski and Č. Brukner, *Phys. Rev. Lett.* **88**, 210401 (2002).
- [9] D. Kaszlikowski, A. Sen De, U. Sen, V. Vedral, and A. Winter, *Phys. Rev. Lett.* **101**, 070502 (2008).
- [10] Z. Walczak, *Phys. Lett. A* **374**, 3999 (2010); Z. Walczak, *Phys. Rev. Lett.* **104**, 068901 (2010); D. Kaszlikowski, A. Sen De, U. Sen, V. Vedral, and A. Winter, *Phys. Rev. Lett.* **104**, 068902 (2010); K. Modi, A. Brodutch, H. Cable, T. Paterek, and V. Vedral, *Rev. Mod. Phys.* **84**, 1655 (2012).
- [11] C. H. Bennett, A. Grudka, M. Horodecki, P. Horodecki, and R. Horodecki, *Phys. Rev. A* **83**, 012312 (2011).
- [12] J. Gruca, W. Laskowski, M. Żukowski, N. Kiesel, W. Wieczorek, C. Schmid, and H. Weinfurter, *Phys. Rev. A* **82**, 012118 (2010).
- [13] W. Laskowski, M. Markiewicz, T. Paterek, and M. Wieśniak, *Phys. Rev. A* **86**, 032105 (2012).
- [14] See Supplemental Material at <http://link.aps.org/supplemental/10.1103/PhysRevLett.114.180501> for additional information on physical meaning of the correlation functions, details about experimentally realized states, derivations of entanglement criteria used, and correlation content of the no-correlation states, which includes Refs. [9,13,15–23].
- [15] G. L. Giorgi, B. Bellomo, F. Galve, and R. Zambrini, *Phys. Rev. Lett.* **107**, 190501 (2011).
- [16] C. Schwemmer, L. Knips, D. Richart, H. Weinfurter, T. Moroder, M. Kleinmann, and O. Gühne, *Phys. Rev. Lett.* **114**, 080403 (2015).
- [17] G. Tóth, *J. Opt. Soc. Am. B*, **24**, 275 (2007).
- [18] C. Schwemmer, G. Tóth, A. Niggebaum, T. Moroder, D. Gross, O. Gühne, and H. Weinfurter, *Phys. Rev. Lett.* **113**, 040503 (2014).
- [19] R. S. Bennink, Y. Liu, D. D. Earl, and W. P. Grice, *Phys. Rev. A* **74**, 023802 (2006); P. Trojek, Ph.D. thesis, Ludwig-Maximilians-Universität München, 2007.
- [20] B. Efron and R. J. Tibshirani, *An Introduction to the Bootstrap* (Chapman & Hall, London, 1994).
- [21] N. Kiesel, Ph.D. thesis, Ludwig-Maximilians-Universität München, 2007.
- [22] D. F. V. James, P. G. Kwiat, W. J. Munro, and A. G. White, *Phys. Rev. A* **64**, 052312 (2001).
- [23] J. Larsson, *J. Phys. A* **47**, 424003 (2014).
- [24] V. Bužek, M. Hillery, and R. F. Werner, *J. Mod. Opt.* **47**, 211 (2000).
- [25] F. de Martini, V. Bužek, F. Sciarrino, and C. Sias, *Nature (London)* **419**, 815 (2002); J. Bang, S.-W. Lee, H. Jeong, and J. Lee, *Phys. Rev. A* **86**, 062317 (2012).
- [26] D. L. Zhou, B. Zeng, Z. Xu, and L. You, *Phys. Rev. A* **74**, 052110 (2006).
- [27] L. Zhao, X. Hu, R.-H. Yue, and H. Fan, *Quantum Inf. Process.* **12**, 2371 (2013).
- [28] M. Eibl, N. Kiesel, M. Bourennane, C. Kurtsiefer, and H. Weinfurter, *Phys. Rev. Lett.* **92**, 077901 (2004).
- [29] T. Yamamoto, K. Tamaki, M. Koashi, and N. Imoto, *Phys. Rev. A* **66**, 064301 (2002); W. Wieczorek, N. Kiesel, C. Schmid, and H. Weinfurter, *Phys. Rev. A* **79**, 022311 (2009).
- [30] N. Kiesel, C. Schmid, G. Tóth, E. Solano, and H. Weinfurter, *Phys. Rev. Lett.* **98**, 063604 (2007); G. Tóth, W. Wieczorek, D. Gross, R. Krischek, C. Schwemmer, and H. Weinfurter, *Phys. Rev. Lett.* **105**, 250403 (2010); R. Krischek, C. Schwemmer, W. Wieczorek, H. Weinfurter, P. Hyllus, L. Pezzé, and A. Smerzi, *Phys. Rev. Lett.* **107**, 080504 (2011).
- [31] R. Krischek, W. Wieczorek, A. Ozawa, N. Kiesel, P. Michelberger, T. Udem, and H. Weinfurter, *Nat. Photonics* **4**, 170 (2010).
- [32] W. Wieczorek, R. Krischek, N. Kiesel, P. Michelberger, G. Tóth, and H. Weinfurter, *Phys. Rev. Lett.* **103**, 020504 (2009).
- [33] Examples of Bell inequalities involving lower order correlations can be found in [34], however, none of them is violated by our state  $\rho_{D_2}^{nc}$ .
- [34] M. Wieśniak, M. Nawaróg, and M. Żukowski, *Phys. Rev. A* **86**, 042339 (2012); J. Tura, R. Augusiak, A. B. Sainz, T. Vértesi, M. Lewenstein, and A. Acín, *Science* **344**, 1256 (2014); J. Tura, A. B. Sainz, T. Vértesi, A. Acín, M. Lewenstein, and R. Augusiak, *J. Phys. A* **47**, 424024 (2014).
- [35] Č. Brukner, M. Żukowski, J.-W. Pan, and A. Zeilinger, *Phys. Rev. Lett.* **92**, 127901 (2004); P. Trojek, C. Schmid, M. Bourennane, Č. Brukner, M. Żukowski, and H. Weinfurter, *Phys. Rev. A* **72**, 050305 (2005); M. Wieśniak, arXiv:1212.2388.

**Genuine  $n$ -partite entanglement without  $n$ -partite correlations  
SUPPLEMENTAL MATERIAL**

Christian Schwemmer,<sup>1,2</sup> Lukas Knips,<sup>1,2</sup> Minh Cong Tran,<sup>3</sup> Anna de Rosier,<sup>4</sup>  
Wiesław Laskowski,<sup>4</sup> Tomasz Paterek,<sup>3,5,6,\*</sup> and Harald Weinfurter<sup>1,2</sup>

<sup>1</sup>*Max-Planck-Institut für Quantenoptik, Hans-Kopfermann-Straße 1, 85748 Garching, Germany*

<sup>2</sup>*Department für Physik, Ludwig-Maximilians-Universität, 80797 München, Germany*

<sup>3</sup>*School of Physical and Mathematical Sciences, Nanyang Technological University, 637371 Singapore*

<sup>4</sup>*Institute of Theoretical Physics and Astrophysics,*

*University of Gdańsk, PL-80-952 Gdańsk, Poland*

<sup>5</sup>*Centre for Quantum Technologies, National University of Singapore, 117543 Singapore*

<sup>6</sup>*MajuLab, CNRS-UNS-NUS-NTU International Joint Research Unit, UMI 3654, Singapore*

### I. PHYSICAL MEANING OF CORRELATION FUNCTIONS

Correlations for two particles are often seen as a measure of predictability of local results when knowing the other result. Yet, this simple statement has to be used carefully. A non-vanishing  $n$ -partite correlation function indicates that we can make an educated guess of the  $n$ th result from the product of the other  $n - 1$  results. The converse statement does not hold and we provide an example of a state with vanishing correlation functions where the inference is still possible.

Let us denote by  $r_j = \pm 1$  the result of the  $j$ th observer. We assume that  $n - 1$  parties cannot infer from the product of their outcomes,  $r_1 \dots r_{n-1}$ , the result of the last observer,  $r_n$ , i.e., the following conditional probabilities hold:

$$P(r_n | r_1 \dots r_{n-1}) = \frac{1}{2}. \quad (1)$$

We show that this implies that the corresponding correlation function,  $T_{j_1 \dots j_n}$ , vanishes. The correlation function is defined as expectation value of the product of all local outcomes

$$T_{j_1 \dots j_n} = \langle r_1 \dots r_n \rangle = P(r_1 \dots r_n = 1) - P(r_1 \dots r_n = -1). \quad (2)$$

Using Bayes' rule

$$P(r_1 \dots r_n = \pm 1) = \sum_{r=\pm 1} P(r_n = \pm r | r_1 \dots r_{n-1} = r) P(r_1 \dots r_{n-1} = r). \quad (3)$$

According to assumption (1) we have  $P(r_n = \pm r | r_1 \dots r_{n-1} = r) = \frac{1}{2}$ , giving  $P(r_1 \dots r_n = \pm 1) = \frac{1}{2}$  and  $T_{j_1 \dots j_n} = 0$ .

As an example of a state with vanishing correlation functions yet allowing to make an educated guess of the result, let us consider the two-qubit mixed state

$$\frac{1}{2} |00\rangle \langle 00| + \frac{1}{4} |01\rangle \langle 01| + \frac{1}{4} |10\rangle \langle 10|, \quad (4)$$

where  $|0\rangle$  and  $|1\rangle$  are the eigenstates of the Pauli operator  $\sigma_z$  with eigenvalues  $+1$  and  $-1$ , respectively. All correlation functions  $T_{kl}$ , with  $k, l = x, y, z$ , of this state vanish. Yet, whenever Alice (Bob) observes outcome  $-1$  in the  $\sigma_z$  measurement, she (he) is sure the distant outcome is  $+1$ , i.e.,  $P(r_2 = +1 | r_1 = -1) = 1$ . Similar examples exist for multiple qubits, but we note that the states  $\rho_\phi^{nc}$  of the main text are an equal mixture of a state and its anti-state. In this case, the vanishing  $n$ -party correlations lead to the impossibility of inferring the  $n$ -th result.

### II. CRITERION FOR GENUINE MULTIPARTITE ENTANGLEMENT

To evaluate entanglement we use the following criterion (see main text) where,  $T^{exp} = T$ , i.e., assuming the ideal experiment producing the required state described by the correlation tensor  $T$ :

$$\max_{T^{bi-prod}} (T, T^{bi-prod}) < (T, T). \quad (5)$$

---

\* tomasz.paterek@ntu.edu.sg

The maximization is performed over all bi-product states keeping in mind also all possible bipartitions. The inner product between two correlation tensors of three qubit states is defined as

$$(V, W) \equiv \sum_{\mu, \nu, \eta=0}^3 V_{\mu\nu\eta} W_{\mu\nu\eta}. \quad (6)$$

#### A. Tripartite entanglement

To keep the statement as general as possible, we prove that all states  $\rho_\phi^{nc} = \frac{1}{2} |\phi\rangle \langle\phi| + \frac{1}{2} |\bar{\phi}\rangle \langle\bar{\phi}|$  with

$$|\phi\rangle = \sin\beta \cos\alpha |001\rangle + \sin\beta \sin\alpha |010\rangle + \cos\beta |100\rangle, \quad (7)$$

$$|\bar{\phi}\rangle = \sin\beta \cos\alpha |110\rangle + \sin\beta \sin\alpha |101\rangle + \cos\beta |011\rangle, \quad (8)$$

are genuinely tripartite entangled as soon as  $|\phi\rangle$  is genuinely tripartite entangled.

First, note that  $|\phi\rangle$  is a bi-product state if at least one amplitude vanishes, i.e., if either

1.  $\beta = 0$  (full product state),
2.  $\beta = \frac{\pi}{2}$  and  $\alpha = 0$  (full product state),
3.  $\beta = \frac{\pi}{2}$  and  $\alpha = \frac{\pi}{2}$  (full product state),
4.  $\beta = \frac{\pi}{2}$  and  $\alpha \in (0, \frac{\pi}{2})$  (bi-product  $A|BC$ ),
5.  $\alpha = 0$  and  $\beta \in (0, \frac{\pi}{2})$  (bi-product  $B|AC$ ),
6.  $\alpha = \frac{\pi}{2}$  and  $\beta \in (0, \frac{\pi}{2})$  (bi-product  $C|AB$ ).

The correlation tensor of the state  $\rho_\phi^{nc}$  contains only bipartite correlations:

$$\begin{aligned} T_{xx0} &= T_{yy0} = \sin(2\beta) \sin(\alpha), \\ T_{x0x} &= T_{y0y} = \sin(2\beta) \cos(\alpha), \\ T_{0xx} &= T_{0yy} = \sin^2(\beta) \sin(2\alpha), \\ T_{zz0} &= \cos(2\alpha) \sin^2(\beta) - \cos^2(\beta), \\ T_{z0z} &= -\cos(2\alpha) \sin^2(\beta) - \cos^2(\beta), \\ T_{0zz} &= \cos(2\beta), \end{aligned} \quad (9)$$

and  $T_{000} = 1$ . Using these expressions, the right-hand side of the entanglement criterion is

$$R = (T, T) = 4. \quad (10)$$

To find the maximum of the left-hand side, we shall follow a few estimations. Consider first the bi-product state in a fixed bipartition, say  $AB|C$ , i.e., of the form  $|\chi\rangle_{AB} \otimes |c\rangle$ , where  $|\chi\rangle_{AB} = \cos(\theta) |00\rangle + \sin(\theta) |11\rangle$ , when written in the Schmidt basis. Let us denote the correlation tensor of  $|\chi\rangle_{AB}$  with  $P$  and its local Bloch vectors by  $\vec{a}$  and  $\vec{b}$ . We therefore have:

$$L = 1 + T_{xx0}(P_{xx} + P_{yy}) + T_{zz0}P_{zz} + T_{x0x}(a_x c_x + a_y c_y) + T_{z0z}a_z c_z + T_{0xx}(b_x c_x + b_y c_y) + T_{0zz}b_z c_z. \quad (11)$$

By optimizing over the states of  $|c\rangle$  we get the following upper bounds:

$$T_{x0x}(a_x c_x + a_y c_y) + T_{z0z}a_z c_z \leq \sqrt{T_{x0x}^2(a_x^2 + a_y^2) + T_{z0z}^2 a_z^2}, \quad (12)$$

and

$$T_{0xx}(b_x c_x + b_y c_y) + T_{0zz}b_z c_z \leq \sqrt{T_{0xx}^2(b_x^2 + b_y^2) + T_{0zz}^2 b_z^2}. \quad (13)$$

The Schmidt decomposition implies for local Bloch vectors:

$$a_x^2 + a_y^2 + a_z^2 = b_x^2 + b_y^2 + b_z^2 = \cos^2(2\theta), \quad (14)$$

and therefore

$$\vec{a} = \cos(2\theta)\vec{n}, \quad \vec{b} = \cos(2\theta)\vec{m}, \quad (15)$$

where  $\vec{n}$  and  $\vec{m}$  are normalized vectors with directions along the local Bloch vectors. This gives the bound

$$\begin{aligned} & \sqrt{T_{x0x}^2(a_x^2 + a_y^2) + T_{z0z}^2 a_z^2} + \sqrt{T_{0xx}^2(b_x^2 + b_y^2) + T_{0zz}^2 b_z^2} \\ &= \cos(2\theta) \left( \sqrt{T_{x0x}^2(n_x^2 + n_y^2) + T_{z0z}^2 n_z^2} + \sqrt{T_{0xx}^2(m_x^2 + m_y^2) + T_{0zz}^2 m_z^2} \right) \\ &\leq \cos(2\theta) (\max(|T_{x0x}|, |T_{z0z}|) + \max(|T_{0xx}|, |T_{0zz}|)), \end{aligned} \quad (16)$$

where the maxima follow from convexity of squared components of a normalized vector.

Now let us focus on the terms depending on the correlations of  $|\chi\rangle_{AB}$ . In order to maximize (11), the Schmidt basis of  $|\chi\rangle_{AB}$  has to be either  $x$ ,  $y$ , or  $z$  as otherwise off-diagonal elements of  $P$  emerge leading to smaller values entering (11). For the diagonal correlation tensor we have  $|P_{xx}| = \sin(2\theta)$ ,  $|P_{yy}| = \sin(2\theta)$ , and  $P_{zz} = 1$ , and with indices permuted. Therefore, there are three cases to be considered in order to optimize  $T_{xx0}(P_{xx} + P_{yy}) + T_{zz0}P_{zz}$ :

- (i)  $|P_{xx}| = 1$  and  $|P_{yy}| = |P_{zz}| = \sin(2\theta)$  with their signs matching those of  $T_{xx0}$  and  $T_{zz0}$  respectively,
- (ii)  $|P_{zz}| = 1$  and  $P_{xx} = P_{yy} = \sin(2\theta)$ ,
- (iii)  $|P_{zz}| = 1$  and  $P_{xx} = -P_{yy} = \sin(2\theta)$ .

Each of these cases leads to an upper bound on  $L$ . For example, for the first case we find

$$\begin{aligned} L_{(i)} &= 1 + |T_{xx0}| + \sin(2\theta)(|T_{xx0}| + |T_{zz0}|) + \cos(2\theta)(\max(|T_{x0x}|, |T_{z0z}|) + \max(|T_{0xx}|, |T_{0zz}|)) \\ &\leq 1 + |T_{xx0}| + \sqrt{(|T_{xx0}| + |T_{z0z}|)^2 + (\max(|T_{x0x}|, |T_{z0z}|) + \max(|T_{0xx}|, |T_{0zz}|))^2}, \end{aligned} \quad (17)$$

where in the last step we optimized over  $\theta$ . The same procedure applied to the other two cases gives:

$$L_{(ii)} \leq 1 + |T_{zz0}| + \sqrt{4T_{xx0}^2 + (\max(|T_{x0x}|, |T_{z0z}|) + \max(|T_{0xx}|, |T_{0zz}|))^2}, \quad (18)$$

$$L_{(iii)} \leq 1 + |T_{zz0}| + \max(|T_{x0x}|, |T_{z0z}|) + \max(|T_{0xx}|, |T_{0zz}|). \quad (19)$$

If instead of the bipartition  $AB|C$  another one was chosen, the bounds obtained are given by those above with the indices correspondingly permuted. Since there are three possible bipartitions, altogether we have nine bounds out of which we should finally choose the maximum as the actual upper bound on the left-hand side.

#### Numerical derivation of bounds

A first approach is to numerically evaluate Eqs. (17)-(19). Fig. 1 shows that only for states  $|\phi\rangle$  that are bi-product the left-hand side reaches  $L = 4$ .

For the  $W$  state we thus obtain  $\max L = 10/3$  which is achieved by the bi-product state  $(\cos\theta|++\rangle - \sin\theta|--\rangle) \otimes |+\rangle$ , where  $|\pm\rangle = \frac{1}{\sqrt{2}}(|0\rangle \pm |1\rangle)$  and  $\tan(2\theta) = 3/4$  in order to optimize case (i) which is the best for the  $W$  state. This bound is used in the main text.

#### Analytic argument

The last step of the proof, showing that only bi-separable states can achieve the bound of 4 in our criterion, involved numerical optimization (Fig. 1). One may complain that due to finite numerical precision there might be genuinely tripartite entangled states for values of  $\alpha$  or  $\beta$  close to 0 and  $\pi/2$  that already achieve the bound of 4. Here, we give a simple analytical argument showing that  $\rho_\phi^{nc}$  is genuinely tripartite entangled if and only if  $|\phi\rangle$  is so.

We first follow the idea of Ref. [9] and note that a mixed state  $\rho_\phi^{nc}$  can only be bi-separable if there are bi-product pure states in its support. The support of  $\rho_\phi^{nc}$  is spanned by  $|\phi\rangle$  and  $|\bar{\phi}\rangle$ , i.e.,  $\rho_\phi^{nc}$  does not have any overlap with the orthogonal subspace  $\mathbb{1} - |\phi\rangle\langle\phi| - |\bar{\phi}\rangle\langle\bar{\phi}|$ . Accordingly any decomposition of  $\rho_\phi^{nc}$  into pure states can only use pure states of the form

$$|\Phi\rangle = a|\phi\rangle + b|\bar{\phi}\rangle. \quad (20)$$

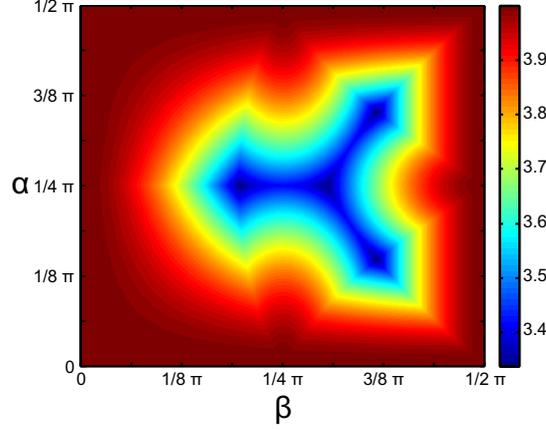


FIG. 1. Contour plot showing the maximal value of the left-hand side of our entanglement criterion for the states  $\rho_\phi^{nc}$  defined above (7). Whenever the value is below 4, i.e., the right-hand side value as given in (10), the criterion detects genuine tripartite entanglement. This shows that all the states  $\rho_\phi^{nc}$  are genuinely tripartite entangled except for those arising from bi-product states  $|\phi\rangle$ , i.e., for  $\alpha, \beta = 0$  or  $\pi/2$ . Numerical optimizations over all bi-separable states yield the same plot.

We now give a simple argument that  $|\Phi\rangle$  is bi-product, and hence  $\rho_\phi^{nc}$  is bi-separable, if and only if  $|\phi\rangle$  is bi-product. In all other infinitely many cases, the no-correlation state is genuinely tripartite entangled. Assume that  $|\Phi\rangle$  is bi-product in the partition  $AB|C$ . Accordingly, all its correlation tensor components factor across this partition. In particular,

$$T_{0xx} = W_{0x}V_x, \quad T_{0yy} = W_{0y}V_y, \quad T_{0xy} = W_{0x}V_y, \quad T_{0yx} = W_{0y}V_x \quad (21)$$

where  $W$  is the correlation tensor of the state of  $AB$  and  $V$  is the correlation tensor corresponding to the state of  $C$ . One directly verifies that for such a bi-product state we have

$$T_{0xx}T_{0yy} = T_{0xy}T_{0yx}. \quad (22)$$

Evaluating condition (22) for the states  $|\Phi\rangle$  gives the following condition on the amplitudes of  $|\phi\rangle$ :

$$\sin^2(2\alpha)\sin^4(\beta) = 0, \quad (23)$$

and indicates that at least one amplitude must be zero. Similar reasoning applies to other partitions and we conclude that  $|\Phi\rangle$  is bi-product if and only if  $|\phi\rangle$  is bi-product.

#### Alternative entanglement criterion

Alternatively we can apply a witness of genuine tripartite entanglement based on angular momentum operators [17],

$$\mathcal{W}_3 = J_x^2 + J_y^2, \quad (24)$$

where e.g.  $J_x = \frac{1}{2}(\sigma_x \otimes \mathbb{1} \otimes \mathbb{1} + \mathbb{1} \otimes \sigma_x \otimes \mathbb{1} + \mathbb{1} \otimes \mathbb{1} \otimes \sigma_x)$ . Maximization of this quantity over bi-separable states gives [17]:

$$\max_{\rho^{\text{bi-sep}}} \langle \mathcal{W}_3 \rangle = 2 + \sqrt{5}/2 \approx 3.12. \quad (25)$$

This criterion detects entanglement of the states  $|\phi\rangle$  and  $|\bar{\phi}\rangle$ , and, consequently, since it uses two-party correlations only, also of the state  $\rho_\phi^{nc}$ . However, entanglement is detected only for a range of roughly  $\alpha \in [0.590, 1.31]$  and  $\beta \in [0.333, 1.24]$ .

### B. Five-partite entanglement

In order to obtain the five-partite bound given in the main text, i.e.,  $\max_{T^{bi-prod}}(T, T^{bi-prod}) = 12.8$ , we have numerically optimized over all bi-product states keeping  $T$  as the correlation tensor of an equal mixture of Dicke states  $|D_5^{(2)}\rangle$  and  $|D_5^{(3)}\rangle$ , where

$$|D_n^{(e)}\rangle = \frac{1}{\sqrt{\binom{n}{e}}} \sum_i |\mathcal{P}_i(1, \dots, 1, 0, \dots, 0)\rangle, \quad (26)$$

with  $\mathcal{P}_i$  denoting all distinct permutations of  $e$  ones and  $n - e$  zeros.

Below, we generalize the analytical argument given above to prove genuine multipartite entanglement of arbitrary mixtures of Dicke and anti-Dicke states. The anti-Dicke state has exchanged roles of zeros and ones as compared with the Dicke state, i.e., it has  $n - e$  ones (excitations). One easily verifies that the Dicke state of  $n$  qubits with  $e$  excitations has the following bipartite correlations:

$$\begin{aligned} T_{0\dots 0xx} = T_{0\dots 0yy} &= \frac{2\binom{n-2}{e-1}}{\binom{n}{e}} = \frac{2e(n-e)}{n(n-1)}, \\ T_{0\dots 0xy} = T_{0\dots 0yx} &= 0. \end{aligned} \quad (27)$$

The correlations of an anti-Dicke state, with  $n - e$  excitations, are the same due to the symmetry  $e \leftrightarrow n - e$  of these correlations. Assume that  $n$  is odd so that (i) the Dicke and anti-Dicke states are orthogonal and (ii) the parity of the number of excitations, i.e., whether there is an even or odd number of them, is opposite in the Dicke and anti-Dicke states. For arbitrary superposition  $\alpha|D_n^{(e)}\rangle + \beta|D_n^{(n-e)}\rangle$  the correlations read:

$$T_{0\dots 0jk} = |\alpha|^2 T_{0\dots 0jk}^D + |\beta|^2 T_{0\dots 0jk}^{\bar{D}} + \alpha^* \beta \langle D_n^{(e)} | \mathbb{1} \otimes \dots \otimes \sigma_j \otimes \sigma_k | D_n^{(n-e)} \rangle + \alpha \beta^* \langle D_n^{(n-e)} | \mathbb{1} \otimes \dots \otimes \sigma_j \otimes \sigma_k | D_n^{(e)} \rangle. \quad (28)$$

Since applying  $\sigma_j \otimes \sigma_k$  with  $j, k = x, y$  to the Dicke states does not change the parity of their excitations, the last two terms vanish, and for the first two terms we have  $T_{0\dots 0jk}^D = T_{0\dots 0jk}^{\bar{D}}$ . Therefore, an arbitrary superposition of Dicke and anti-Dicke states has the same correlations as in (27) and therefore none of such superposed states is bi-product. Since the Dicke states are invariant under exchange of parties (and so are their superpositions), the same holds for other partitions. Finally, the lack of bi-product states in a subspace spanned by Dicke and anti-Dicke states implies that their mixtures are also genuinely multipartite entangled.

### III. GENUINE TRIPARTITE CORRELATIONS

While the conventional full correlation function vanishes for  $\rho_\phi^{nc}$ , this is not necessarily so for other types of correlation functions introduced recently. For a comparison we analyze the correlation content of the states of our family also according to the three measures given in Ref. [15], namely: (a) genuine tripartite correlations  $T^{(3)}(\rho_\phi^{nc})$ , (b) genuine tripartite classical correlations  $J^{(3)}(\rho_\phi^{nc})$ , and (c) genuine tripartite quantum correlations  $D^{(3)}(\rho_\phi^{nc})$ . The results are presented and discussed in Fig. 2.

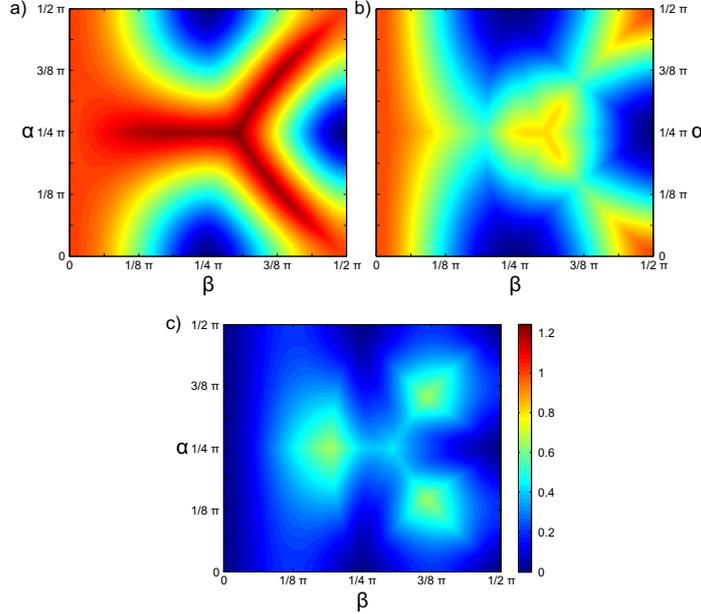


FIG. 2. Correlation content [15] of the states  $\rho_\phi^{nc} = \frac{1}{2} |\phi\rangle\langle\phi| + \frac{1}{2} |\bar{\phi}\rangle\langle\bar{\phi}|$  with the pure states given in Eq. (7). (a) *Total genuine tripartite correlations*. The genuine tripartite correlations vanish only for mixtures of bi-product states. The highest value (1.2516) is obtained for the state  $(|W\rangle\langle W| + |\bar{W}\rangle\langle\bar{W}|)/2$ . (b) *Genuine tripartite classical correlations*. The genuine classical correlations also vanish only for mixtures of bi-product states. The highest value (1.0) is observed for fully separable states. The local maximum (0.8127) is achieved by the state  $(|W\rangle\langle W| + |\bar{W}\rangle\langle\bar{W}|)/2$ . (c) *Genuine tripartite quantum correlations*. The genuine quantum correlations vanish for mixtures of bi-product states and for fully separable states. The highest values (0.6631) correspond to the mixture of the state  $\sqrt{1/6}|001\rangle + \sqrt{1/6}|010\rangle + \sqrt{2/3}|100\rangle$  with its antistate (and permutations). The state  $(|W\rangle\langle W| + |\bar{W}\rangle\langle\bar{W}|)/2$  achieves the local maximum (0.4389).

#### IV. EXPERIMENTAL THREE AND FIVE QUBIT STATES

The experimentally prepared states  $|W\rangle^{exp}$ ,  $|\bar{W}\rangle^{exp}$ ,  $\rho_W^{nc,exp}$ , and  $\rho_{D_5^{(2)}}^{nc,exp}$  were characterized by means of quantum state tomography. Their corresponding density matrices can be seen in Fig. 3 and Fig. 4. The fidelities of the observed three qubit states with respect to their target states are  $0.939 \pm 0.011$  for  $|W\rangle^{exp}$ ,  $0.919 \pm 0.010$  for  $|\bar{W}\rangle^{exp}$ , and  $0.961 \pm 0.003$  for  $\rho_W^{nc,exp}$ . Note that the value of the fidelity for the state  $\rho_W^{nc,exp}$  was obtained from a maximum likelihood (ML) reconstruction together with non-parametric bootstrapping. This value thus might be slightly incorrect due to the bias of the maximum likelihood data evaluation [16].

Fig. 4 shows the real part of the tomographically determined no-correlation state from which all further five qubit results are deduced. The five-qubit fidelity of  $\rho_{D_5^{(2)}}^{nc,exp}$  is determined via a ML reconstruction from five-fold coincidences to be  $0.911 \pm 0.004$ .

To obtain a correlation function value, e.g.,  $T_{zzz} = \text{Tr}(\rho \sigma_z \otimes \sigma_z \otimes \sigma_z)$ , we analyze the three photons in the respective set of bases (here all  $\hat{z}$ ). Fig. 5 shows the relative frequencies for observing all the possible results for such a polarization analysis. Clearly one recognizes the complementary structure of the the detection frequencies for the states  $|W\rangle^{exp}$  and  $|\bar{W}\rangle^{exp}$  which results in approximately the same magnitude of the correlations, yet with different sign. Mixing the two states, one thus obtains a vanishingly small correlation. Fig. [2] of the main text then shows the full set of correlations.

For the analysis of the five qubit no correlation state, we see from an eigen decomposition that this state indeed comprises of a mixture of two states ( $|\Theta^{(2)}\rangle^{exp}$  and  $|\Theta^{(3)}\rangle^{exp}$ ), which are in very good agreement with  $|D_5^{(2)}\rangle$  and  $|D_5^{(3)}\rangle$ . Fig. 6 (a) and (b) show all symmetrized correlations for the five-qubit states  $|\Theta^{(2)}\rangle$  and  $|\Theta^{(3)}\rangle$  and  $\rho_{D_5^{(2)}}^{nc,exp}$  with good agreement with the ideal states. Also the respective fidelity of the eigenvectors of the experimentally determined state are quite high ( $F_{|D_5^{(2)}\rangle}(|\Theta^{(2)}\rangle) = 0.978 \pm 0.012$  and  $F_{|D_5^{(3)}\rangle}(|\Theta^{(3)}\rangle^{exp}) = 0.979 \pm 0.012$ ). Equally mixing the states

$|\Theta^{(2)}\rangle^{exp}$  and  $|\Theta^{(3)}\rangle^{exp}$  indeed would result in a state with vanishingly small correlations as seen in Fig. 6 (c). However, due to asymmetry in the coupling of signal and idler states from the down conversion source [19] the correlations are still present, albeit smaller by a factor of 10 compared with  $|D_5^{(2)}\rangle$  and  $|D_5^{(3)}\rangle$ . In the main text we show that the very same state is genuinely five-party entangled.

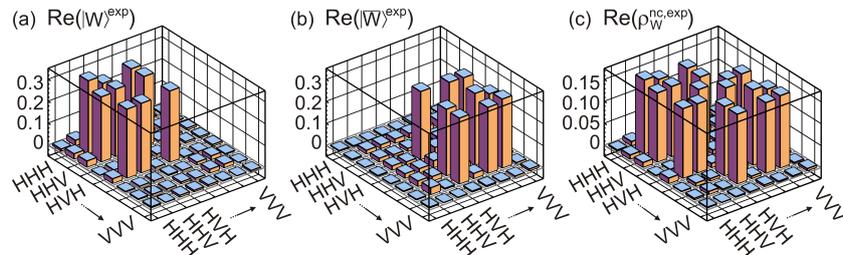


FIG. 3. Experimental three qubit states as obtained from the state  $|D_4^{(2)}\rangle^{exp}$ . (a) The state  $|W\rangle^{exp}$  is obtained by projection of the fourth qubit of  $|D_4^{(2)}\rangle^{exp}$  on  $V$ . (b) The state  $|\bar{W}\rangle^{exp}$  is prepared by projecting the fourth qubit of  $|D_4^{(2)}\rangle^{exp}$  on  $H$ . (c) When the fourth qubit of  $|D_4^{(2)}\rangle^{exp}$  is traced out, a mixture of  $|W\rangle^{exp}$  and  $|\bar{W}\rangle^{exp}$  is obtained, i.e., the state  $\rho_W^{nc,exp}$ . The corresponding fidelities with respect to their target states are  $0.939 \pm 0.011$  for  $|W\rangle^{exp}$ ,  $0.919 \pm 0.010$  for  $|\bar{W}\rangle^{exp}$ , and  $0.961 \pm 0.003$  for  $\rho_W^{nc,exp}$ .

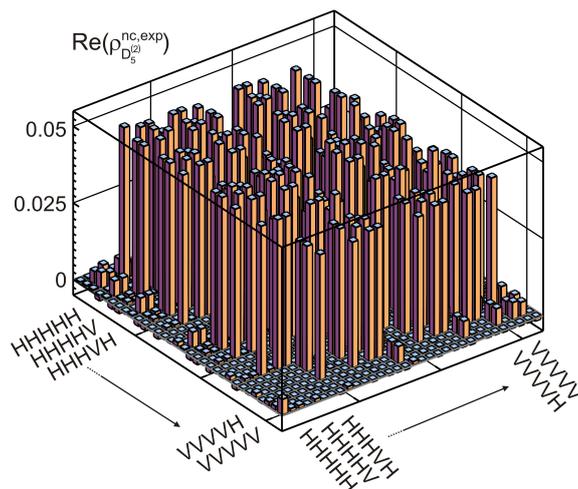


FIG. 4. Experimental state  $\rho_{D_5^{(2)}}^{nc,exp}$  determined from five-fold coincidences together with permutational invariant tomography [18]. The fidelity with respect to the target state is  $0.911 \pm 0.004$ .

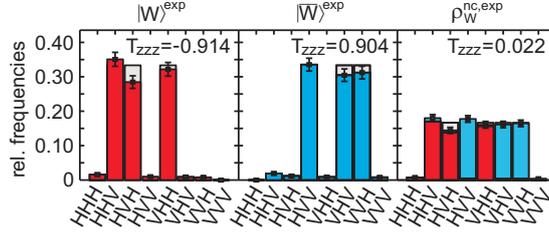


FIG. 5. (color online). Detection frequencies when observing the states  $|W\rangle^{exp}$  (red) and  $|\bar{W}\rangle^{exp}$  (blue) and  $\rho_W^{nc,exp}$  (red and blue) in the  $\sigma_z^{\otimes 3}$  basis. From these data  $T_{zzz}$  values can be calculated showing how the correlations of  $|W\rangle^{exp}$  and  $|\bar{W}\rangle^{exp}$  average to approximately 0. For comparison, the theoretically expected values are shown in gray. The correlation value  $T_{zzz}$  of the state  $\rho_W^{nc,exp}$  was determined as the weighted sum of the correlation values  $T_{zzz}$  of the states  $|W\rangle^{exp}$  and  $|\bar{W}\rangle^{exp}$ . The state  $|W\rangle^{exp}$  was observed with a slightly lower probability (0.485) than the state  $|\bar{W}\rangle^{exp}$  (0.515) leading to a value of  $T_{zzz} = 0.022$  for the state  $\rho_W^{nc,exp}$ . In contrast, in Fig. 2 of the main text the states  $|W\rangle^{exp}$  and  $|\bar{W}\rangle^{exp}$  were obtained from the state  $|D_4^{(2)}\rangle^{exp}$  by projection of the fourth qubit onto horizontal/vertical polarization, i.e., from measuring  $\sigma_z$  on the fourth qubit. There,  $\rho_W^{nc,exp}$  was obtained by tracing out the fourth qubit and hence measurements of  $\sigma_x, \sigma_y, \sigma_z$  on the fourth qubit of  $|D_4^{(2)}\rangle^{exp}$  contribute, leading to approximately three times better statistics for the state  $\rho_W^{nc,exp}$ .

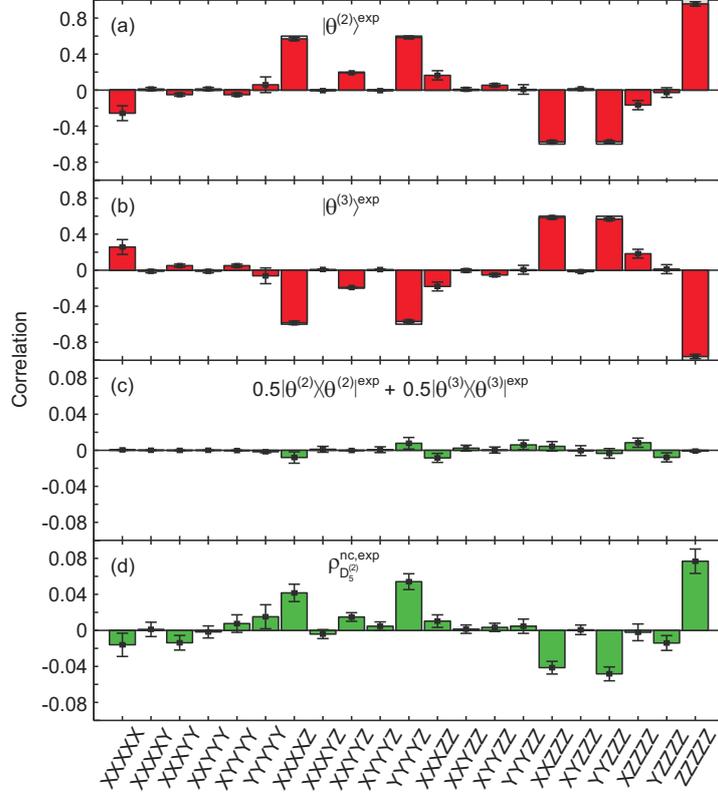


FIG. 6. Experimental five-partite symmetric correlations for the two most prominent states (a)  $|\Theta^{(2)}\rangle^{exp}$  and (b)  $|\Theta^{(3)}\rangle^{exp}$  in the eigen decomposition of the experimental density matrix  $\rho_{D_5^{(2)}}^{nc,exp}$  shown in Fig. 4. The correlations of these states are compared with the ones of the states (a)  $|D_5^{(2)}\rangle$  and (b)  $|D_5^{(3)}\rangle$ , respectively, shown in gray. The agreement between the actual and expected correlations is evident and also the fidelities of  $|\Theta^{(2)}\rangle^{exp}$  and  $|\Theta^{(3)}\rangle^{exp}$  with the respective target states are high:  $F_{|D_5^{(2)}\rangle}(|\Theta^{(2)}\rangle^{exp}) = 0.978 \pm 0.012$  and  $F_{|D_5^{(3)}\rangle}(|\Theta^{(3)}\rangle^{exp}) = 0.979 \pm 0.012$ . (c) When both states are evenly mixed, the resultant state has practically vanishing correlations. (d) Since the collection efficiencies for signal and idler photons generated via spontaneous parametric down-conversion differ slightly [19], the states  $|\Theta^{(2)}\rangle^{exp}$  and  $|\Theta^{(3)}\rangle^{exp}$  are observed with relative weights of 0.54 and 0.46 leading to largely suppressed but not entirely vanishing full correlations. Hence, the experimentally prepared state  $\rho_{D_5^{(2)}}^{nc,exp}$  is a very good approximation to a no-correlation state. Please note that the correlations shown in (c) and (d) are magnified by a factor of 10 compared with the scale of (a) and (b). The errors given in subfigures (a)-(c) were obtained by non-parametric bootstrapping [20] whereas for (d) Gaussian error propagation was used.

## V. STATISTICAL ANALYSIS

### A. Error analysis

In order to carry out  $n$ -qubit quantum state tomography, we measured in the eigenbases of all  $3^n$  combinations of local Pauli settings  $s_i$  with  $s_1 = x\dots xx$ ,  $s_2 = x\dots xy$ , ...,  $s_{3^n} = z\dots zz$ . In each setting  $s_i$  we performed projection measurements on all the  $2^n$  eigenvectors of the corresponding operators. The single measurement results are enumerated by  $r_j$  representing the binary numbers from 0 to  $2^n - 1$  in increasing order, i.e.,  $r_1 = 0\dots 00$ ,  $r_2 = 0\dots 01$ , ...,  $r_{2^n} = 1\dots 11$ . The observed counts for the outcome  $r_j$  when measuring  $s_i$  are labeled as  $c_{r_j}^{s_i}$  and the total number of counts  $N_{s_i}$  for setting  $s_i$  is given by  $N_{s_i} = \sum_{j=1}^{2^n} c_{r_j}^{s_i}$ . From these data the density matrix can be obtained as

$$\rho = \sum_{i=1}^{3^n} \sum_{j=1}^{2^n} \frac{c_{r_j}^{s_i}}{N_{s_i}} M_{r_j}^{s_i} \quad (29)$$

where the elements of the generating set of operators  $M_{r_j}^{s_i}$  are defined as  $M_{r_j}^{s_i} = \frac{1}{2^n} \bigotimes_{k=1}^n \left( \frac{\mathbf{1}}{3} + (-1)^{r_j(k)} \sigma_{s_i(k)} \right)$  [21, 22], where  $\mathbf{1}$  denotes the  $2 \times 2$  identity matrix and  $r_j(k)$  is the  $k$ -th entry in the string  $r_j$ . Then, the fidelity  $F_{|\psi\rangle}$  with respect to a pure target state  $|\psi\rangle$  can be calculated as

$$F_{|\psi\rangle} = \langle \psi | \rho | \psi \rangle = \sum_{i=1}^{3^n} \sum_{j=1}^{2^n} \frac{c_{r_j}^{s_i}}{N_{s_i}} \langle \psi | M_{r_j}^{s_i} | \psi \rangle. \quad (30)$$

For Poissonian measurement statistics, i.e.,  $\Delta c_{r_j}^{s_i} = \sqrt{c_{r_j}^{s_i}}$ , the error to the fidelity  $\Delta F_{|\psi\rangle} = \sqrt{\Delta^2 F_{|\psi\rangle}}$  can be deduced via Gaussian error propagation as  $\Delta^2 F_{|\psi\rangle} = \sum_{i=1}^{3^n} \sum_{j=1}^{2^n} \left( \frac{1}{N_{s_i}} - \frac{1}{N_{s_i}^2} \right)^2 \langle \psi | M_{r_j}^{s_i} | \psi \rangle^2 c_{r_j}^{s_i}$  which is approximately

$$\Delta^2 F_{|\psi\rangle} = \sum_{i=1}^{3^n} \Delta^2 F_{|\psi\rangle}^{s_i} = \sum_{i=1}^{3^n} \sum_{j=1}^{2^n} \frac{c_{r_j}^{s_i}}{N_{s_i}^2} \langle \psi | M_{r_j}^{s_i} | \psi \rangle^2 \quad (31)$$

for large number of counts per setting as in our experiment. As an example, in table I we give the corresponding values for  $c_{r_j}^{s_i}$  and  $|\langle \psi | M_{r_j}^{s_i} | \psi \rangle|$  for the  $2^3 = 8$  possible results of the  $zzz$  measurement of the three qubit  $|W\rangle$  state to get an impression of the size of the  $3^3 = 27$  terms in Eq. (31).

TABLE I. The values of  $c_{r_j}^{s_i}$  and  $|\langle \psi | M_{r_j}^{s_i} | \psi \rangle|$  for the measurement of the setting  $zzz$  of the experimentally observed state  $|W\rangle^{exp}$ . The first row shows all possible results  $r_j$  associated with the eigenvectors on which projection measurements are performed, labeled in binary representation. Please note that the observed counts  $c_{r_j}^{s_i}$  are not integers since the slightly differing relative detection efficiencies of the single photon counters were included. From these data we obtain for  $s_i = zzz$  a contribution for Eq. (31) of  $\Delta^2 F_{|W\rangle}^{zzz} = 2.46\text{e-}05$ .

	$r_j$	000	001	001	011	100	101	110	111
$zzz$	$ \langle \psi   M_{r_j}^{zzz}   \psi \rangle $	1.48e-01	1.48e-01	1.48e-01	1.11e-01	1.48e-01	1.11e-01	1.11e-01	7.41e-02
	counts $c_{r_j}^{zzz}$	14	309	250	8.71	283	8	7.07	0

Similarly, also the error of the  $4^3 = 64$  correlations of the given state are evaluated. For example, we obtain for the correlation value  $T_{zzz} = -0.914 \pm 0.034$ . The error for the maximum likelihood estimate was determined by non-parametric bootstrapping, for details see [20].

### B. Hypothesis testing

#### *Vanishing correlations*

After having calculated the experimental error of the  $zzz$  correlation, we find that the measurements of the remaining 26 full correlations have similar errors. We test our hypothesis of vanishing full correlations by comparing our measured

correlation values with a normal distribution with mean  $\mu = 0$  and standard deviation  $\sigma = 0.0135$ , which corresponds to the average experimental standard deviation. If our data are in agreement with this distribution, we can retain the hypothesis of vanishing full correlations.

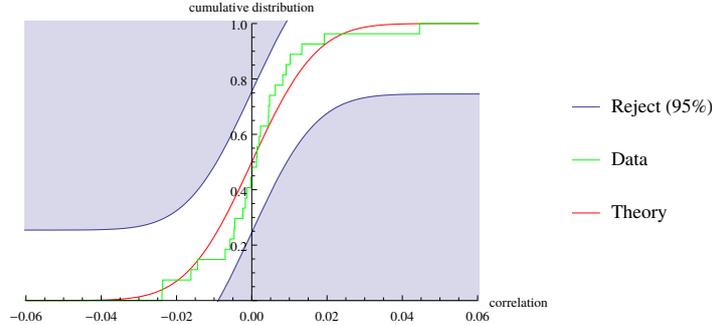


FIG. 7. The cumulative distribution of the experimentally determined correlations is compared to the cumulative distribution of the expected correlations ( $\mu = 0$ ,  $\sigma = 0.0135$ ). The shaded blue region contains points that would be sampled from the normal distribution with probability smaller than 5%. Since the empirical function lies in between the shaded regions, our hypothesis of vanishing correlations can be retained with significance level of 0.05.

To test the hypothesis

$$H_0^{(nc,3)} : \text{all full correlations of the state } \rho_W^{nc,exp} \text{ vanish,}$$

according to the Kolmogorov-Smirnov method, the cumulative distribution of the 27 measured full correlations is compared with the cumulative probability distribution of the assumed normal distribution, see Fig. 7, quantifying the hypothesis of vanishing full correlations. We can directly see that the data do not enter the region of rejection given by a significance level of 0.05. This clearly indicates that the hypothesis of normal distribution with mean  $\mu = 0$  and  $\sigma = 0.135$  cannot be rejected. While this test (Kolmogorov-Smirnov hypothesis test) is demonstrative, the Anderson-Darling test is considered to be more powerful, i.e., to decrease the probability of errors of second kind. Since the Anderson-Darling test gives a  $p$ -value of 0.44 far above a 0.05 significance level, we can retain the claim that our measured data indeed correspond to vanishing full correlations, while their scatter can be fully explained by the experimental error.

#### Testing for genuine multipartite entanglement

Furthermore, we also check our hypotheses of the main text that the tripartite and five-partite states are genuinely multipartite entangled. For that purpose, we calculate the probability that a state without genuine multipartite entanglement achieves values comparable to the measured value based on the assumption that the measurement errors are normally distributed. Let us formulate for the tripartite state the null hypothesis

$$H_0^{(3)} : \text{state } \rho_W^{nc,exp} \text{ is not genuinely tripartite entangled.}$$

To show the genuine tripartite entanglement of that state, we want to reject the null hypothesis  $H_0^{(3)}$ . In order to estimate the error of first kind, i.e., the probability that  $H_0^{(3)}$  is *true*, we calculate the probability that a state without tripartite entanglement achieves the measured value of  $(T, T_W^{nc,exp}) = 3.858$ . The calculation is based on the assumption of a normal distributed result of the indicator with mean  $\mu = \frac{10}{3}$ , i.e., the bi-separable bound, and with standard deviation given by our experimental error of  $\sigma = 0.079$ . The probability of the error of first kind is then at most

$$p = \Pr \left[ (T, T_W^{nc,exp}) \geq 3.858 \mid H_0^{(3)} \right] < \frac{1}{\sqrt{2\pi}\sigma} \int_{3.858}^{\infty} dx \exp \left( -\frac{(x - \mu)^2}{2\sigma^2} \right) = 1.55 \times 10^{-11} \ll 0.05. \quad (32)$$

Since  $p$  is far below the significance level of 0.05, our experimentally implemented state  $\rho_W^{nc}$  is genuine tripartite entangled.

Analogously, we test if the state  $\rho_{D_5^{(2)}}^{nc,exp}$  is indeed genuinely five-partite entangled. For that purpose, we formulate the null hypothesis

$$H_0^{(5)} : \text{state } \rho_{D_5^{(2)}}^{nc,exp} \text{ is not genuinely five-partite entangled.}$$

In order to test the probability that a bi-separable state can achieve  $\left(T, T_{D_5^{(2)}}^{nc,exp}\right) = 13.663$ , we now use a normal distribution centered around the bi-separable bound of  $\mu = 12.8$ . The standard deviation is chosen according to the experimental error of  $\sigma = 0.340$ , such that the probability for a false rejection of the null hypothesis  $H_0^{(5)}$  is estimated to be at most

$$p = \Pr \left[ \left( T, T_{D_5^{(2)}}^{nc,exp} \right) \geq 13.663 \middle| H_0^{(5)} \right] < \frac{1}{\sqrt{2\pi}\sigma} \int_{13.663}^{\infty} dx \exp \left( -\frac{(x-\mu)^2}{2\sigma^2} \right) = 5.6 \times 10^{-3} \ll 0.05, \quad (33)$$

clearly indicating the five-partite entanglement of our state with high significance.

#### *Bell inequality*

Finally, we test whether we can retain our claim that the five-partite state is non-classical due to its violation of the Bell inequality. In order to show the violation, we formulate the null hypothesis

$$H_0^B : \text{violation of the Bell inequality can be explained by LHV model (finite statistics loophole).}$$

For the considered Bell inequality [13]

$$\mathcal{B} = E_{\mathcal{P}(11110)} + E_{\mathcal{P}(22220)} + E_{\mathcal{P}(12220)} - E_{\mathcal{P}(21110)} - E_{\mathcal{P}(11000)} - E_{\mathcal{P}(22000)} \leq 6 \quad (34)$$

with  $\mathcal{P}$  denoting the summation over all permutations, e.g.  $E_{\mathcal{P}(11110)} = E_{11110} + E_{11101} + E_{11011} + E_{10111} + E_{01111}$ , we calculate the probability that an LHV model can achieve the measured value of  $\mathcal{B} = 6.358$ , which was estimated with a standard deviation of  $\Delta\mathcal{B} = 0.149$ . Following Ref. [23] we assume that the LHV model gives the maximal allowed expectation value of our Bell parameter, equal to  $\mu = 6$ , and that the standard deviation of a normal distribution about this mean value is equal to our experimental standard deviation  $\Delta\mathcal{B}$ . Therefore, the probability that the LHV model gives values at least as high as observed is found to be

$$p = \Pr \left[ \mathcal{B} \geq 6.358 \middle| H_0^B \right] < \frac{1}{\sqrt{2\pi}\sigma} \int_{6.358}^{\infty} dx \exp \left( -\frac{(x-\mu)^2}{2\sigma^2} \right) = 0.0083 \ll 0.05. \quad (35)$$

This small  $p$ -value clearly indicates that the null hypothesis  $H_0^B$  is to be rejected and thus the non-classicality of the no-correlation state is confirmed.

#### **C. Vanishing full correlations with arbitrary measurement directions**

The measurements presented in the main text show not only vanishing full correlations for measurements in  $x, y, z$  directions, but also for measurements of one qubit rotated in the  $yz$ -plane. Here, we show that full correlations have to vanish for arbitrary measurement directions. Since the 2-norm of the correlation tensor is invariant under local rotations, its entries vanish in all local coordinate systems if they do in one. Moreover,  $l$ -fold correlations in one set of local coordinate system only depend on  $l$ -fold correlations of another set. As an example, we explicitly show this for the case of three qubits.

$$T_{(\theta_1, \phi_1)(\theta_2, \phi_2)(\theta_3, \phi_3)} = \text{Tr}(\rho \sigma_{(\theta_1, \phi_1)} \otimes \sigma_{(\theta_2, \phi_2)} \otimes \sigma_{(\theta_3, \phi_3)}) \quad (36)$$

with

$$\sigma_{(\theta_i, \phi_i)} = \sin(\theta_i) \cos(\phi_i) \sigma_x + \sin(\theta_i) \sin(\phi_i) \sigma_y + \cos(\theta_i) \sigma_z. \quad (37)$$

Consequently,

$$\begin{aligned}
T_{(\theta_1, \phi_1) (\theta_2, \phi_2) (\theta_3, \phi_3)} &= \sin(\theta_1) \cos(\phi_1) \sin(\theta_2) \cos(\phi_2) \sin(\theta_3) \cos(\phi_3) T_{xxx} \\
&+ \sin(\theta_1) \cos(\phi_1) \sin(\theta_2) \cos(\phi_2) \sin(\theta_3) \sin(\phi_3) T_{xxy} \\
&+ \dots \\
&+ \cos(\theta_1) \cos(\theta_2) \cos(\theta_3) T_{zzz},
\end{aligned} \tag{38}$$

which has to vanish since all full correlations along Pauli directions vanish.

- 
- [9] D. Kaszlikowski, A. Sen De, U. Sen, V. Vedral, and A. Winter, Phys. Rev. Lett. **101**, 070502 (2008).
  - [13] W. Laskowski, M. Markiewicz, T. Paterek, and M. Wieśniak. Phys. Rev. A **86**, 032105 (2012).
  - [15] G. L. Giorgi, B. Bellomo, F. Galve, and R. Zambrini, Phys. Rev. Lett. **107**, 190501 (2011).
  - [16] C. Schwemmer, L. Knips, D. Richart, T. Moroder, M. Kleinmann, O. Gühne, and H. Weinfurter, Phys. Rev. Lett. **114**, 080403 (2015).
  - [17] G. Tóth, J. Opt. Soc. Am. B, **24**, 275 (2007).
  - [18] C. Schwemmer, G. Tóth, A. Niggelbaum, T. Moroder, D. Gross, O. Gühne, and H. Weinfurter, Phys. Rev. Lett. **113**, 040503 (2014).
  - [19] R. S. Bennink, Y. Liu, D. D. Earl, and W. P. Grice, Phys. Rev. A **74**, 023802 (2006); P. Trojek, Ph.D. thesis, Ludwig-Maximilians-Universität München, 2007.
  - [20] B. Efron and R. J. Tibshirani, *An introduction to the bootstrap* (Chapman & Hall, 1994).
  - [21] N. Kiesel, Ph.D. thesis, Ludwig-Maximilians-Universität München (2007).
  - [22] D. F. V. James, P. G. Kwiat, W. J. Munro, and A. G. White Phys. Rev. A **64**, 052312 (2001).
  - [23] J. Larsson, J. Phys. A: Math. Theor. **47**, 424003 (2014).

PHYSICAL REVIEW A **95**, 062331 (2017)

## Genuine $N$ -partite entanglement without $N$ -partite correlation functions

Minh Cong Tran,<sup>1,2</sup> Margherita Zuppardo,<sup>1,3</sup> Anna de Rosier,<sup>4</sup> Lukas Knips,<sup>5,6</sup> Wiesław Laskowski,<sup>4</sup>  
Tomasz Paterek,<sup>1,7</sup> and Harald Weinfurter<sup>5,6</sup>

<sup>1</sup>*School of Physical and Mathematical Sciences, Nanyang Technological University, 637371 Singapore*

<sup>2</sup>*Joint Center for Quantum Information and Computer Science, University of Maryland, College Park, Maryland 20742, USA*

<sup>3</sup>*Science Institute, University of Iceland, Dunhaga 3, IS-107 Reykjavik, Iceland*

<sup>4</sup>*Institute of Theoretical Physics and Astrophysics, Faculty of Mathematics, Physics and Informatics,  
University of Gdańsk, 80-308 Gdańsk, Poland*

<sup>5</sup>*Max-Planck-Institut für Quantenoptik, Hans-Kopfermann-Straße 1, 85748 Garching, Germany*

<sup>6</sup>*Department für Physik, Ludwig-Maximilians-Universität, 80797 München, Germany*

<sup>7</sup>*MajuLab, CNRS-UNS-NUS-NTU International Joint Research Unit, UMI 3654 Singapore, Singapore*

(Received 11 April 2017; published 26 June 2017)

A genuinely  $N$ -partite entangled state may display vanishing  $N$ -partite correlations measured for arbitrary local observables. In such states the genuine entanglement is noticeable solely in correlations between subsets of particles. A straightforward way to obtain such states for odd  $N$  is to design an “antistate” in which all correlations between an odd number of observers are exactly opposite. Evenly mixing a state with its antistate then produces a mixed state with no  $N$ -partite correlations, with many of them genuinely multipartite entangled. Intriguingly, all known examples of “entanglement without correlations” involve an *odd* number of particles. Here we further develop the idea of antistates, thereby shedding light on the different properties of even and odd particle systems. We conjecture that there is no antistate to any pure even- $N$ -party entangled state making the simple construction scheme unfeasible. However, as we prove by construction, higher-rank examples of entanglement without correlations for arbitrary even  $N$  indeed exist. These classes of states exhibit genuine entanglement and even violate an  $N$ -partite Bell inequality, clearly demonstrating the nonclassical features of these states as well as showing their applicability for quantum information processing.

DOI: [10.1103/PhysRevA.95.062331](https://doi.org/10.1103/PhysRevA.95.062331)

### I. INTRODUCTION

Quantum entanglement is present in quantum states that cannot be obtained from uncorrelated states by local operations and classical communication [1,2]. It turns out that for pure states the existence of entanglement is fully captured by  $N$ -partite correlation functions only: A pure state is entangled if and only if the sum of squared  $N$ -partite correlation functions exceeds certain bounds [3–7]. One may then wonder if similar detection methods could exist for mixed states, i.e., whether appropriate processing of only  $N$ -partite correlation functions detects entanglement in all mixed states. The states we consider here demonstrate vividly that such a universal entanglement criterion does not exist. Despite vanishing  $N$ -partite correlation functions in all possible local measurements, these states can be even genuinely  $N$ -partite entangled. As a matter of fact the genuine  $N$ -partite entanglement is due to nonvanishing correlations between less than  $N$  particles, the so-called lower-order correlations.

The first example of such a state was given in Ref. [8] and consists of an even mixture of two  $W$  states between an odd number of qubits. The two states have exactly opposite  $N$ -partite correlations such that they average out in the even mixture. More recently it was shown that any pure quantum state has an “antistate” where all correlation functions have opposite signs, but only between an *odd* number of observers [9,10]. Then, the equal mixture of a pure state of odd number of qubits and its antistate produces a mixed state with vanishing  $N$ -partite correlation functions. Many of such

“no-correlation” states are genuinely  $N$ -partite entangled and even an infinite family of such states with two continuous parameters could be constructed [10].

Here we generalize the notion of antistates and study their relations to entanglement without correlations. A number of problems were raised in Refs. [8–10] which have now been solved. In particular, we provide an analytical plausibility argument (Sec. V) and strong numerical evidence (Conjecture 1) that there is no antistate to any genuinely multipartite entangled pure state of an even number of qubits. This explains why previous relatively simple examples of entanglement without correlations could be constructed for only an odd number of qubits. Therefore, in the case of an even number of particles, this phenomenon requires mixing of at least three pure quantum states. We provide here analytical examples of rank-4 mixed states that are both genuinely  $N$ -partite entangled and have vanishing all  $N$ -partite correlation functions, for arbitrary even  $N$  (Sec. III A). Up to numerical precision also rank-3 mixed states with this property exist. Remarkably, they violate a suitably designed Bell-type inequality (Sec. III B). In order to further emphasize that entanglement without correlations is not very unusual, we extend the previous example of the infinite family with two continuous parameters to exponentially many in  $N$  continuous parameters (Sec. IV). This is achieved with the help of a simple identifier of genuine multipartite entanglement that also illustrates limits to entanglement detection with only bipartite correlation functions [11–16].

MINH CONG TRAN *et al.*PHYSICAL REVIEW A **95**, 062331 (2017)

## II. BASIC NOTIONS

### A. Genuine multipartite entanglement

A mixed quantum state of  $N$  particles is genuinely  $N$ -partite entangled if it cannot be written as

$$\rho \neq \sum_j p_j \rho_{A_j}^j \otimes \rho_{B_j}^j, \quad (1)$$

where  $A_j : B_j$  is a partition of the  $N$  particles and  $p_j$  are probabilities. Note that different terms in this convex decomposition may involve different partitions. The states  $\rho_{A_j}^j \otimes \rho_{B_j}^j$  can always be chosen pure, in which case they are called biproduct states. All our examples will exploit the fact that if the support of  $\rho$  does not contain a single biproduct state, then  $\rho$  must be genuinely  $N$ -partite entangled.

### B. Correlation functions

The correlation function is a standard statistical quantifier defined as the expectation value of a product of measurement results. Consider dichotomic observables, i.e., the measurement results are  $\pm 1$ , conducted on multiple qubits. Such observables are parametrized by unit vectors on a sphere. We denote by  $\vec{m}_n$  the vector encoding the observable of the  $n$ th party. If such observables are measured on every particle from an  $N$ -partite quantum system in state  $\rho$ , one obtains the  $N$ -partite (quantum) correlation function:

$$E(\vec{m}_1, \dots, \vec{m}_N) = \text{Tr}(\rho \vec{m}_1 \cdot \vec{\sigma} \otimes \dots \otimes \vec{m}_N \cdot \vec{\sigma}), \quad (2)$$

where  $\vec{\sigma} = (\sigma_x, \sigma_y, \sigma_z)$  is the vector of Pauli operators. We shall also write  $\sigma_x, \sigma_y, \sigma_z$  as  $\sigma_1, \sigma_2, \sigma_3$ , respectively. It is customary to introduce correlation tensor  $T$  or, respectively, its coefficients

$$T_{j_1 \dots j_N}(\rho) = \text{Tr}(\rho \sigma_{j_1} \otimes \dots \otimes \sigma_{j_N}), \quad (3)$$

for the  $N$ -partite correlation functions measured explicitly along the  $\vec{x}, \vec{y}, \vec{z}$  axes. Here  $j_n = 1, 2, 3$ . By writing  $\rho$  in the basis of tensor products of Pauli operators, one easily verifies the tensor transformation law:

$$E(\vec{m}_1, \dots, \vec{m}_N) = \sum_{j_1, \dots, j_N=1}^3 T_{j_1 \dots j_N}(\vec{m}_1)_{j_1} \dots (\vec{m}_N)_{j_N}, \quad (4)$$

where  $(\vec{m}_n)_{j_n}$  is the component of the vector  $\vec{m}_n$  along the  $j_n$ th axis. In the present context this implies that it is sufficient to ensure that  $T_{j_1 \dots j_N} = 0$  for all  $j_1, \dots, j_N = 1, 2, 3$  to guarantee that  $N$ -partite correlation functions vanish for arbitrary local measurements.

One could of course also measure subsets of all  $N$  particles, in which case the resulting correlation functions are called lower-order correlations. We will be only interested in these correlations along the  $\vec{x}, \vec{y}, \vec{z}$  axes, in which case they can be calculated as follows:

$$T_{\mu_1 \dots \mu_N}(\rho) = \text{Tr}(\rho \sigma_{\mu_1} \otimes \dots \otimes \sigma_{\mu_N}), \quad (5)$$

where index  $\mu_n = 0, 1, 2, 3$ , i.e., additionally to Pauli operators it also includes  $\sigma_0$ , the identity operator, for those parties who do not conduct measurements. For example, bipartite correlation functions between the first two observers are denoted by  $T_{j_1 j_2 0 \dots 0}$ .

### C. Antistates

Given a pure or mixed state  $\rho$ , with the  $N$ -partite correlation functions  $T_{j_1 \dots j_N}$ , we define its antistate  $\bar{\rho}$  by the requirement that all its  $N$ -partite correlation functions have opposite sign, i.e.,  $T_{j_1 \dots j_N}(\bar{\rho}) = -T_{j_1 \dots j_N}(\rho)$  for all indices  $j_n = 1, 2, 3$ . No assumptions are made about the lower-order correlation functions.

Reference [10] presented a method to build an antistate to an arbitrary pure state with an odd number  $N$  of qubits. However, this method does not apply to cases where  $N$  is even. We therefore need to use different approaches depending on the parity of  $N$ , as we will discuss in the next sections.

## III. $N$ EVEN

Let us begin with the problem of existence of antistates for an even number of qubits. We argue that most likely all genuinely  $N$ -partite entangled pure states of even  $N$  do not admit antistates. Nevertheless, this does not imply impossibility of entanglement without correlations. It just says that more than two pure states have to be present in the mixture. Indeed, we will provide such examples for every even  $N \geq 4$ .

We start with bipartite systems where one can easily exclude existence of an antistate to arbitrary pure entangled state.

*Theorem 1.* There is no antistate to an arbitrary, entangled pure state of two qubits.

*Proof.* Any pure state can be written in the Schmidt form  $|\psi\rangle = a|00\rangle + b|11\rangle$ , with real coefficients. In this basis, the only nonzero elements of the correlation tensor are

$$T_{zz}(\psi) = 1, \quad (6)$$

$$T_{xx}(\psi) = -T_{yy}(\psi) = 2ab. \quad (7)$$

Therefore, the hypothetical antistate (mixed states allowed) has to have  $T_{zz} = -1$ , and hence it lies in the subspace spanned by  $|01\rangle$  and  $|10\rangle$ . Since all such states have  $T_{xx} = T_{yy}$ , only the product state with  $ab = 0$  has an antistate. ■

There is strong numerical evidence that a pure genuinely  $N$ -partite entangled state of  $N = 4$  and  $N = 6$  qubits does not admit an antistate. We are therefore conjecturing this in general.

*Conjecture 1.* There is no antistate to an arbitrary genuinely  $N$ -partite entangled pure state of even- $N$  qubits.

*Evidence.* Our aim is to verify to a high precision whether an antistate exists to a preselected state  $|\psi\rangle$ . In our numerical approach we parametrize a candidate state  $|\phi\rangle$  and use simulated annealing [17] to globally minimize the length of correlation [6] of the even mixture  $\rho = \frac{1}{2}|\psi\rangle\langle\psi| + \frac{1}{2}|\phi\rangle\langle\phi|$ . The length of correlation is defined as

$$L(\rho) = \sum_{j_1, \dots, j_N=1}^3 T_{j_1 \dots j_N}^2(\rho). \quad (8)$$

If antistates to  $|\psi\rangle$  exist, then  $L(\rho)$  will converge to 0 while  $|\phi\rangle$  converges to an antistate.

We tested this algorithm on states of  $N = 3$  and  $N = 5$  qubits, using the genuinely  $N$ -partite entangled input state  $|\psi\rangle$ . The candidate state  $|\phi\rangle$  converged to an approximate antistate, in accordance with what is known about antistates of

an odd number of qubits. For  $N = 4$  and  $N = 6$  we tested both the states for which it is known that they have no antistate, the Greenberger-Horne-Zeilinger states [10], the  $W$  states, and the Dicke states as well as a thousand randomly chosen states. We found that their length of correlation always converged to a finite value larger than 0, which indicates that there is no antistate.

On the other hand, in addition to  $|\phi\rangle$  we also varied the state  $|\psi\rangle$  as we minimized the length of correlation. In this case, for all choices of initial states, the algorithm quickly converged to a pair of state-antistate. However, all such states were not genuinely  $N$ -partite entangled. Instead, each pair was of the form  $\{|\psi_{N-1}\rangle \otimes |\phi_1\rangle, |\bar{\psi}_{N-1}\rangle \otimes |\phi_1\rangle\}$ , i.e., a biproduct of a state or antistate between (odd)  $N - 1$  qubits and a common single-qubit state. This strongly suggests that a genuinely multipartite entangled state of  $N = 4$  and  $N = 6$  qubits does not have an antistate. ■

#### A. Entanglement without correlations

Although antistates to *pure*  $N$ -partite entangled states most likely do not exist, one can find antistates to *mixed* entangled states. These can subsequently be used to construct examples of states with no  $N$ -partite correlation functions yet containing genuine  $N$ -partite entanglement. The simplest such example involves four qubits. For two qubits, while antistates to mixed entangled states can easily be constructed, all states with vanishing correlation functions are separable. A simple antistate example can be seen as follows. Consider the state  $\rho = \frac{1}{2}|\psi^+\rangle\langle\psi^+| + \frac{1}{2}|11\rangle\langle 11|$ . Being a mixture of a pure entangled state and a product state,  $\rho$  is entangled [18,19]. Its antistate is given by  $\bar{\rho} = \frac{1}{2}|\psi^-\rangle\langle\psi^-| + \frac{1}{2}|11\rangle\langle 11|$  as can be directly verified. The antistate is also entangled by the same argument, but the even mixture of the two states,  $\frac{1}{2}(\rho + \bar{\rho})$ , is separable.

The following theorem proves in general the absence of entanglement in bipartite states without bipartite correlation functions.

**Theorem 2.** Two-qubit states with vanishing bipartite correlation functions are separable.

*Proof.* The most general bipartite state with vanishing bipartite correlation functions is of the form

$$\rho = \frac{1}{4}(\mathbb{1} + \vec{a} \cdot \vec{\sigma} \otimes \sigma_0 + \sigma_0 \otimes \vec{b} \cdot \vec{\sigma}), \quad (9)$$

where  $\mathbb{1}$  denotes the identity operator in the space of two qubits,  $|\vec{a}| \leq 1$ , and similarly  $|\vec{b}| \leq 1$ . It has eigenvalues  $\frac{1}{4}(1 \pm |\vec{a}|^2 \pm |\vec{b}|^2)$  with all four sign combinations allowed. The same eigenvalues are obtained after partially transposing  $\rho$ . Hence all of such states are separable [20,21]. ■

Theorem 2 does not generalize to  $N > 2$ . A similar code to the one used in the evidence for Conjecture 1 returned a rank-3 genuinely four-party entangled state of four qubits with no four-partite correlation functions. Here we provide an analytical example of the rank-4 state for arbitrary even  $N \geq 4$ , giving rise to entanglement without correlations. Consider a mixed state

$$\rho_0 = \frac{1}{4}|\psi_1\rangle\langle\psi_1| + \dots + \frac{1}{4}|\psi_4\rangle\langle\psi_4|, \quad (10)$$

which mixes the following pure states:

$$\begin{aligned} |\psi_1\rangle &= \frac{1}{\sqrt{2}}(|0\dots 0\rangle|\psi\rangle + |\psi\rangle|0\dots 0\rangle), \\ |\psi_2\rangle &= \frac{1}{\sqrt{2}}(|1\dots 1\rangle|\bar{\psi}\rangle - |\bar{\psi}\rangle|1\dots 1\rangle), \\ |\psi_3\rangle &= |\phi\rangle|\psi\rangle, \\ |\psi_4\rangle &= |\psi\rangle|\phi\rangle, \end{aligned} \quad (11)$$

where we take the  $|\psi\rangle$  state and its antistate  $|\bar{\psi}\rangle$  as generalized  $W$  states of  $N/2$  qubits

$$\begin{aligned} |\psi\rangle &= \alpha_1|10\dots 0\rangle + \alpha_2|01\dots 0\rangle + \dots + \alpha_{N/2}|00\dots 1\rangle, \\ |\bar{\psi}\rangle &= \alpha_1|01\dots 1\rangle + \alpha_2|10\dots 1\rangle + \dots + \alpha_{N/2}|11\dots 0\rangle. \end{aligned}$$

It is assumed that all the coefficients are real and strictly positive, i.e.,  $\alpha_n > 0$ . The state  $|\phi\rangle$  is any product state containing an odd number of excitations, i.e., ones. To show that  $\rho_0$  is genuinely multipartite entangled, one may attempt to seek suitable entanglement witnesses. Here we present a much simpler approach. For that, we need the following theorem.

**Theorem 3.** If a state  $\rho$  lies in the subspace spanned by  $\{|\psi_1\rangle, |\psi_2\rangle, |\psi_3\rangle, |\psi_4\rangle\}$  given in Eq. (11) and  $\rho$  is biseparable, then  $\rho$  is orthogonal to  $|\psi_1\rangle$ , i.e.,

$$\text{Tr}(\rho|\psi_1\rangle\langle\psi_1|) = 0. \quad (12)$$

*Proof.* We first prove that all biproduct pure states in this subspace are orthogonal to  $|\psi_1\rangle$ . Suppose there exists a biproduct state  $|\xi\rangle_A|\eta\rangle_B \in \text{span}\{|\psi_1\rangle, \dots, |\psi_4\rangle\}$ , i.e.,

$$|\xi\rangle_A|\eta\rangle_B = c_1|\psi_1\rangle + c_2|\psi_2\rangle + c_3|\psi_3\rangle + c_4|\psi_4\rangle. \quad (13)$$

Here  $A, B$  form an *arbitrary* bipartition of all  $N$  qubits. We emphasize that, say,  $A$  contains any subset of qubits, not even neighboring ones. Denote  $a_0 = \langle 00\dots 0|\xi\rangle_A$  and  $b_0 = \langle 00\dots 0|\eta\rangle_B$ . Since the discussed subspace is orthogonal to the  $|0\dots 0\rangle$  state of all  $N$  qubits, by taking the inner product with both sides of Eq. (13), we conclude that

$$a_0b_0 = 0. \quad (14)$$

Now consider the vector  $\langle 00\dots 0|_A\langle 10\dots 0|_B$ , where the excitation 1 is in the first qubit of the subsystem  $B$ . Since this vector has only one excitation, it is orthogonal to  $|\psi_2\rangle$  (which has  $N - 1$  excitations) and both  $|\psi_3\rangle$  and  $|\psi_4\rangle$  (they have an even number of excitations). The inner product with both sides of Eq. (13) gives

$$a_0b_1 = c_1\alpha_k, \quad (15)$$

where  $b_1 = \langle 10\dots 0|\eta\rangle_B$  and index  $k$  depends on which bipartition is chosen. For example, if  $A$  contains first half of the qubits, then comparison with Eq. (11) shows that  $k = 1$ , or if  $A$  contains first  $N - 1$  qubits, then the same analysis reveals that  $k = N$ . Since we assume that all  $\alpha_n > 0$ , the state  $|\psi_1\rangle$  is a superposition of one excitation on every qubit and hence for arbitrary bipartition there exists index  $k$  such that Eq. (15) holds. Similarly, by taking the inner product with the vector  $\langle 10\dots 0|_A\langle 00\dots 0|_B$ , we obtain

$$a_1b_0 = c_1\alpha_l, \quad (16)$$

where  $a_1 = \langle 10\dots 0|\xi\rangle_A$  and  $\alpha_l$  is the suitable coefficient of  $|\psi_1\rangle$ . Multiplying Eq. (15) by (16) shows that  $a_0b_0a_1b_1 = c_1^2\alpha_k\alpha_l$ . The left-hand side of this equation is zero, as we

MINH CONG TRAN *et al.*PHYSICAL REVIEW A **95**, 062331 (2017)

have shown in Eq. (14). Since both  $\alpha_k$  and  $\alpha_l$  are strictly positive numbers, we conclude that  $c_1 = 0$ . In other words, all biproduct states in the discussed subspace are orthogonal to  $|\psi_1\rangle$ . Hence, arbitrary mixture of such states is also orthogonal to  $|\psi_1\rangle$  and the theorem follows. ■

According to this theorem, if a general state  $\rho$  has nonzero overlap with  $|\psi_1\rangle$ , then either  $\rho$  is genuinely  $N$ -qubit entangled or it does *not* belong to the subspace spanned by  $\{|\psi_1\rangle, \dots, |\psi_4\rangle\}$  or both. Since the state  $\rho_0$  presented in Eq. (10) clearly belongs to this subspace, it has to be genuinely  $N$ -qubit entangled. Furthermore, we prove in the Appendix that this state has no  $N$ -partite correlation functions. This concludes construction of entanglement without correlations for any even  $N \geq 4$ .

The construction just given also sheds light on the kind of operations required to produce an antistate. In particular, one could consider a mixed state  $\rho = \frac{1}{2}|\psi_3\rangle\langle\psi_3| + \frac{1}{2}|\psi_4\rangle\langle\psi_4|$ , which is clearly biseparable. By our construction, its antistate is  $\bar{\rho} = \frac{1}{2}|\psi_1\rangle\langle\psi_1| + \frac{1}{2}|\psi_2\rangle\langle\psi_2|$ , which is genuinely  $N$ -partite entangled. Hence, at least some of the antistates cannot be obtained by local operations and classical communication because this class of maps is not capable of producing entanglement.

#### B. Violation of local realism

Another remarkable property of states in Eq. (10) is their ability to violate a Bell inequality. The lack of  $N$ -partite correlation functions makes many standard tools inapplicable to these states. This was first pointed out in [8] and only recently suitable Bell inequalities were found [9,10] and were experimentally implemented to test the no-correlation states of an odd number of qubits [10]. We present now a Bell-type inequality which is violated by appropriate quantum measurements on states (10) for arbitrary even  $N$ .

Consider the following Bell-type inequality introduced in Ref. [22]:

$$0 \leq P(+\dots + |A_1 \dots A_{N-2}) \text{CH}_{N-1,N}^{+\dots+}, \quad (17)$$

where  $P(+\dots + |A_1 \dots A_{N-2})$  is the probability that the first  $N-2$  parties all detect  $+1$  outcomes when they measure observables  $A_1, \dots, A_{N-2}$ , respectively;  $\text{CH}_{N-1,N}^{+\dots+}$  denotes the Clauser-Horne expression [23] between the last two parties, which is calculated in the subensemble of experiments in which the first  $N-2$  observers all obtain  $+1$ .

For simplicity let us choose  $|\psi\rangle$  as the symmetric  $W$  state of  $N/2$  qubits, i.e., all  $\alpha_n = 1/\sqrt{N/2}$ . In order to demonstrate a violation of Eq. (17), each of the first  $N-2$  observers performs measurement  $A_n = \sigma_z$ . Therefore,  $P(+\dots + |A_1 \dots A_{N-2}) = \frac{1}{4} \frac{1}{N}$  with the sole contribution from the state  $|\psi_1\rangle$ . In the subensemble where all these  $N-2$  results are  $+1$  the state of the last two qubits collapses to  $\frac{1}{\sqrt{2}}(|01\rangle + |10\rangle)$ . The last two observers perform measurements that lead to the maximal violation of the CH inequality given by  $-\frac{\sqrt{2}-1}{2}$ . Finally, the right-hand side of Eq. (17) is equal to  $-\frac{\sqrt{2}-1}{4N}$ , which violates the lower bound 0. We also verified, using the software described in Ref. [24], that the above inequality is optimal in the sense that it is violated for the highest admixture of white noise to the state  $\rho$ .

We note that additionally to fundamental interest this also demonstrates practical applicability of states (10). It is well-known that such states reduce communication complexity, improve security of cryptographic key distribution, or enable device-independent protocols [25].

#### IV. N ODD

Reference [10] demonstrated a continuous family of mixed states which are genuinely tripartite entangled and give rise to vanishing tripartite correlation functions. In this section we will extend this example to a larger family of states described by exponentially many, in  $N$ , parameters. This example will then be shown to elucidate limits on entanglement detection with bipartite correlation functions only, such as those discussed in Refs. [11–16].

Consider the family of generalized Dicke states of  $N$  qubits:

$$|D_N^e\rangle = \sum_{\mathcal{P}} \alpha_{\mathcal{P}(1\dots 10\dots 0)} |\mathcal{P}(\underbrace{1\dots 1}_e \underbrace{0\dots 0}_{N-e})\rangle, \quad (18)$$

where the sum is over all permutations of  $e$  excitations, i.e., in every term in superposition we have  $e$  ones and  $N-e$  zeros. We assume that all the coefficients are strictly positive and we shall collectively denote them by  $\alpha_{\mathcal{P}}$ , i.e., we take  $\alpha_{\mathcal{P}} > 0$ . Note that the highest number of terms in the superposition is obtained for  $e = (N \pm 1)/2$  (recall that  $N$  is odd) and according to the Stirling approximation it scales as  $2^N/\sqrt{N}$ . We show that for all these exponentially many continuous parameters, the following even mixture

$$\rho = \frac{1}{2} |D_N^e\rangle\langle D_N^e| + \frac{1}{2} |D_N^{N-e}\rangle\langle D_N^{N-e}| \quad (19)$$

has vanishing all  $N$ -partite correlation functions and simultaneously it is genuinely  $N$ -partite entangled.

The former statement follows immediately from the results in Ref. [10]. Namely, one verifies that  $|D_N^{N-e}\rangle$  is the antistate to the generalized Dicke state  $|D_N^e\rangle$ . Any state of odd number of qubits equally mixed with its antistate has no  $N$ -partite correlation functions. The following theorem proves genuine multipartite entanglement.

*Theorem 4.* For all  $\alpha_{\mathcal{P}} > 0$  the state (19) is genuinely  $N$ -partite entangled.

*Proof.* We shall prove that no biproduct state exists in the subspace spanned by  $\{|D_N^e\rangle, |D_N^{N-e}\rangle\}$  if all  $\alpha_{\mathcal{P}} > 0$ . The following simple observation will be utilized: correlation functions of a biproduct state across the  $A : B$  partition satisfy

$$T_{0\dots 0x0\dots 0} T_{0\dots 0yy0\dots 0} = T_{0\dots 0xy0\dots 0} T_{0\dots 0yx0\dots 0}, \quad (20)$$

where the first nonzero index is for the last particle in  $A$ , and the second nonzero index is for the first particle in  $B$ . We now prove that this condition is not satisfied by any pure state in the considered subspace. Hence it contains no biproduct states and it follows that also all the mixed states with this support are genuinely  $N$ -partite entangled.

An arbitrary pure state in the considered subspace can be written as

$$|\phi\rangle = a |D_N^e\rangle + b |D_N^{N-e}\rangle, \quad (21)$$

where  $a$  and  $b$  are normalized complex coefficients. Without loss of generality we focus on the correlation functions

between the last two particles:

$$\begin{aligned} T_{0\dots 0jk} &= |a|^2 T_{0\dots 0jk}(D_N^e) + |b|^2 T_{0\dots 0jk}(D_N^{N-e}) \\ &\quad + a^* b \langle D_N^e | \sigma_0 \otimes \dots \otimes \sigma_0 \otimes \sigma_j \otimes \sigma_k | D_N^{N-e} \rangle \\ &\quad + ab^* \langle D_N^{N-e} | \sigma_0 \otimes \dots \otimes \sigma_0 \otimes \sigma_j \otimes \sigma_k | D_N^e \rangle, \end{aligned} \quad (22)$$

for  $j, k = x, y$ . Note that applying  $\sigma_j \otimes \sigma_k$  to the states  $|D_N^e\rangle$  and  $|D_N^{N-e}\rangle$  does not change their excitation parity. Since  $N$  is odd,  $|D_N^e\rangle$  and  $|D_N^{N-e}\rangle$  have opposite excitation parity. Thus the last two terms in Eq. (22) vanish. Furthermore, the bipartite correlation functions of the antistate to the generalized Dicke state are the same as in the original state. We conclude that the bipartite correlations of any state  $|\phi\rangle$  are the same as those of the generalized Dicke state  $|D_N^e\rangle$ . One now readily verifies that for the generalized Dicke state we have

$$\begin{aligned} T_{0\dots 0xy} &= T_{0\dots 0yx} = 0, \\ T_{0\dots 0xx} &= T_{0\dots 0yy} = \sum_P \alpha_{\mathcal{P}(1\dots 10\dots 0)01} \alpha_{\mathcal{P}(1\dots 10\dots 0)10}, \end{aligned} \quad (23)$$

where the sum is over all permutations of  $e - 1$  excitations on  $N - 2$  positions. Since all  $\alpha_{\mathcal{P}} > 0$ , Eq. (20) is never satisfied. The same argument holds for arbitrary partitions  $A : B$ . ■

#### A. Limits on entanglement witnesses based on bipartite correlations

Note that the proof of genuine  $N$ -partite entanglement of the state in Eq. (19) uses solely its bipartite correlation functions. Furthermore, it relies on the fact that some of these correlations vanish, as in Eq. (23). Naturally one would wonder if it is possible to conclude the genuine multipartite entanglement from only nonzero bipartite correlation functions. This is important especially in view of entanglement witnesses which are combinations of correlation functions and therefore are insensitive to the vanishing correlation functions [11–16].

We now show that in general the vanishing bipartite correlation functions are important for revealing genuine  $N$ -partite entanglement. Without taking them into account even entanglement of some manifestly genuinely  $N$ -partite entangled Dicke states is not detectable. The Dicke state with  $e$  excitations is defined by all the coefficients  $\alpha_{\mathcal{P}} = 1/\sqrt{\binom{N}{e}}$ . It is a permutation-invariant state with the following nonvanishing bipartite correlation functions:

$$\begin{aligned} T_{\mathcal{P}(xx0\dots 0)} &= T_{\mathcal{P}(yy0\dots 0)} = \frac{2}{\binom{N}{e}} \binom{N-2}{e-1}, \quad (24) \\ T_{\mathcal{P}(zz0\dots 0)} &= \frac{1}{\binom{N}{e}} \left\{ \binom{N-2}{e} + \binom{N-2}{e-2} - 2 \binom{N-2}{e-1} \right\}. \end{aligned} \quad (25)$$

Using the property of the binomial coefficients  $\binom{n-1}{k-1} + \binom{n-1}{k} = \binom{n}{k}$ , one verifies that

$$T_{xx0\dots 0} + T_{yy0\dots 0} + T_{zz0\dots 0} = 1. \quad (26)$$

Therefore, as long as the correlation functions in Eq. (26) are non-negative, we can always construct a pure single-qubit state  $|\phi\rangle$ , with Bloch vector  $(\sqrt{T_{xx0\dots 0}}, \sqrt{T_{yy0\dots 0}}, \sqrt{T_{zz0\dots 0}})$ , so

that the tensor product  $|\phi\rangle \otimes \dots \otimes |\phi\rangle$  mimics all the nonzero bipartite correlations of the Dicke state. The non-negativity of all the terms in Eq. (26) is satisfied for

$$N \geq \lceil \frac{1}{2}(1 + 4e + \sqrt{1 + 8e}) \rceil, \quad (27)$$

where  $\lceil x \rceil$  denotes the smallest integer greater than or equal to  $x$ . For example, the nonzero bipartite correlation functions of the  $|W\rangle$  state, i.e., Dicke state with  $e = 1$ , are compatible with the correlation functions of the product state for all  $N \geq 4$ , hence practically for all the  $|W\rangle$  states. For such states the nonzero bipartite correlations alone are not able to reveal genuine  $N$ -partite entanglement. However, when combined with the vanishing bipartite correlations a suitable proof may be found as we illustrated above.

#### V. GENERAL $N$

We would like to present here an observation which in a simple way characterizes all known facts about the existence of antistates for both  $N$  even and odd. It provides yet another piece of evidence that arbitrary genuinely  $N$ -partite entangled pure state of even number of qubits does not admit an antistate.

Consider a state  $|\psi\rangle$  endowed with correlation tensor  $T_{j_1\dots j_N}$ . Recall that its antistate is defined by having correlation tensor elements given by  $-T_{j_1\dots j_N}$ , for all indices  $j_n = x, y, z$ . One way of obtaining an antistate would be to apply onto an odd number of qubits a local operation which maps

$$\vec{x} \rightarrow -\vec{x}, \quad \vec{y} \rightarrow -\vec{y}, \quad \vec{z} \rightarrow -\vec{z}. \quad (28)$$

However, it is well known that such a local operation, called a universal-NOT gate [26], is not present within quantum formalism because it is antiunitary. On the level of multiple qubits one can to some degree overcome this restriction. Namely, note that mathematically one obtains Eq. (28) by applying the  $\sigma_y$  operation and partial transposition. The effect of  $\sigma_y$  is to invert  $\vec{x} \rightarrow -\vec{x}$  and  $\vec{z} \rightarrow -\vec{z}$ , and the effect of partial transposition is to flip the remaining axis  $\vec{y} \rightarrow -\vec{y}$ . If partial transposition is applied on a subsystem  $A$  of a pure state entangled across  $A : B$  it results in a matrix with negative eigenvalues [20,21]. Hence, this method leads to a physically meaningful antistate only for original states with odd total number of qubits (as applying partial transposition on every individual qubit results in a transposition, which is a completely positive map) or having a subsystem  $A$  with an odd number of qubits in a product state. For example, by applying  $\sigma_y$  and partial transposition on every single qubit or by taking  $A$  as the first qubit this procedure will produce an antistate to the three-qubit  $|0\rangle|\psi^-\rangle$  state, but no antistate to a four-qubit  $|\psi^-\rangle|\psi^-\rangle$  state, and indeed any genuinely multipartite entangled pure state of even number of qubits. Of course global operations may exist that produce antistates in a completely different way, but nevertheless it is appealing that this simple procedure recovers all that is presently known about antistates.

#### A. Impossibility of inverting all correlation functions between even number of observers

We would like to finish with one more observation contrasting even and odd lower-order correlation functions

MINH CONG TRAN *et al.*PHYSICAL REVIEW A **95**, 062331 (2017)

in an antistate for general  $N$ . The antistates constructed in Ref. [10] have opposite correlation functions between arbitrary odd number of observers, as compared to the original state. The correlation functions between an arbitrary even number of observers are the same as in the original state. In contrast, there is no state in which all the correlation functions between arbitrary even number of observers are opposite.

*Theorem 5.* Any pure state  $|\psi\rangle$  of  $N$  qubits does not admit state  $|\psi'\rangle$  in which all the  $k$ -partite correlation functions, for all even  $k$ , are opposite.

*Proof.* By contradiction. Let us build an antistate to the hypothetical state  $|\psi'\rangle$  according to prescription of Ref. [10]. Denote it  $|\bar{\psi}'\rangle$  and note that it has opposite all the correlation functions between even and odd number of observers, as compared to the original state  $|\psi\rangle$ . Therefore the even mixture

$$\rho = \frac{1}{2}|\psi\rangle\langle\psi| + \frac{1}{2}|\bar{\psi}'\rangle\langle\bar{\psi}'| \quad (29)$$

has no correlations whatsoever, including expectation values of local observables, i.e.,  $\rho$  is a white noise  $\mathbb{I}/2^N$ . However, this is not possible since the rank of  $\rho$  is 2, while the white noise must have rank  $2^N$ . ■

## VI. CONCLUSIONS

We provided nontrivial examples of genuinely multipartite entangled states of even number  $N$  of qubits that simultaneously have vanishing  $N$ -partite correlation functions. We showed that they violate suitable Bell-type inequalities. The states have rank 4 and rank 3, respectively, and we gave compelling evidence supporting the conjecture that rank-2 examples do not exist. This is in contrast to multipartite systems with odd number of qubits and explains why only such cases were considered up to date. We also extended previously known examples using techniques that show limits to entanglement detection with bipartite correlation functions only.

The states discussed here opened a debate on rigorous quantification of genuine multipartite classical and quantum correlations that led to the formulation of the natural postulates such quantifiers should satisfy [27]. We hope that the examples provided here will be a useful test bed for candidate identifiers and will help to find computable measures that will enable a deeper analysis of multipartite experiments.

## ACKNOWLEDGMENTS

This work is supported by the Singapore Ministry of Education Academic Research Fund Tier 2 project MOE2015-T2-2-034 and NCN Grant No. 2014/14/M/ST2/00818. We acknowledge the support of this work by the EU (ERC QOLAPS). L.K. thanks the international Ph.D. program ExQM from the Elite Network of Bavaria for support. M.C.T acknowledges support from the NSF-funded Physics Frontier Center at the JQI.

## APPENDIX

*Theorem 6.* All  $N$ -partite correlation functions of the state in Eq. (10) vanish.

*Proof.* To simplify notation we divide all  $N$  observers into Alice and Bob, each in possession of  $N/2$  qubits. The  $N$ -partite quantum correlation functions are written as  $T_{AB}(\rho)$ , with  $A$  and  $B$  being sequences, each of length  $N/2$ , of indices  $x, y, z$ . For example,

$$2 T_{AB}(\psi_2) = \langle\psi_2|\sigma_A \otimes \sigma_B|\psi_2\rangle \quad (A1)$$

$$\begin{aligned} &= \langle 1 \dots 1 | \sigma_A | 1 \dots 1 \rangle \langle \bar{\psi} | \sigma_B | \bar{\psi} \rangle \\ &\quad + \langle \bar{\psi} | \sigma_A | \bar{\psi} \rangle \langle 1 \dots 1 | \sigma_B | 1 \dots 1 \rangle \\ &\quad - \langle 1 \dots 1 | \sigma_A | \bar{\psi} \rangle \langle \bar{\psi} | \sigma_B | 1 \dots 1 \rangle \\ &\quad - \langle \bar{\psi} | \sigma_A | 1 \dots 1 \rangle \langle 1 \dots 1 | \sigma_B | \bar{\psi} \rangle. \end{aligned} \quad (A2)$$

Note that  $|\bar{\psi}\rangle = \sigma_X \otimes \sigma_X |\psi\rangle$ , where each  $\sigma_X \equiv \sigma_x \otimes \dots \otimes \sigma_x$  operates on all the qubits of Alice and Bob. Furthermore, for Alice we have  $\sigma_X \sigma_A \sigma_X = (-1)^a \sigma_A$ , where  $a$  is the number of  $x$  indices appearing in the sequence  $A$ . Similarly,  $\sigma_X \sigma_B \sigma_X = (-1)^b \sigma_B$ , where  $b$  is the number of  $x$  indices appearing in the sequence  $B$ . Therefore, if  $a + b$  is even the  $N$ -partite correlation functions of  $\rho$  read

$$T_{AB}(\rho) = \frac{1}{4} \sum_{i=1}^4 T_{AB}(\psi_i) \quad (A3)$$

$$\begin{aligned} &= \langle 00 \dots 0 | \sigma_A | 00 \dots 0 \rangle \langle \psi | \sigma_B | \psi \rangle \\ &\quad + \langle \psi | \sigma_A | \psi \rangle \langle 00 \dots 0 | \sigma_B | 00 \dots 0 \rangle \\ &\quad + \langle \phi | \sigma_A | \phi \rangle \langle \psi | \sigma_B | \psi \rangle + \langle \psi | \sigma_A | \psi \rangle \langle \phi | \sigma_B | \phi \rangle. \end{aligned} \quad (A4)$$

But due to an odd number of excitations in  $|\phi\rangle$  we have that  $\langle 00 \dots 0 | \sigma_A | 00 \dots 0 \rangle$  and  $\langle \phi | \sigma_A | \phi \rangle$  are either both zero or have opposite sign (and the same for Bob). We thus arrive at vanishing  $N$ -partite correlation functions of  $\rho$ .

If  $a + b$  is odd, we instead have

$$\begin{aligned} T_{AB}(\psi_1) + T_{AB}(\psi_2) &= \langle 0 \dots 0 | \sigma_A | \psi \rangle \langle \psi | \sigma_B | 0 \dots 0 \rangle \\ &\quad + \langle \psi | \sigma_A | 0 \dots 0 \rangle \langle 0 \dots 0 | \sigma_B | \psi \rangle. \end{aligned} \quad (A5)$$

Since by our assumption  $|\psi\rangle$  is a superposition of states with only one excitation, both terms above vanish unless  $A$  and  $B$  each has only one  $x$  or  $y$  index. Hence together they must have in total an even number of  $x$  and  $y$  indices. But the number of  $x$  indices, i.e.,  $a + b$  is assumed to be odd, so the number of  $y$  indices must also be odd. Therefore both terms in Eq. (A5) are imaginary and since they are complex adjoints of each other the sum  $T_{AB}(\psi_1) + T_{AB}(\psi_2)$  vanishes. Meanwhile, the contribution from  $|\psi_3\rangle$  is

$$T_{AB}(\psi_3) = \langle \phi | \sigma_A | \phi \rangle \langle \psi | \sigma_B | \psi \rangle. \quad (A6)$$

For this to be nonzero,  $\sigma_A$  must be  $\sigma_z \otimes \dots \otimes \sigma_z$ , and therefore has no  $x$  in the sequence:  $a = 0$ . Since  $a + b$  is odd,  $b$  must be odd. But  $\langle \psi | \sigma_B | \psi \rangle$  is nonzero only if  $B$  contains an even number of  $x$  and  $y$  indices in total. Thus the number of  $y$  indices must be odd leading to an imaginary  $T_{AB}(\psi_3)$ . Since the correlation function is defined as the average of real numbers it is always real valued. We conclude that  $T_{AB}(\psi_3) = 0$ . The same argument applies to  $|\psi_4\rangle$ . ■

GENUINE  $N$ -PARTITE ENTANGLEMENT WITHOUT  $N$ - ...PHYSICAL REVIEW A **95**, 062331 (2017)

- [1] C. H. Bennett, D. P. DiVincenzo, J. A. Smolin, and W. K. Wootters, *Phys. Rev. A* **54**, 3824 (1996).
- [2] R. Horodecki, P. Horodecki, M. Horodecki, and K. Horodecki, *Rev. Mod. Phys.* **81**, 865 (2009).
- [3] A. S. M. Hassan and P. S. Joag, *Quant. Inf. Comp.* **8**, 773 (2007).
- [4] A. S. M. Hassan and P. S. Joag, *Phys. Rev. A* **77**, 062334 (2008).
- [5] A. S. M. Hassan and P. S. Joag, *Phys. Rev. A* **80**, 042302 (2009).
- [6] M. C. Tran, B. Dakić, F. Arnault, W. Laskowski, and T. Paterek, *Phys. Rev. A* **92**, 050301(R) (2015).
- [7] M. C. Tran, B. Dakić, W. Laskowski, and T. Paterek, *Phys. Rev. A* **94**, 042302 (2016).
- [8] D. Kaszlikowski, A. S. De, U. Sen, V. Vedral, and A. Winter, *Phys. Rev. Lett.* **101**, 070502 (2008).
- [9] W. Laskowski, M. Markiewicz, T. Paterek, and M. Wieśniak, *Phys. Rev. A* **86**, 032105 (2012).
- [10] C. Schwemmer, L. Knips, M. C. Tran, A. de Rosier, W. Laskowski, T. Paterek, and H. Weinfurter, *Phys. Rev. Lett.* **114**, 180501 (2015).
- [11] G. Tóth and O. Gühne, *Phys. Rev. A* **72**, 022340 (2005).
- [12] G. Tóth, *J. Opt. Soc. Am. B* **24**, 275 (2007).
- [13] M. Wieśniak, M. Nawareg, and M. Żukowski, *Phys. Rev. A* **86**, 042339 (2012).
- [14] M. Markiewicz, W. Laskowski, T. Paterek, and M. Żukowski, *Phys. Rev. A* **87**, 034301 (2013).
- [15] J. Tura, R. Augusiak, A. B. Sainz, T. Vertesi, M. Lewenstein, and A. Acin, *Science* **344**, 1256 (2014).
- [16] J. Tura, A. B. Sainz, T. Vertesi, A. Acin, M. Lewenstein, and R. Augusiak, *J. Phys. A* **47**, 424024 (2014).
- [17] S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi, *Science* **220**, 671 (1983).
- [18] P. Horodecki, J. A. Smolin, B. M. Terhal, and A. V. Thapliyal, *Theor. Comput. Sci.* **292**, 589 (2003).
- [19] M. C. Tran, W. Laskowski, and T. Paterek, *J. Phys. A* **47**, 424025 (2014).
- [20] A. Peres, *Phys. Rev. Lett.* **77**, 1413 (1996).
- [21] M. Horodecki, P. Horodecki, and R. Horodecki, *Phys. Lett. A* **223**, 1 (1996).
- [22] W. Laskowski, T. Vertesi, and M. Wieśniak, *J. Phys. A* **48**, 465301 (2015).
- [23] J. F. Clauser and M. A. Horne, *Phys. Rev. D* **10**, 526 (1974).
- [24] J. Gruca, W. Laskowski, M. Żukowski, N. Kiesel, W. Wieczorek, C. Schmid, and H. Weinfurter, *Phys. Rev. A* **82**, 012118 (2010).
- [25] N. Brunner, D. Cavalcanti, S. Pironio, V. Scarani, and S. Wehner, *Rev. Mod. Phys.* **86**, 419 (2014).
- [26] V. Bužek, M. Hillery, and R. F. Werner, *J. Mod. Opt.* **47**, 211 (2000).
- [27] C. H. Bennett, A. Grudka, M. Horodecki, P. Horodecki, and R. Horodecki, *Phys. Rev. A* **83**, 012312 (2011).

## 4.5 Conclusion

Entanglement is a fascinating characteristic of quantum systems, which cannot be explained using classical systems. While detection and characterization of entanglement between two observers, usually implemented using a two-qubit quantum state, can be done fairly easily, the task becomes far more difficult for larger systems.

In this chapter, different tools for detecting entanglement between multiple observers have been proposed, which are optimized for specific purposes. The first method, derived, explained, and discussed in [P3], which is reprinted in Sec. 4.2.5, is optimal with respect to the number of experimentally required measurements and can, for suitable states, reveal genuine multipartite entanglement after measurements in only two measurement settings, even irrespective of the number of qubits. A similar approach can be generalized for higher dimensional states and has been adopted to systems with multiple qutrits, see [P4], which is reprinted in Sec. 4.2.7.

In Sec. 4.3, a scenario has been described in which all observers are unable to fix or even estimate their local reference frames. This restriction renders usual schemes for entanglement detection impossible as long as they are based on specific measurements or require to repeat a previous measurement later on. Still, the distribution of possible measurement outcomes contains valuable information and can be used to reveal not only the purity of the quantum state, but also certify genuine multipartite entanglement. It is an open question if this method allows to reconstruct the actual quantum state up to local unitary transformations when additional assumptions are added.

Finally, in Sec. 4.4.6 and 4.4.7, which are reprints of [P6] and [P7], the connection between entanglement and correlations has been discussed. Intuitively, a highly entangled state will lead to several correlations between the outcomes of various observers. However, in those publications, counterexamples to this perception are given. Substantiated by experimental results, a multipartite quantum state has been provided, which is genuinely multipartite entangled, while at the same time not containing any correlations between those observers. This puts the usual notion of quantum correlations into question.



# Chapter 5

## Quantum Measurements

### Contents

---

<b>5.1</b>	<b>Introduction</b>	<b>182</b>
<b>5.2</b>	<b>Weak Measurements and Weak Values</b>	<b>183</b>
5.2.1	Weak Amplification	186
5.2.2	Direct Tomography	186
5.2.3	Weak Value Controversies	186
5.2.4	Weak Value for Mixed Pre- and Postselection	187
5.2.5	Universal Description of Local Interactions	187
<b>5.3</b>	<b>Quantum Instruments and Universal Cloning</b>	<b>188</b>
5.3.1	Uncertainty Principle and Robertson Inequality	189
5.3.2	Heisenberg Microscope	190
5.3.3	Measurement-Disturbance Tradeoff Relations	191
5.3.4	[P10] Procedures for the Optimal Meas.-Disturbance Tradeoff	198
<b>5.4</b>	<b>Conclusion</b>	<b>209</b>

---

## 5.1 Introduction

All physical experiments eventually produce classical information, which can be distributed by spoken word, on paper or electronically. Yet, in order to retrieve this classical, clonable information from a system governed by the rules of quantum mechanics, some interaction between the classical measurement device and the quantum system has to take place. In this chapter, different types of measurements will be discussed, which allow to gain information about the current state of some quantum system, especially with respect to the change of the state due to the measurement.

In 1935, Dirac described how a measurement disturbs a quantum system by forcing it into one of its eigenstates [206]. He asserts that,

[w]hen we measure a real dynamical variable  $\xi$ , the disturbance involved in the act of measurement causes a jump in the state of the dynamical system. From physical continuity, if we make a second measurement of the same dynamical variable  $\xi$  immediately after the first, the result of the second measurement must be the same as that of the first. Thus after the first measurement has been made, there is no indeterminacy in the result of the second. Hence, after the first measurement has been made, the system is in an eigenstate of the dynamical variable  $\xi$ , the eigenvalue it belongs to being equal to the result of the first measurement. This conclusion must still hold if the second measurement is not actually made. In this way we see that a measurement always causes the system to jump into an eigenstate of the dynamical variable that is being measured, the eigenvalue this eigenstate belongs to being equal to the result of the measurement.

In this chapter, an introduction to the formalism of *weak measurements* and *weak values* will be given. There, the interaction between the system and the measurement device, the so-called *pointer system*, is kept weak or is restricted to a short duration. This way, in the ideal case, the *jump* as described by Dirac and with it the distortion to the quantum system are avoided. Additionally, weak measurements usually consider systems, which are pre- and postselected, i.e., the pointer system is considered only conditioned on a specific final state of the system itself.

In the second part of this chapter, the tradeoff relation between measurements and the thereby induced disturbance to the quantum system will be investigated based on the description of quantum instruments. This parametrization allows to derive the optimal instruments in terms of the measurement-disturbance tradeoff, which have also been analyzed experimentally in the course of this work.

Weak measurements consider pre- and postselected systems and an interaction strength which is in the ideal case negligibly small. Due to the small interaction strength, one does not gain much information from a single measurement. Additionally, depending on the postselection, the probability to obtain a result may become vanishingly small. On the other hand, every measurement that extracts information about a quantum system necessarily entangles the measurement system with the quantum system, which leads to

an unavoidable disturbance to the state of the quantum system when viewed on its own. In our work, we have studied how optimal measurement instruments can optimize this interaction, allowing to retrieve information, while minimizing the amount of disturbance.

## 5.2 Weak Measurements and Weak Values

Weak values have been introduced in 1988 by Aharonov, Albert, and Vaidman [207] in order to describe the outcome of a weak measurement on a pre- and postselected quantum system [208, 209]. In this situation, the usual measurement scenario, in which a quantum system is in a well defined, albeit possibly unknown, initial state, is extended by adding a postselection, i.e., a conditioning of the results on the outcome of a second (final) measurement. In Fig. 5.1, the scenario of a standard measurement is illustrated. The quantum

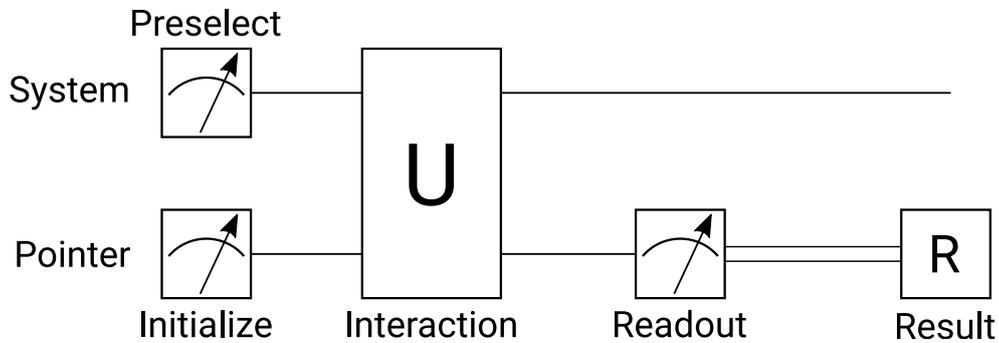


FIGURE 5.1: Standard measurement situation with interaction between quantum system of interest and pointer (or measurement) system, which is eventually read. The preselection and initialization is indicated as a (strong projective) measurement process as this ensures a well defined initial state. Adapted from [208].

system, which is to be measured, is brought to interaction with another quantum system, usually referred to as the *pointer system*, the *measurement device*, or the *meter*. From the measurement outcome of the pointer system measurement, one can deduce the value of the observable of the system.

In the weak measurement scenario, however, two changes occur to the conventional measurement scenario. Usually, the interaction between system and pointer is considered to be weak or short in order to avoid the measurement induced disturbance on the quantum system. This now becomes necessary as contrary to the scenario depicted in Fig. 5.1 also a postselection on the quantum system is introduced, which therefore is supposed to remain almost undisturbed. Only when the system is finally found in a specific state  $|\phi\rangle$ , the outcome of the measurement on the pointer is considered, see Fig. 5.2.

If the pointer is initially in state  $|\Phi_i\rangle$ , if the system is pre- and postselected to the states  $|\psi\rangle$  and  $|\phi\rangle$ , respectively, and if the interaction between system and pointer can be described by the Hamiltonian  $H = gA \otimes p$  (with some interaction parameter  $g$ ), which

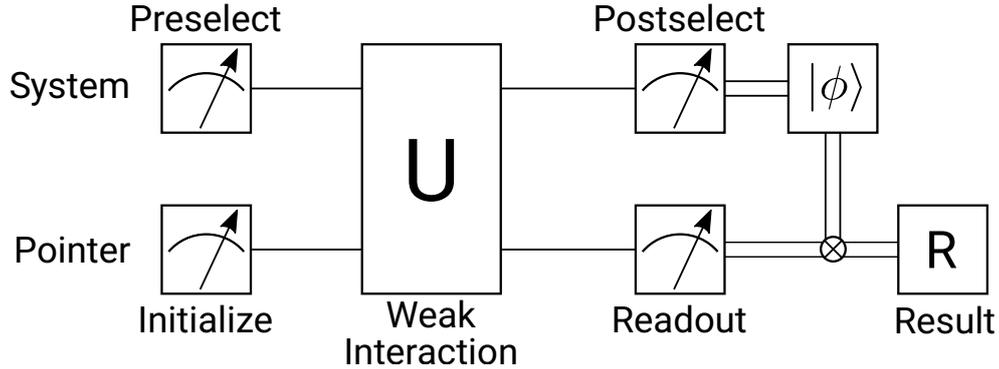


FIGURE 5.2: Weak measurement situation with weak interaction between quantum system of interest and pointer (or measurement) system, which is eventually read conditioned on the outcome of the quantum system itself. Adapted from [208].

leads to some unitary transformation  $U = \exp(-i\gamma A \otimes p)$ , one obtains for the final state of the pointer [207–210]

$$|\Phi_f\rangle = \langle\phi|U[|\psi\rangle \otimes |\Phi_i\rangle] \quad (5.1)$$

$$= \langle\phi| \exp(-i\gamma A \otimes p)[|\psi\rangle \otimes |\Phi_i\rangle] \quad (5.2)$$

$$\approx \langle\phi|(\mathbb{1} - i\gamma A \otimes p)[|\psi\rangle \otimes |\Phi_i\rangle] \quad (5.3)$$

$$= \langle\phi|\psi\rangle \left(1 - i\gamma \frac{\langle\phi|A|\psi\rangle}{\langle\phi|\psi\rangle} \otimes p\right) |\Phi_i\rangle \quad (5.4)$$

$$\approx \langle\phi|\psi\rangle \exp\left(-i\gamma \frac{\langle\phi|A|\psi\rangle}{\langle\phi|\psi\rangle} \otimes p\right) |\Phi_i\rangle, \quad (5.5)$$

where the approximations are due to first order expansions and become identities in the limit of infinitesimally small coupling between system and pointer. The expression

$$(A)_w \equiv \frac{\langle\phi|A|\psi\rangle}{\langle\phi|\psi\rangle} \quad (5.6)$$

is called *weak value* of operator  $A$  for the pre- and postselection states  $|\psi\rangle$  and  $|\phi\rangle$ , respectively.

As a direct consequence of its definition, the weak value is not restricted to real values within the set or range of eigenvalues of  $A$  as opposed to expectation values  $\langle A \rangle$ . Rather, weak values are in principle not limited at all and can also take complex values. Weak values outside the range of eigenvalues are referred to as *anomalous weak values*.

It is illustrative to demonstrate this concept using a polarized Gaussian beam, where the spatial degree of freedom provides the pointer system for the qubit system encoded in polarization. Using polarizers, a birefringent crystal and considering the spatial distribution on some screen, the weak measurement procedure can be easily visualized, see Fig. 5.3.

Considering the position  $x$  conjugate to the generator of the shift  $p$  in the Hamiltonian as given above, the shift of the pointer distribution is approximately given by

$$\langle \Phi_f | x | \Phi_f \rangle - \langle \Phi_i | x | \Phi_i \rangle = \gamma \Re [(A)_w], \quad (5.7)$$

while the momentum of the Gaussian pointer changes according to

$$\langle \Phi_f | p | \Phi_f \rangle - \langle \Phi_i | p | \Phi_i \rangle = 2\gamma (\Delta p)^2 \Im [(A)_w], \quad (5.8)$$

where  $\Re [(A)_w]$  and  $\Im [(A)_w]$  denote the real and imaginary parts of the weak value of the operator  $A$  with the pre- and postselection states  $|\psi\rangle$  and  $|\phi\rangle$ , respectively, and  $\Delta p$  is taken with respect to the initial pointer state  $|\Phi_i\rangle$ .

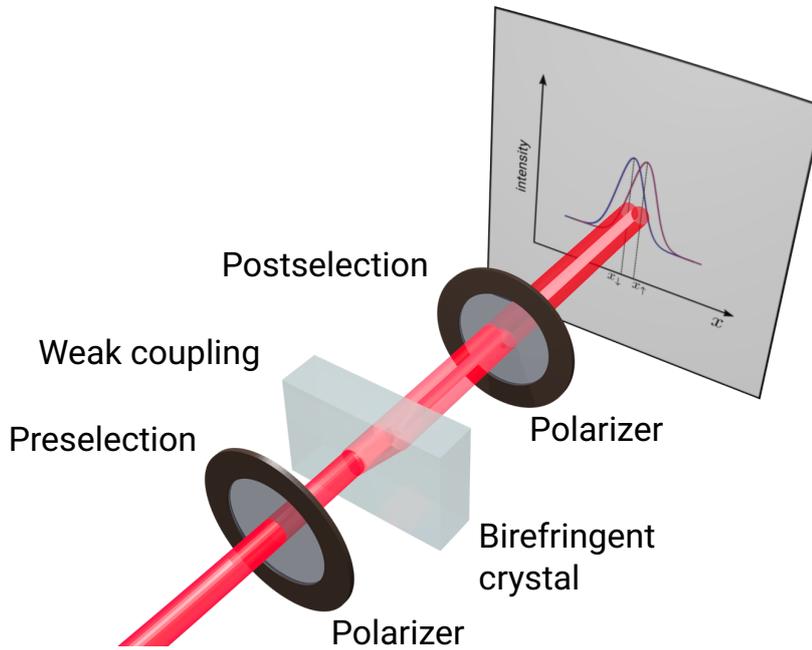


FIGURE 5.3: The concept of weak measurements illustrated using optical beams. The system is encoded in the polarization degree of freedom of the light with the preselection given by the first polarizer. The pointer system, encoded in the spatial distribution, is initially in a Gaussian mode with a given center. The waist corresponds to the initial uncertainty of the pointer. Using a birefringent crystal, the spatial distribution is changed depending on the polarization of the light. A second polarizer realizes the postselection and ensures that the pointer distribution on the screen is only considered conditioned on successful postselection. In the regime of weak coupling, the spatial distribution after postselection can still be described by a Gaussian distribution with a shift depending on the weak value. For illustrative purposes, the coupling shown here is already rather large.

Please note that, although weak values are introduced and usually considered to be the outcome of weak measurements, weak values can also be considered using strong measurements [211]. The authors of aforementioned reference disassociate the concepts of weak

values and weak measurements and allow the experimental determination of first without usage of latter.

### 5.2.1 Weak Amplification

The technique of weak amplification, sometimes called *weak value amplification*, makes use of the fact that for properly chosen pre- and postselection states a small shift can be significantly enhanced. According to  $\delta x = \langle \Phi_f | x | \Phi_f \rangle - \langle \Phi_i | x | \Phi_i \rangle = \gamma \Re [(A)_w]$ , see Eq. (5.7), the pointer position obtains a shift proportional to the parameter  $\gamma$ , which is given by the coupling strength and time, and to the real part of the weak value  $(A)_w$ , which depends on pre- and postselection [208, 209, 212–214]. Hence, if small effects leading to a small coupling between two different systems are to be observed, the weak amplification technique can be used.

This way, for example, the spin Hall effect of light can be measured as this leads, similar to above mentioned birefringent crystal, to a polarization dependent pointer displacement [215]. Even though  $\gamma$  is small in this case, choosing almost orthogonal pre- and postselection states results in a significant amplification of the pointer deflection signal. Similarly, very small beam deflections and displacements of the order of hundreds of femtorad and tens of femtometers have been detected [216].

### 5.2.2 Direct Tomography

Usually, for tomographically reconstructing a quantum state, a set of measurements are conducted, which allow to reconstruct a density matrix. There exists a unique connection between the true state and the probabilities for measurement outcomes, depending on the choice of the measurement operators. Using and inverting this relation, the outcomes can be used to reconstruct the state. For more details, please see Sec. 3.2.1.

One of the applications of weak values is the so-called *direct (state) tomography* [209, 217–222], as it does not involve the inversion of some reconstruction matrices. Instead, the outcome distributions are directly used to retrieve the elements of the density matrix or the amplitudes of the state vector, respectively.

### 5.2.3 Weak Value Controversies

Although weak measurements and weak values are of interest for some specific purposes and applications, there is still an ongoing debate about their usefulness [223]. Some part of the criticism questions the quantumness of weak values since a similar amplification effect can be reproduced using classical probability distributions with some uncertainty and a conditioning of results [224]. Thus, it is argued that instead of being a property of a quantum system, weak values are a mere statistical effect. This, of course, caused some controversial discussion [223–227].

Furthermore, the process of weak amplification is usually considered to be useful to detect small signals which could not have been detected without amplification. However,

as a fundamental property of the weak measurement process, a large portion of the original signal has to be removed, resulting in only a largely shifted pointer, which, however, had already been contained in the original pointer distribution. This argument triggered another discussion whether the weak amplification technique can be informationally useful [228]. Subsequently, several works discussed the role of different types of noise such as technical noise during the measurement and argued in favor of the usefulness of weak amplification in the presence of such noise [212, 229–231]. Yet, the authors of [232] argue that from a strictly metrological point of view the performance even is reduced by the usage of weak amplification.

### 5.2.4 Weak Value for Mixed Pre- and Postselection

The definition of the weak value of operator  $A$  in Eq. (5.6) is based on pure pre- and postselection states. Nevertheless, the concept can be generalized to mixed pre- and postselection states using the density matrices  $\varrho_{\text{pre}}$  and  $\varrho_{\text{post}}$ , respectively. A more general expression for the weak value allowing for mixed pre- and postselection is given by [P8]

$$(A)_w = \frac{\text{tr}(\varrho_{\text{post}} A \varrho_{\text{pre}})}{\text{tr}(\varrho_{\text{post}} \varrho_{\text{pre}})}. \quad (5.9)$$

The validity of this expression is proven and discussed in [P8], see also [213].

### 5.2.5 Universal Description of Local Interactions

As weak values are not restricted to the spectrum of eigenvalues, large (anomalous) weak values allow for the amplification of small effects as discussed in Sec. 5.2.1. This allows to measure small displacements and deflections [215, 216] and has been used to amplify some phase shifts by a factor of 8 [233]. However, by increasing the coupling strength between the pointer and the system, the coherence between the original pointer state and the pointer state after interaction decreases, limiting the achievable amplification.

To model the loss of coherence, consider a Mach-Zehnder interferometer with two arms,  $|A\rangle$  and  $|B\rangle$ . The initial superposition state (*preselection*) is given by

$$|\psi\rangle = \cos \alpha |A\rangle + \sin \alpha e^{i\varphi} |B\rangle, \quad (5.10)$$

where  $\alpha$  and  $\varphi$  can be tuned to vary the relative intensities of the arms and the relative phase, respectively. By introducing a small interaction in arm  $|A\rangle$ , the initial pointer state  $|\chi\rangle$  changes according to

$$|A\rangle \otimes |\chi\rangle \rightarrow |A\rangle \otimes |\chi'\rangle = \eta |A\rangle \otimes (|\chi\rangle + \epsilon |\chi^\perp\rangle) = \eta (|A\rangle \otimes |\chi\rangle + |A\rangle \otimes \epsilon |\chi^\perp\rangle) \quad (5.11)$$

with small  $\epsilon > 0$  and  $\eta = \langle \chi' | \chi \rangle = \frac{1}{\sqrt{1+\epsilon^2}}$ .

When considering now the pointer state in one of the output ports (*postselection*),

$$|\phi\rangle = \frac{1}{\sqrt{2}} (|A\rangle + |B\rangle), \quad (5.12)$$

one observes an amplification of the amplitude of the orthogonal component according to

$$|\chi\rangle \rightarrow |\chi\rangle + \epsilon (P_A)_w |\chi^\perp\rangle + \mathcal{O}(\epsilon^2), \quad (5.13)$$

where  $(P_A)_w$  denotes the weak value of the projection operator  $P_A = |A\rangle\langle A|$ ,

$$(P_A)_w = \frac{\langle\phi|P_A|\psi\rangle}{\langle\phi|\psi\rangle} = \frac{1}{1 + e^{i\varphi} \tan \alpha}. \quad (5.14)$$

Please note that this holds in the limit of small coupling  $\epsilon$ .

Contrary, for finite coupling, the system and pointer get entangled, rendering a pure state description of the pointer as in Eq. (5.13) impossible. By properly considering the arising entanglement between those two systems, the pre- and postselected state have to be described using density matrices as in Eq. (5.9) [214], see also [P8,P9]. Finally, the weak value of the projection onto path  $|A\rangle$  is then given by

$$(P_A)_w = \frac{\text{Tr}(|\phi\rangle\langle\phi|P_A\rho)}{\text{Tr}(|\phi\rangle\langle\phi|\rho)} = \frac{1 + \tan \alpha \eta e^{-i\varphi}}{1 + \tan^2 \alpha + 2 \tan \alpha \eta \cos \varphi}. \quad (5.15)$$

Consequently, the parameter  $\epsilon$  enters via  $\eta$  the weak value formula Eq. (5.15). Hence, for finite coupling ( $\epsilon > 0$ ,  $\eta < 1$ ), the entanglement between pointer and system causes decoherence, reducing the maximally achievable weak value. For three different values of  $\eta$ , Eq. (5.15) is plotted in Fig. 5.4. For more details, please see [P9] as well as [214].

Using the language of interferometer paths and weak interactions between the spatial superposition state and some external system, one can make use of the concept of weak values to compensate for unwanted misalignments. In [P9], the concept of weak amplification is used to devise a strategy to align optical interferometers with a single position resolving detector at an arbitrary, but fixed distance with a single phase scan of the interferometer. This allows to retrieve possible misalignments in position and angle between the different arms as well as to detect for other misalignments such as differences in the polarization of the arms [P9].

### 5.3 Quantum Instruments and Universal Cloning

Measurements are a fundamental ingredient to understand quantum theory. Contrary to classical physics, measurements cannot be considered as a mere observation of predefined values of a system. Instead, the backaction due to a measurement has to be taken into account. As stated by Dirac, see Sec. 5.1, the measurement process forces the system to one of the measurement operator's eigenstates, which generally leads to an unavoidable disturbance.

In this section, the concept of strong measurements is introduced and discussed in the context of measurement-disturbance relations. Based on the theoretic work of [234] and the joint theoretical and experimental work of [P10], different means of performing measurements on quantum systems are compared. This allows to derive measurement instruments attaining the optimal measurement-disturbance relation, which eventually is to be tested experimentally.

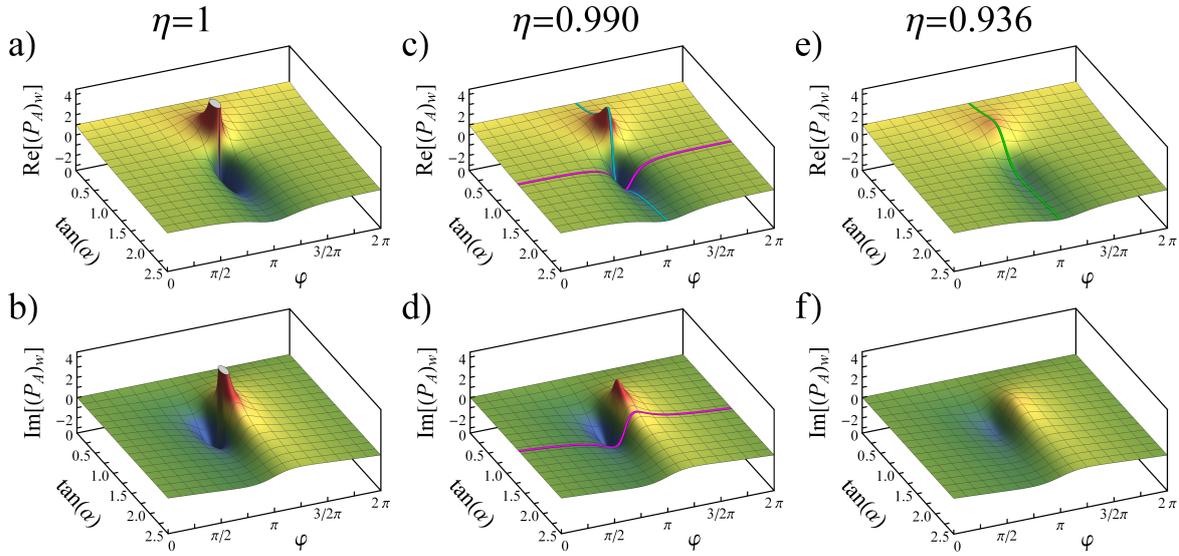


FIGURE 5.4: Weak value of the projection operator onto path  $|A\rangle$  for preselection  $|\psi\rangle = \cos\alpha|A\rangle + \sin\alpha e^{i\varphi}|B\rangle$  and postselection  $|\phi\rangle = \frac{1}{\sqrt{2}}(|A\rangle + |B\rangle)$ . With increasing coupling strength  $\epsilon$ , the coherence  $\eta$  between the pointer state in arm  $|A\rangle$  and the pointer state in arm  $|B\rangle$  decreases, limiting the maximal amplification. For infinitesimal small coupling,  $\eta = 1$ , a singularity in the surface plot occurs, leading to arbitrary large amplifications. In the realistic scenario, however, those plots are smoothed and flattened, restricting the amplification.

### 5.3.1 Uncertainty Principle and Robertson Inequality

For the product of the uncertainties (standard deviations)  $\sigma_x$  and  $\sigma_p$  of position  $x$  and momentum  $p$ , respectively, Heisenberg introduced the famous uncertainty relation [15, 235]

$$\sigma_x \sigma_p \sim \hbar, \quad (5.16)$$

where  $\sim$  suggests some lower bound on the mutual uncertainties, which, however, has not been properly derived by him<sup>1</sup>. Later on, Kennard [236] and Weyl [237] derived and proved the relation

$$\sigma_x \sigma_p \geq \frac{\hbar}{2}. \quad (5.17)$$

While the uncertainty principle was originally derived for position and momentum operators, the general mathematical formulation for two arbitrary observables  $A$  and  $B$  is denoted as *Robertson uncertainty relation* [238],

$$\sigma_A \sigma_B \geq \frac{|\langle \psi | [A, B] | \psi \rangle|}{2}, \quad (5.18)$$

<sup>1</sup>Please note that the symbol  $\sigma_x$  is used here to indicate the standard deviation of the position variable  $x$  and is not to be confused with the first Pauli matrix.

where  $\sigma_A \equiv \sigma_{A,\psi} = \sqrt{\langle \psi | A^2 | \psi \rangle - \langle \psi | A | \psi \rangle^2}$  denotes the standard deviation [239]. For the above example of position  $x$  and momentum  $p_x$ ,  $[x, p_x] = i\hbar$ , one obtains Eq. (5.17), which differs from Heisenberg's heuristically derived inequality in Eq. (5.16) by a constant factor.

Please note that the equation given above are lower bounds for the uncertainty when preparing two incommensurable observables simultaneously. They do not imply that measuring a first observable induces a disturbance to another one.

### 5.3.2 Heisenberg Microscope

One of the first approaches to understand the effect of inevitable backaction due to a measurement process has also been provided by Heisenberg by his famous gedankenexperiment of measuring the position of an electron using a microscope [4, 240, 241]. Position and momentum of an electron are to be measured by scattering a photon. For this purpose, a photon with initial wavelength  $\lambda_{\text{initial}}$  is sent towards an electron. The position of the electron can be determined by detection of the scattered photon, e.g., using a microscope with half opening angle  $\theta$  and lens diameter  $D$ . On the other hand, the electron's momentum in  $x$  direction will be changed due to the scattering process.

After scattering, the observed photon may have a wavelength between  $\lambda_{\text{final},1}$  and  $\lambda_{\text{final},2}$ , which in turn affects the momentum of the electron, which will be between  $-\frac{\lambda_{\text{final},1}}{h} \sin \theta$  and  $\frac{\lambda_{\text{final},2}}{h} \sin \theta$ . Thus, when using the small angle approximation  $\sin \theta \approx \theta$  and neglecting wavelength changes  $\lambda_{\text{final},1} \approx \lambda_{\text{final},2} \approx \lambda_{\text{initial}} \equiv \lambda$ , the momentum kick  $p_x$  of the electron in  $x$  direction will be

$$-\frac{h}{\lambda} \theta \leq p_x \leq \frac{h}{\lambda} \theta, \quad (5.19)$$

which thus introduces an uncertainty of  $\Delta p_x = 2\frac{h}{\lambda} \theta$ . On the other hand, the resolution of the microscope is limited by diffraction according to

$$\Delta x = \frac{\lambda}{\theta}, \quad (5.20)$$

while the backaction due to the momentum transfer leads to  $\Delta p_x = \frac{2h\theta}{\lambda}$ . Hence, the product of the uncertainties cannot be reduced arbitrarily, and is bounded as

$$\Delta x \Delta p_x \geq 2h. \quad (5.21)$$

The Heisenberg microscope is helpful to visualize the measurement induced disturbance to a quantum system. While it also leads to slightly different results than the uncertainty principle in Eq. (5.17), as mentioned above, it is important to point out that this gedankenexperiment is not a visualization of Heisenberg's uncertainty relation. The uncertainty principle compares the uncertainty in the preparation of a first observable (position) with the uncertainty in the *preparation* of a second observable (momentum) conjugate to the first one. Prevalently, the uncertainty principle is misunderstood as the tradeoff relation between the measured accuracy of a first observable and the thereby introduced disturbance

to a second observable [43], which is *not* described by Heisenberg's uncertainty principle, but by Heisenberg's microscope.

In [239] Heisenberg's measurement-disturbance relation is brought into a form similar to Eq. (5.18),

$$\epsilon(A)\eta(B) \geq \frac{|\langle\psi|[A,B]|\psi\rangle|}{2} \quad (5.22)$$

for an input state  $|\psi\rangle$  and the observables  $A$  and  $B$ .  $\epsilon(A) \equiv \epsilon(A, \psi)$  denotes the noise of the measurement, while  $\eta(B) \equiv \eta(B, \psi)$  corresponds to the backaction due to it.

Please note that the measurement-disturbance relation of Eq. (5.22) has been derived only heuristically and could be violated experimentally [242]. A more rigorous version of Eq. (5.22) has been derived in [239, 243]. This *generalized noise-disturbance uncertainty relation* [239] reads

$$\epsilon(A)\eta(B) + \eta(A)\sigma(B) + \sigma(A)\eta(B) \geq \frac{|\langle\psi|[A,B]|\psi\rangle|}{2}, \quad (5.23)$$

where additional terms contain the pre-measurement uncertainties  $\sigma(A)$  and  $\sigma(B)$  of the observables  $A$  and  $B$ , respectively. Though, this version is not tight in general. A tight measurement-disturbance relation, which is a generalization of Eq. (5.23), can be found in [244].

Depending on the context, there are different possibilities to pose the question of an optimal tradeoff relation between the error of a measurement and the disturbance induced by it. In general, one can distinguish between the different approaches based on whether the measurement error and the state disturbance are evaluated with respect to some reference measurements [234]. For example, the measurement-disturbance relation given by Ozawa is based on two reference observables. On the other hand, both quantities can be assessed without any reference.

### 5.3.3 Measurement-Disturbance Tradeoff Relations

In the theoretical approach of [234], a reference for the measurement error is used, whereas the disturbance introduced by the measurement is quantified absolutely, or, more formally, with the identity channel as a reference. Hence, if a perfect or ideal measurement is taken as reference, the measurement error indicates the information loss, i.e., the amount of information which is not retrieved via the actual measurement, but could have been obtained by an ideal measurement. Yet, if it is not fixed which measurement is to be conducted afterwards on the postmeasurement state, it is natural not to use any reference observable for quantifying the disturbance.

In this section and the subsequent publication, several strategies for the measurement of a quantum system are introduced and evaluated. An optimal strategy is derived and it is shown both theoretically and experimentally that this optimal strategy outperforms common strategies such as the quantum cloning and the coherent swap. The general structure of these measurement schemes is the following. An input quantum state  $\rho$  is

processed by the measurement apparatus. As output, one obtains the postmeasurement quantum state together with the result, formulated as classical information channel.

### 5.3.3.1 Measures

The derivation of optimal measurement protocols in terms of measurement-disturbance relations is supposed not to depend on a specific choice of measures to quantify the measurement error and the disturbance, respectively. While protocols are derived which are optimal for most meaningful measures, it is helpful to use specific measures in order to illustrate the obtained tradeoff relations.

For this purpose, the measurement error  $\delta$  depending on a measurement  $E'$  and a reference POVM  $E$  (see also Sec. 2.2.5.1) is defined by the worst-case total variational distance,

$$\delta(E') := \sup_{\rho} \frac{1}{2} \sum_{i=1}^2 |\operatorname{tr} E'_i \rho - \operatorname{tr} E_i \rho|. \quad (5.24)$$

The disturbance  $\Delta$  induced on the quantum channel  $T_s$  accompanying the measurement process is described by the worst-case trace norm distance, see also Sec. 2.2.3.4,

$$\Delta(T_s) := \frac{1}{2} \sup_{\rho} \|T_s(\rho) - \rho\|_1. \quad (5.25)$$

Using those measures, the measurement-disturbance tradeoff relations can be given in terms of achievable regions parametrized using  $\delta$  and  $\Delta$ .

### 5.3.3.2 Universal Optimal Asymmetric Quantum Cloning

As proven in [245], perfect quantum cloning does not exist. This section describes some approximate protocol, being optimal in the sense that (for the given asymmetry parameter) both clones are as close to the original state as possible according to the restrictions of quantum physics. For fixing one of the outputs, no other protocol can in general obtain a second output closer to the original state than the second output of the universal optimal asymmetric quantum cloning protocol.

This protocol, generally assumed to perform well with respect to measurement-disturbance relations [246], is denoted *universal optimal asymmetric quantum cloning* [247, 248] and allows to approximately clone a single input qubit to two output qubits. *Universal* indicates that the cloner should work equally well irrespective of the input state, while *optimal* underlines that the best approximate cloning is to be considered. Additionally, by introducing an asymmetry, one has the freedom to choose if either of the clones should be closer to the initial state at the expense of the quality of the other clone.

The channel for universal optimal asymmetric quantum cloning of a single input qubit into two output qubits can be written as [248]

$$T_{\text{cl}}(\rho) = (a_2 \mathbb{1} + a_1 \mathbb{F}) \left( \rho \otimes \frac{\mathbb{1}}{2} \right) (a_2 \mathbb{1} + a_1 \mathbb{F}), \quad (5.26)$$

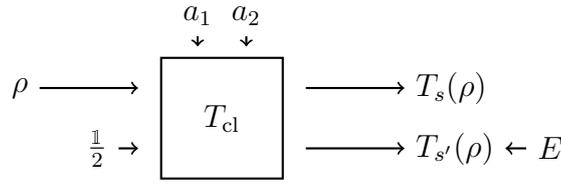


FIGURE 5.5: Universal optimal asymmetric quantum cloning. The quantum state  $\rho$  is approximately cloned by the universal optimal asymmetric quantum cloning device  $T_{\text{cl}}$ . The two parameters  $a_1$  and  $a_2$  are determining the asymmetry between the clones  $T_s(\rho)$  and  $T_{s'}(\rho)$ . As an auxiliary system, the state  $\frac{1}{2}$  is provided. Finally, the ideal measurement  $E$  can be applied to the clone used for the measurement procedure, resulting in classical information, while the other clone is compared to  $\rho$ .

with  $a_1^2 + a_2^2 + a_1 a_2 = 1$ ,  $a_1, a_2 \in \mathbb{R}$ , where  $\mathbb{F}$  swaps two qubits,

$$\mathbb{F} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{5.27}$$

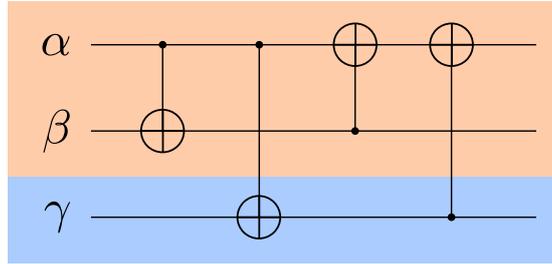


FIGURE 5.6: A quantum circuit model for universal optimal asymmetric quantum cloning, adapted from [247]. An incoming system qubit, denoted by  $\alpha$ , is to be cloned onto  $\beta$ . In order to achieve the cloning between the two (orange) qubits, a third auxiliary system, shown in blue, is needed. Four CNOT gates result in the transformation of basis states as given in Eq. (5.29).

Please note that the (approximate) cloning operation cannot be realized by a 2 qubit unitary transformation. Instead, additional auxiliary systems are needed. As shown in Fig. 5.6, the cloning operation can be achieved by three qubits interacting via in total four controlled-NOT (CNOT) gates [247], where each CNOT gate corresponds to the unitary transformation

$$\text{CNOT} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}, \tag{5.28}$$

as it flips (application of  $\sigma_x$ ) the, say, second qubit if the first qubit is in the excited state. The transformation of the basis states in the computational basis during the cloning operation then reads [247, 249]

$$|\alpha\rangle \otimes |\beta\rangle \otimes |\gamma\rangle \rightarrow |\alpha \oplus \beta \oplus \gamma\rangle \otimes |\alpha \oplus \beta\rangle \otimes |\alpha \oplus \gamma\rangle \quad (5.29)$$

with  $\alpha, \beta, \gamma \in \{0, 1\}$  and  $\oplus$  denoting addition modulo 2. The transformation according to Eq. (5.29) allows to deduce the transformation of arbitrary input states. Depending on the joint state of the second qubit and the third (auxiliary) qubit, the state of the first qubit will remain unchanged or will be transferred to the second qubit. Superposition states thereof can control the asymmetry of the cloning.

However, if one is not interested in the specific implementation using the auxiliary qubit, but only in the two approximate clones, one can describe the cloning transformation channel by Eq. (5.26). For example, consider the asymmetric cloning of  $\rho = |0\rangle\langle 0|$  for different choices of the parameters  $a_1$  and  $a_2$ . For  $a_1 = 0, a_2 = 1$ , the cloning channel acts as the identity channel,

$$T_{\text{cl}}^{(a_1=0, a_2=1)}(|0\rangle\langle 0|) = \mathbb{1} \left( |0\rangle\langle 0| \otimes \frac{\mathbb{1}}{2} \right) \mathbb{1} = |0\rangle\langle 0| \otimes \frac{\mathbb{1}}{2}. \quad (5.30)$$

Thus, the first clone is equal to the input state, whereas the second clone does not contain any information about the input state. If instead  $a_1 = 1, a_2 = 0$ , the cloning results in a swap of the two qubits,

$$T_{\text{cl}}^{(a_1=1, a_2=0)}(|0\rangle\langle 0|) = \mathbb{F} \left( |0\rangle\langle 0| \otimes \frac{\mathbb{1}}{2} \right) \mathbb{F} = \frac{\mathbb{1}}{2} \otimes |0\rangle\langle 0|. \quad (5.31)$$

Hence, the second clone is perfect, while the first is white noise.

The intermediate case, which corresponds to the universal optimal *symmetric* cloner, is attained for  $a_1 = a_2 = \frac{1}{\sqrt{3}}$ . Then,

$$T_{\text{cl}}^{(a_1=a_2=\frac{1}{\sqrt{3}})}(|0\rangle\langle 0|) = \frac{\mathbb{1} + \mathbb{F}}{\sqrt{3}} \left( |0\rangle\langle 0| \otimes \frac{\mathbb{1}}{2} \right) \frac{\mathbb{1} + \mathbb{F}}{\sqrt{3}} = \frac{1}{6} \begin{pmatrix} 4 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad (5.32)$$

which describes an entangled state. Tracing out either of the qubits leads to

$$\tilde{\rho} = \tilde{\rho}' = \frac{1}{6} \begin{pmatrix} 5 & 0 \\ 0 & 1 \end{pmatrix}. \quad (5.33)$$

The fidelity between either of the clones,  $\tilde{\rho}$  and  $\tilde{\rho}'$ , and the initial state is  $\frac{5}{6}$ .

A possible scheme to perform measurements based on the universal optimal asymmetric cloner first prepares two in general different clones. One of them can then be used to evaluate the disturbance, which is only caused by the cloning process itself. The other clone, now independent of the first clone, can be measured strongly in order to determine the measurement error.

### 5.3.3.3 Coherent Swap

A different measurement strategy which will be used as a reference for the performance of measurement schemes is the coherent swap. For a product state  $\varrho_{\text{prod}} = \rho \otimes \tilde{\rho}$ , the transformation leads to

$$\varrho_{\text{prod}} = \rho \otimes \tilde{\rho} \xrightarrow{\mathbb{F}} \chi_{\text{prod}} = \tilde{\rho} \otimes \rho. \quad (5.34)$$

In general, the swap operation  $\mathbb{F}$  as given in Eq. (5.27) swaps the two qubits of state  $\varrho$  with each other,

$$\varrho = \frac{1}{4} \sum_{\mu, \nu=0}^3 T_{\mu, \nu} \sigma_{\mu} \otimes \sigma_{\nu} \longrightarrow \chi = \mathbb{F} \varrho \mathbb{F} = \frac{1}{4} \sum_{\mu, \nu=0}^3 T_{\nu, \mu} \sigma_{\mu} \otimes \sigma_{\nu}. \quad (5.35)$$

If, instead, the swap operation is only partially applied, the system will be in an intermediate state between the original state and the state after swap operation. The transformation of a product state reads

$$T_{\text{cs}}(\rho, \tilde{\rho}) = e^{it\mathbb{F}}(\rho \otimes \tilde{\rho})e^{-it\mathbb{F}} = (a_2\mathbb{1} + ia_1\mathbb{F})(\rho \otimes \tilde{\rho})(a_2\mathbb{1} - ia_1\mathbb{F}), \quad (5.36)$$

where  $t \in [0, \pi/2]$  determines the amplitude of the swap operation. In order to obtain a similar form as in Eq. (5.26), the parameters  $a_1$  and  $a_2$  with  $a_1^2 + a_2^2 = 1$  are introduced.

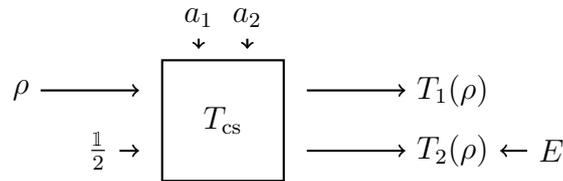


FIGURE 5.7: The coherent swap operation with variable interaction time  $t \in [0, \pi/2]$  encoded by the parameters  $a_1$  and  $a_2$ . The initial quantum state  $\rho$  and the maximally mixed state  $\tilde{\rho} = \mathbb{1}/2$  are coherently swapped, allowing to use one of the operation's marginals for the measurement, while the other marginal provides the distorted quantum state.

To use Eq. (5.36) in the framework of a measurement, one can consider the secondary quantum input  $\tilde{\rho}$  as an auxiliary state, which, as shown in [P10], is optimally chosen to be  $\tilde{\rho} = \mathbb{1}/2$ . Analogously to the quantum cloner, the ideal measurement can be performed on one of the outputs of the coherent swap channel in Eq. (5.36), leading to a channel carrying classical information. The other output can be used to evaluate the measurement induced disturbance. In Sec. 5.3.4, which is a reprint of [P10], the coherent swap and the cloner serve as reference.

### 5.3.3.4 Optimal Instruments

In [234], it has been shown that an optimal tradeoff relation between the measurement error and the disturbance induced by the measurement process is attainable by a set of quantum

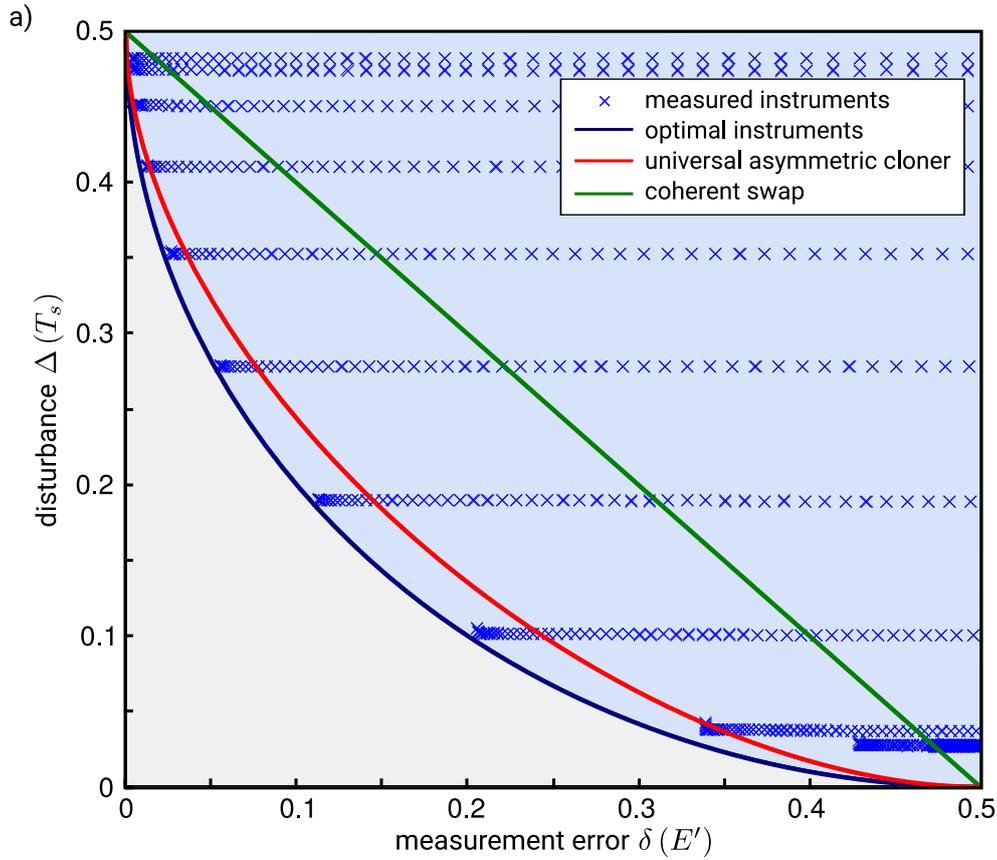


FIGURE 5.8: The optimal instruments with Kraus operators given in Eqs. (5.39) result in the measurement-disturbance tradeoff relation of Eq. (5.40), shown as the dark blue quarter-circle. It clearly outperforms the measurement-disturbance relation of the cloning protocol (red) as well as of another measurement protocol, the coherent swap (green).

devices parametrized by only two parameters. The optimal tradeoff is then obtained for quantum channels of the form

$$T(\rho) \equiv \sum_{i=1}^d \left[ z \langle i | \rho | i \rangle \frac{\mathbb{1}_d - |i\rangle\langle i|}{d-1} + (1-z) K_i \rho K_i \right] \otimes |i\rangle\langle i| \quad (5.37)$$

with  $K_i = \mu \mathbb{1}_d + \nu |i\rangle\langle i|$  and  $z \in [0, 1]$  and  $\mu, \nu \in \mathbb{R}$ . The parameters  $\mu$  and  $\nu$  are not independent from each other as the channel has to be trace preserving. Although the theoretical analysis is valid for various measures quantifying measurement error and disturbance and for different system dimensions, this thesis focuses on qubit measurements only. Hence, the discussion is restricted to  $d = 2$ , leading to the general channel

$$T(\rho) \equiv [z \langle 1 | \rho | 1 \rangle |2\rangle\langle 2| + (1-z) K_1 \rho K_1] \otimes |1\rangle\langle 1| + [z \langle 2 | \rho | 2 \rangle |1\rangle\langle 1| + (1-z) K_2 \rho K_2] \otimes |2\rangle\langle 2| \quad (5.38)$$

with  $K_1 = (\mu + \nu) |1\rangle\langle 1| + \mu |2\rangle\langle 2| = \nu |1\rangle\langle 1| + \mu \mathbb{1}$  and  $K_2 = \mu |1\rangle\langle 1| + (\mu + \nu) |2\rangle\langle 2| = \nu |2\rangle\langle 2| + \mu \mathbb{1}$  with the condition  $\mu + \nu = \pm 1$ .

According to the parametrization used in [P10], the optimal instruments are then given by

$$K_1 = \frac{1}{\sqrt{2}} \left( \sqrt{1-\gamma} |1\rangle\langle 1| + \sqrt{1+\gamma} |2\rangle\langle 2| \right), \quad (5.39a)$$

$$K_2 = \frac{1}{\sqrt{2}} \left( \sqrt{1+\gamma} |1\rangle\langle 1| + \sqrt{1-\gamma} |2\rangle\langle 2| \right), \quad (5.39b)$$

leading to the optimal tradeoff relation

$$\Delta \geq \begin{cases} \frac{1}{2} \left( \sqrt{1-\delta} - \sqrt{\delta} \right)^2 & \text{if } \delta \leq \frac{1}{2}, \\ 0 & \text{if } \delta \geq \frac{1}{2}. \end{cases} \quad (5.40)$$

The optimal tradeoff relation of Eq. (5.40), achieved by the optimal instruments with the Kraus operators of Eqs. (5.39), outperforms the tradeoff relation of the universal optimal asymmetric cloning, as best visualized in Fig. 5.8. In the following section, the publication *Procedures for the Optimal Measurement-Disturbance Tradeoff* [P10] is preprinted, in which those optimal instruments are derived, implemented, and compared against the measurement procedures of quantum cloning and coherent swap as introduced above.

This publication originated from a discussion of Anna-Lena Hashagen, Jan Dziewior and myself and is based on theoretical work of Anna-Lena Hashagen and Michael Wolf. Jan Dziewior adapted with help from me the theoretical approach to an experimental setup which was conceived mainly by myself with help of Jan Dziewior. In close collaboration, we refined the theory and adapted the experiment. We conducted the experimental measurements and evaluated the results together. The manuscript was prepared and edited by all authors.

## Procedures for the Optimal Measurement-Disturbance Tradeoff

Lukas Knips,<sup>1,2,3</sup> Jan Dziewior,<sup>1,2,3</sup> Anna-Lena K. Hashagen,<sup>3,4</sup> Jasmin  
D. A. Meinecke,<sup>1,2,3</sup> Harald Weinfurter,<sup>1,2,3</sup> and Michael M. Wolf<sup>3,4</sup>

<sup>1</sup>Max-Planck-Institut für Quantenoptik, Hans-Kopfermann-Straße 1, 85748 Garching, Germany

<sup>2</sup>Department für Physik, Ludwig-Maximilians-Universität, 80797 München, Germany

<sup>3</sup>Munich Center for Quantum Science and Technology (MCQST), Schellingstr. 4, D-80799 München

<sup>4</sup>Fakultät für Mathematik, Technische Universität München, Germany

One of the characteristic features of quantum mechanics is that every measurement that extracts information about a general quantum system necessarily causes an unavoidable disturbance to the state of this system. A plethora of different approaches has been developed to characterize and optimize this tradeoff for particular measurement scenarios. Yet, the framework of quantum instruments not only allows to investigate the optimal tradeoff, but furthermore to derive and identify the general, optimal procedures themselves. We focus our analysis on binary measurements on qubits as commonly used in communication and computation protocols and demonstrate theoretically and in an experiment that the optimal universal asymmetric quantum cloner, albeit ideal for cloning, can be outperformed with high significance by the optimal procedures derived with the quantum instrument framework.

*Introduction.*—The work of Heisenberg, best visualized by the Heisenberg microscope [1], teaches us that every measurement is accompanied by a fundamental disturbance of a quantum system. The question about the precise relation between the information gained about the quantum system and the resulting disturbance has since inspired numerous studies [2–19]. A central problem is to find a tight, quantitative tradeoff relation, e.g., for the maximally achievable information for a given disturbance or, vice versa, for the minimal disturbance for a certain amount of extracted information. Obviously, this is not only relevant for quantum foundations, but also for many applications in quantum communication [20, 21] and quantum computation [22–24]. Initially studied in the context of which-path information and loss of visibility in interferometers [2, 3], quantifying the information-disturbance tradeoff was based on various measures such as the traditional root mean squared distance [4, 5], the distance of probability distributions [6], operation and estimation fidelities [7–9], entropic quantities [8–13], reversibility [13–15], stabilized operator norms [16, 17], state discrimination probability [10], probability distribution fidelity [18], and Fisher information [19]. While the above approaches are focused on finding quantitative tradeoff relations for a given set of measures and a particular measurement procedure, they do not give a clear candidate for the most fundamental one.

We propose to go one step further and to consider the structure of procedures itself asking the question what makes a particular quantum measurement procedure optimal, i.e., fulfill a tradeoff relation optimally. For this purpose, we build upon a novel, comprehensive analysis of measurement processes and their information-disturbance relations by two of us [25]. There, not only optimal tradeoff relations have been found, but optimal measurement devices have been identified and proven to be independent of the chosen quality measures, as

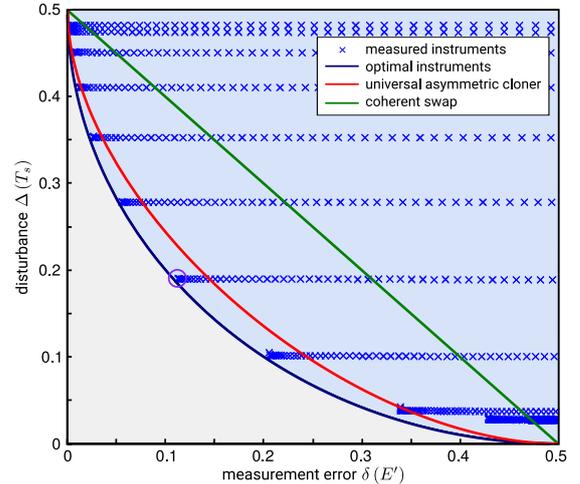


FIG. 1. Error-disturbance dependencies for various measurement schemes. Our optimal quantum instruments clearly outperform the optimal asymmetric cloner (red) and the coherent swap operation (green). Our measurements (blue crosses) come close to the theoretical curve (blue curve). The violet marked instrument is discussed in Fig. 5 in more detail. The error bars are too small to be visible; for a detailed discussion see Appendix VIII.

long as these fulfill some reasonable assumptions, such as convexity and basis-independence. In addition, our approach is also unique with respect to employing reference observables. On one hand, since information eventually is obtained via measurements of observables, we base the quantification of the measurement error on a reference observable. On the other hand, in order not to restrict the further usage of the post-measurement state, the measurement induced disturbance is defined

without relying on any reference observable. For a finite-dimensional von Neumann measurement, our approach gives the whole class of quantum instruments which achieve the optimal tradeoff.

In this letter, we describe how optimal instruments can be derived for typical measures of measurement error, i.e., inverse information, and state disturbance and how they can be implemented in an experiment. Furthermore, we demonstrate the practical advantage of our experimentally implemented optimal procedures by comparing them with commonly employed schemes, namely the coherent swap and the universal asymmetric quantum cloner. We experimentally test the tradeoff relation using a tunable Mach-Zehnder-Interferometer which allows to implement a large range of quantum instruments applied to a two-dimensional quantum system encoded in the polarization of photons. Evaluating the relation between the error of the measurement and the disturbance of the qubit state clearly shows the superiority of the optimal instruments derived here.

*Measurements as quantum instruments.*—To generally quantify both the measurement error and the measurement induced disturbance, we describe the measurement of observables on a quantum system by means of quantum instruments [26, 27] as illustrated in Fig. 2. Formally, a quantum instrument  $I$  is defined as a set of completely positive linear maps  $I := \{I_j\}_{j=1}^m$  that fulfill the normalization condition  $\sum_{j=1}^m I_j^*(\mathbb{1}) = \mathbb{1}$ , where  $I_j^*$  denotes the dual map to  $I_j$  with respect to the Hilbert-Schmidt inner product. This description naturally encompasses the connection between the observable given by a positive operator valued measure (POVM)  $E' := \{E'_j\}_{j=1}^m$  and the quantum channel  $T_s$ , which describes the measurement induced change of the state.

In general, a quantum channel is a completely positive trace preserving linear map. In the context of quantum instruments, the channel is given by the sum of the linear maps with  $T_s := \sum_{j=1}^m I_j$ , where each map corresponds to one measurement operator  $E'_j$  of the POVM. The normalization condition of the quantum instrument ensures that the corresponding quantum channel is trace-preserving. Expressing the channel in terms of  $I$  as above reflects the decohering effect of the measurement on the quantum state of the measured system.

The measurement operators  $\{E'_j\}_{j=1}^m$  themselves are fully determined by  $I$  via  $E'_j := I_j^*(\mathbb{1})$ , where the probability distribution for outcomes  $\{j\}_{j=1}^m$  on state  $\rho$  is given by  $\text{tr}(I_j(\rho)) = \text{tr}(I_j(\rho)\mathbb{1}) = \text{tr}(\rho I_j^*(\mathbb{1})) = \text{tr}(\rho E'_j)$ . From this point of view, the normalization condition of the quantum instrument ensures that the distribution  $\{\text{tr}(E'_j \rho)\}_{j=1}^m$  is normalized. The instrument description based on the normalized set of maps  $I$ , which implies the pair  $(E', T_s)$ , is sufficient to exhaustively describe all possible quantum measurement processes.

*Distance measures.*—From the notion of quantum instruments it becomes immediately clear that  $E'$  and  $T_s$  are not independent, i.e., the change of the state has a fundamental dependence on the information gained and

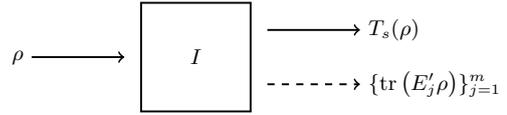


FIG. 2. General description of a measurement using a quantum instrument  $I$ . Obtaining information about the quantum state via the POVM  $E'$  (dashed, classical output) induces a change of the quantum state described by the quantum channel  $T_s$  (solid, quantum output).

vice versa. For a thorough quantitative analysis of this measurement-disturbance tradeoff, we use distance measures to assess the quality of the approximate measurement and to quantify the disturbance. The optimality of the instruments is independent of the chosen measures as long as they fulfill the above mentioned reasonable assumptions (see also Appendix IV). For the evaluation of instruments we choose some of the measures recommended in [16], i.e., the worst-case total variational distance and the worst-case trace norm distance. For other measures, see Appendix V. We quantify the disturbance  $\Delta$  caused to the system by the deviation of the channel  $T_s$  from the identity channel  $T_{\text{id}}(\rho) := \rho$ . The measurement error  $\delta$  quantifies the deviation of the measurement  $E'$  from a reference measurement  $E$ . This approach utilizes a reference POVM  $E$  to quantify the measurement error, but not the disturbance, in contrast to all other approaches found in the literature, where either a reference system is used for both, measurement error and disturbance, or none is used at all.

The measurement error  $\delta$  can be quantified by defining a worst-case total variational distance based on the  $l_1$ -distance between probability distributions. The  $l_1$ -distance, also called total variational distance, displays the largest possible difference between the probabilities that two probability distributions assign to the same event and therefore is the relevant distance measure for hypothesis testing [27, 28]. In our case, these two probability distributions stem from the reference measurement  $E$  and the actual measurement  $E'$  for some quantum state. To generalize the measure for the measurement error to take into account all possible quantum states  $\rho$  of the system we additionally take the worst case with respect to all states, which is natural when considering the maximal difference, i.e., worst-case characteristic of the  $l_1$ -distance itself. Thus our worst-case total variational distance is defined as

$$\delta(E') := \sup_{\rho} \frac{1}{2} \sum_{i=1}^2 |\text{tr}(E'_i \rho) - \text{tr}(E_i \rho)|. \quad (1)$$

The quantum analogue of the worst-case total variational distance is the worst-case trace norm distance, which we thus use to quantify the distance between the quantum channel  $T_s$  and the identity channel  $T_{\text{id}}$ ,

$$\Delta(T_s) := \frac{1}{2} \sup_{\rho} \|T_s(\rho) - \rho\|_1. \quad (2)$$

This disturbance measure quantifies how well the quantum channel  $T_s$  can be distinguished from the identity channel  $T_{\text{id}}$  in a statistical experiment, if no auxiliary systems are allowed [29].

*Optimal instruments and tradeoff.*—As reference measurement, we choose the ideal projective measurement of the qubit with  $E = \{|j\rangle\langle j|\}_{j=1}^2$ . As proven in [25] for the optimal quantum instruments each element  $I_j$  can be expressed by a single Kraus operator, agreeing with the intuition that additional Kraus operators introduce noise to the system. For a qubit this leads to

$$T_s(\rho) = \sum_{j=1}^2 K_j \rho K_j^\dagger \quad \text{and} \quad \{E'_j = K_j^\dagger K_j\}_{j=1}^2. \quad (3)$$

The Kraus operators of an optimal instrument can be chosen diagonal in the basis  $\{|j\rangle\}_{j=1}^2$  given by the reference measurement [25]. Since for a qubit there are only two of them and they must satisfy the normalization condition, in general their form is

$$K_1 = \sqrt{1 - b_2^2} |1\rangle\langle 1| + e^{i\beta_1} b_1 |2\rangle\langle 2|, \quad (4a)$$

$$K_2 = b_2 |1\rangle\langle 1| + e^{i\beta_2} \sqrt{1 - b_1^2} |2\rangle\langle 2|, \quad (4b)$$

with  $0 \leq b_1^2, b_2^2 \leq 1$  and two arbitrary phases  $\beta_1$  and  $\beta_2$ .

As proven in Appendix I, for such an instrument, the worst-case total variational distance  $\delta$  and its trace-norm analogue  $\Delta$ , Eqs. (1,2), quantifying measurement error and disturbance respectively, satisfy

$$\Delta \geq \begin{cases} \frac{1}{2} \left( \sqrt{1 - \delta} - \sqrt{\delta} \right)^2 & \text{if } \delta \leq \frac{1}{2}, \\ 0 & \text{if } \delta > \frac{1}{2}. \end{cases} \quad (5)$$

The inequality is tight and cannot be exceeded by any quantum measurement procedure. The family of optimal instruments, which attain equality in Eq. (5), is found to be

$$K_1 = \frac{1}{\sqrt{2}} \left( \sqrt{1 - \gamma} |1\rangle\langle 1| + \sqrt{1 + \gamma} |2\rangle\langle 2| \right), \quad (6a)$$

$$K_2 = \frac{1}{\sqrt{2}} \left( \sqrt{1 + \gamma} |1\rangle\langle 1| + \sqrt{1 - \gamma} |2\rangle\langle 2| \right), \quad (6b)$$

with  $\gamma \in [0, 1]$ , leading to  $\delta(\gamma) = (1 - \gamma)/2$ .

*Comparison with other schemes.*—For comparison with the optimal procedures derived here, we have chosen coherent swap and quantum cloning operations. Cloning has been already analyzed in connection with quantum measurements [30] and is discussed in the context of joint measurements of observables [31, 32]. For perfect cloning, there would be no measurement-disturbance tradeoff, as one of the perfect clones could be measured without error with the other clone staying undisturbed. The quantum channel  $T_s(\rho) = \text{tr}_{s'}(T_{\text{clo}}(\rho))$ , a marginal of the cloning channel  $T_{\text{clo}}$ , corresponds to the evolution of the system state, obtained when tracing out the second (primed) clone. The channel of the second clone,

$T_{s'}(\rho) = \text{tr}_s(T_{\text{clo}}(\rho))$ , provides an approximate copy to which the reference POVM  $E$  is applied, see Fig. 3. Asymmetry within the quality of the clones determines the tradeoff between the measurement error and the disturbance.

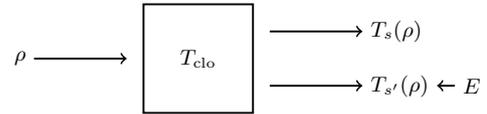


FIG. 3. Universal asymmetric quantum cloning. The initial quantum state  $\rho$  is asymmetrically, approximately cloned to the auxiliary system, initially in state  $\mathbb{1}/2$ . The reference measurement is performed on one of the clones, while the other is compared to the initial quantum state  $\rho$ .

The optimal universal asymmetric quantum cloning channel  $T_{\text{clo}}$  for any initial quantum state  $\rho$  reads [33]

$$T_{\text{clo}}(\rho) = (a_2 \mathbb{1} + a_1 \mathbb{F}) \left( \rho \otimes \frac{\mathbb{1}}{2} \right) (a_2 \mathbb{1} + a_1 \mathbb{F}), \quad (7)$$

with  $a_1^2 + a_2^2 + a_1 a_2 = 1$ ,  $a_1, a_2 \in \mathbb{R}$ , and the flip (or swap) operator  $\mathbb{F} := \sum_{i,j=1}^2 |ji\rangle\langle ij|$ . The parameter  $a_1$  determines the amplitude of a swap operation between both qubits.

With our measures, the measurement-disturbance tradeoff for the asymmetric quantum cloning channel satisfies  $\Delta = \frac{1}{4} \left( \sqrt{2 - 3\delta} - \sqrt{\delta} \right)^2$  for  $\delta \leq \frac{1}{2}$  (and 0 otherwise) with  $\delta(a_2) = a_2^2/2$ , see Appendix II.

As the cloning operation cannot be realized by a unitary two-qubit transformation, any real implementation is embedded in a larger system. Let us thus consider an obvious analogue to the cloning operation, which can be realized by the unitary two-qubit swap operation. For this channel,  $T_{\text{cs}}$ , the system interacts with the auxiliary system via a Heisenberg Hamiltonian as

$$T_{\text{cs}}(\rho) = e^{it\mathbb{F}} (\rho \otimes \tilde{\rho}) e^{-it\mathbb{F}} = (a_2 \mathbb{1} + ia_1 \mathbb{F}) (\rho \otimes \tilde{\rho}) (a_2 \mathbb{1} - ia_1 \mathbb{F}), \quad (8)$$

with  $t \in [0, \pi/2]$  or using a parametrization analogous to the cloning scheme with  $a_1^2 + a_2^2 = 1$ ,  $a_1, a_2 \in \mathbb{R}$ . The extreme cases are no swap ( $t = 0$ ,  $a_2 = 1$ ) and full swap ( $t = \pi/2$ ,  $a_1 = 1$ ).

The  $\delta$ - $\Delta$ -tradeoff for the coherent swap considering the reference measurement  $E = \{|j\rangle\langle j|\}_{j=1}^2$  performed on one of the outputs satisfies  $\Delta = \frac{1}{2} - \delta$  with  $\delta(a_1) = (1 - a_1^2)/2$ , see Appendix III. This is evidently also inferior to our optimal instruments, Eq. (6), with the tradeoff given in Eq. (5).

*Experimental implementation.*—For our experimental evaluation of the procedures for the optimal measurement-disturbance tradeoff we want to realize a broad range of quantum instruments including the optimal ones. For that purpose we consider the polarization degree of freedom of photons to encode  $\rho$ , with  $|1\rangle \leftrightarrow |H\rangle$

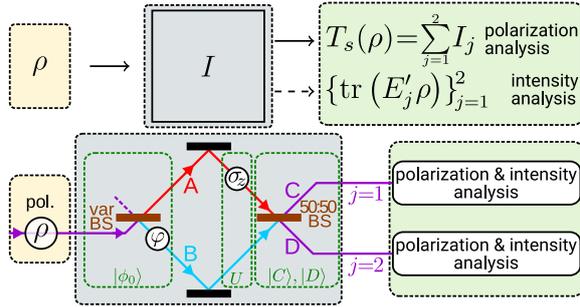


FIG. 4. Theoretical description and conceptual setup. The state  $\rho$  is encoded in the polarization degree of freedom of a photon, which is sent to a variable beam splitter (var BS). The spatial superposition state  $|\phi_0\rangle$  can be tuned in terms of relative intensities and phase. For the interaction  $U$  between the path and the polarization degrees of freedom we apply a  $\sigma_z$  operation to the polarization in one path. Projections onto the output ports  $|C\rangle$  and  $|D\rangle$  of a balanced 50:50 beam splitter conclude the realization of the Kraus operators as given in Eqs. (10). Polarization and intensity measurements are performed at the output ports of the interferometer. The actual experiment, while equivalent to the shown setup, is structured differently such that the polarization state  $\rho$  is created inside of the interferometer, see Appendix VI.

and  $|2\rangle \leftrightarrow |V\rangle$ , where  $|H\rangle$  ( $|V\rangle$ ) denotes horizontally (vertically) polarized light. The Kraus operators describing the chosen set of instruments are thus given by

$$K_{1,2} = \frac{1}{\sqrt{2}} \left[ \sqrt{1 \pm \gamma} |H\rangle\langle H| + e^{i\beta} \sqrt{1 \mp \gamma} |V\rangle\langle V| \right] \quad (9)$$

with an arbitrary phase  $\beta$ . The optimal cases Eqs. (6) are achieved for  $\beta = 0$ .

To experimentally realize a quantum instrument and to enable analysis of the two outputs  $T_s$  and  $E'$ , it is necessary to employ an additional auxiliary quantum system, which is not yet explicitly present in the instrument description of Fig. 2. For the measurement of photon polarization a natural candidate is the path degree of freedom of the photons. Since in our case a two dimensional auxiliary system is sufficient, we employ a Mach-Zehnder interferometer providing the path states  $|A\rangle$  and  $|B\rangle$ , see Fig. 4. The instruments are then determined by the initial state of this auxiliary system,  $|\phi_0\rangle = \cos\alpha|A\rangle + e^{i\varphi}\sin\alpha|B\rangle$ , the measurement performed on it, i.e., the detection in the output path states  $|C\rangle$  and  $|D\rangle$ , as well as by an intermediate interaction between path and polarization. The interaction is given by a unitary evolution  $U$ , which exchanges information between the systems. We use  $U = i\sigma_z \otimes |A\rangle\langle A| + \mathbb{1} \otimes |B\rangle\langle B|$ , which introduces a polarization dependent phase shift in arm  $|A\rangle$  with  $\sigma_z = |H\rangle\langle H| - |V\rangle\langle V|$ .

For an initial path state  $|\phi_0\rangle$  the Kraus operators,

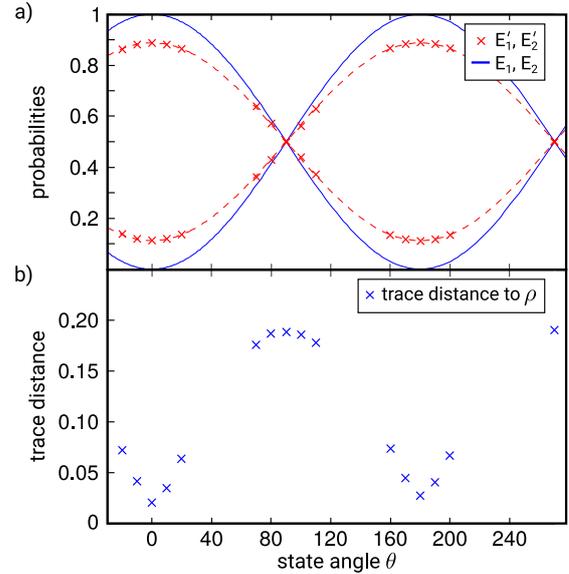


FIG. 5. Evaluating measurement error  $\delta$  and disturbance  $\Delta$ . a) The measurement error corresponds to the maximal distance between the outcomes of the actual measurements  $E'_1$  and  $E'_2$  (red crosses) and the outcomes of the reference measurements  $E_1$  and  $E_2$  (blue line). b) The disturbance is obtained by taking the supremum of the trace distance between the prepared polarization states and the tomographically reconstructed states of  $T_s$ . Please note that the suprema in a) and b) are achieved for different states. Statistical error bars are negligibly small. For a detailed discussion, see Appendix VIII.

which act on the polarization, can then be obtained as

$$K_1 = \text{tr}_{\text{path}} [(\mathbb{1} \otimes |C\rangle\langle C|) U (\mathbb{1} \otimes |\phi_0\rangle\langle\phi_0|)], \quad (10a)$$

$$K_2 = \text{tr}_{\text{path}} [(\mathbb{1} \otimes |D\rangle\langle D|) U (\mathbb{1} \otimes |\phi_0\rangle\langle\phi_0|)]. \quad (10b)$$

Relating these expressions with Eq. (9), the parameters  $\gamma$  and  $\beta$  are given by the experimental parameters  $\alpha$  and  $\varphi$  by  $\gamma = \sin(2\alpha)\sin\varphi$  and  $\beta = \arctan[\tan(2\alpha)\cos\varphi]$ . The outcome of the measurement  $E'$  is then obtained by determining the total intensity in the output  $C$  ( $E'_1$ ) and  $D$  ( $E'_2$ ), respectively, the action of the quantum channel  $T_s$  by polarization state tomography.

*Measurements and results.*—According to Eqs. (1) and (2), the measures  $\delta$  and  $\Delta$  use the supremum over different input states  $\rho$ . To obtain data for a specific instrument described by the Kraus operators of Eq. (9), we fix the experimental parameters  $\alpha$  and  $\varphi$  (i.e., the instrument parameters  $\gamma$  and  $\beta$ ) and vary the angle  $\theta$  defining the state  $\rho = |\psi\rangle\langle\psi|$  with  $|\psi\rangle = \cos\frac{\theta}{2}|H\rangle + \sin\frac{\theta}{2}|V\rangle$ , see Appendix VII. Fig. 5 shows the evaluation for one specific instrument. The measurement error  $\delta$  is obtained from the maximal deviation of the measurement (red crosses) to the reference measurement (blue solid line), see Eq. (1). Similarly, we obtain the disturbance  $\Delta$  from the maximum trace distance when varying  $\theta$ .

The obtained values for measurement error and state disturbance are shown in Fig. 1 for the set of experimentally prepared quantum instruments. Each data point here identifies one quantum instrument, for which the supremum of the prepared quantum states in terms of measurement error and disturbance is determined. We could show that there exist quantum instruments, also experimentally accessible, which significantly outperform the optimal universal asymmetric cloner and the coherent swap operation in terms of the considered distances.

*Conclusion.*—We applied the novel approach derived in [25] to the setting of binary qubit measurements to retrieve quantum instruments achieving an optimal measurement-disturbance tradeoff. In this setting a reference measurement is used to quantitatively obtain the measurement error. The disturbance, on the other hand, does not depend on any reference measurement, but solely on comparing the state before and after the measurement. Our protocol is tailored for applications based on a specific measurement without restricting subsequent use of the post-measurement state.

Furthermore, we have demonstrated that the strategies of optimal universal asymmetric quantum cloning and coherent swap do not perform optimally when considering the tradeoff relation between measurement error and disturbance. Those protocols are optimal for their re-

spective purposes such as approximate quantum cloning, but cannot compete with the optimal instruments in the measurement scenario as in general they result in worse measurement-disturbance tradeoff relations. This is in compelling agreement with less general findings, e.g., in [31, 32]. We have shown that the advantage of optimal instruments over other schemes is not only a mere theoretical improvement, but experimentally accessible. In future applications our findings allow to identify the optimal procedures which retrieve information at the physically lowest cost in terms of state disturbance.

#### ACKNOWLEDGMENTS

We thank Jonas Goeser and Katharina Senkalla for stimulating discussions. This research was supported in part by the National Science Foundation under Grant No. NSF PHY11-25915, by the DFG (German Research Foundation) Beethoven 2 Project No. WE2541/7-1, and by the DFG under Germanys Excellence Strategy EXC-2111 390814868. LK and AKH are supported by the PhD program *Exploring Quantum Matter* of the Elite Network of Bavaria. JD acknowledges support by the International Max-Planck Research Program for Quantum Science and Technology (IMPRS-QST). JDMA is supported by an LMU research fellowship.

- 
- [1] Werner Heisenberg, *The Physical Principles of the Quantum Theory* (University of Chicago Press, 1930).
  - [2] Gregg Jaeger, Abner Shimony, and Lev Vaidman, “Two interferometric complementarities,” *Phys. Rev. A* **51**, 54–67 (1995).
  - [3] Berthold-Georg Englert, “Fringe Visibility and Which-Way Information: An Inequality,” *Phys. Rev. Lett.* **77**, 2154–2157 (1996).
  - [4] Masanao Ozawa, “Universally valid reformulation of the Heisenberg uncertainty principle on noise and disturbance in measurement,” *Phys. Rev. A* **67**, 042105 (2003).
  - [5] Cyril Branciard, “Error-tradeoff and error-disturbance relations for incompatible quantum measurements,” *Proceedings of the National Academy of Sciences* **110**, 6742–6747 (2013).
  - [6] Paul Busch, Pekka Lahti, and Reinhard F. Werner, “Proof of Heisenberg’s Error-Disturbance Relation,” *Phys. Rev. Lett.* **111**, 160405 (2013).
  - [7] Konrad Banaszek, “Fidelity Balance in Quantum Operations,” *Phys. Rev. Lett.* **86**, 1366–1369 (2001).
  - [8] Christopher A. Fuchs and Asher Peres, “Quantum-state disturbance versus information gain: Uncertainty relations for quantum information,” *Phys. Rev. A* **53**, 2038–2045 (1996).
  - [9] Lorenzo Maccone, “Information-disturbance tradeoff in quantum measurements,” *Phys. Rev. A* **73**, 042307 (2006).
  - [10] Francesco Buscemi, Masahito Hayashi, and Michał Horodecki, “Global Information Balance in Quantum Measurements,” *Phys. Rev. Lett.* **100**, 210504 (2008).
  - [11] Francesco Buscemi, Michael J. W. Hall, Masanao Ozawa, and Mark M. Wilde, “Noise and Disturbance in Quantum Measurements: An Information-Theoretic Approach,” *Phys. Rev. Lett.* **112**, 050401 (2014).
  - [12] Jun Zhang, Yang Zhang, and Chang-Shui Yu, “The Measurement-Disturbance Relation and the Disturbance Trade-off Relation in Terms of Relative Entropy,” *International Journal of Theoretical Physics*, **55** (2016).
  - [13] Giacomo M. D’Ariano, “On the Heisenberg principle, namely on the information-disturbance trade-off in a quantum measurement,” *Fortschritte der Physik* **51**, 318–330 (2003).
  - [14] Andrew N. Jordan and Alexander N. Korotkov, “Uncollapsing the wavefunction by undoing quantum measurements,” *Contemporary Physics* **51**, 125–147 (2010).
  - [15] Yong Wook Cheong and Seung-Woo Lee, “Balance Between Information Gain and Reversibility in Weak Measurement,” *Phys. Rev. Lett.* **109**, 150402 (2012).
  - [16] Alexei Gilchrist, Nathan K. Langford, and Michael A. Nielsen, “Distance measures to compare real and ideal quantum processes,” *Phys. Rev. A* **71**, 062310 (2005).
  - [17] Dennis Kretschmann, Dirk Schlingemann, and Reinhard F. Werner, “The Information-Disturbance Trade-off and the Continuity of Stinespring’s Representation,” *IEEE Transactions on Information Theory* **54**, 1708–1717 (2008).
  - [18] Longfei Fan, Wenchao Ge, Hyunchul Nha, and M. Suhail Zubairy, “Trade-off between information gain and fidelity under weak measurements,” *Phys. Rev. A* **92**, 022114 (2015).

- [19] Tomohiro Shitara, Yui Kuramochi, and Masahito Ueda, “Trade-off relation between information and disturbance in quantum measurement,” *Phys. Rev. A* **93**, 032134 (2016).
- [20] Nicolas Gisin, Grégoire Ribordy, Wolfgang Tittel, and Hugo Zbinden, “Quantum cryptography,” *Reviews of Modern Physics* **74**, 145–195 (2002).
- [21] Jian-Wei Pan, Zeng-Bing Chen, Chao-Yang Lu, Harald Weinfurter, Anton Zeilinger, and Marek Żukowski, “Multiphoton entanglement and interferometry,” *Reviews of Modern Physics* **84**, 777–838 (2012).
- [22] Artur Ekert and Richard Jozsa, “Quantum computation and Shor’s factoring algorithm,” *Reviews of Modern Physics* **68**, 733–753 (1996).
- [23] Vlatko Vedral and Martin B. Plenio, “Basics of quantum computation,” *Progress in Quantum Electronics* **22**, 1–39 (1998).
- [24] Andrew Steane, “Quantum computing,” *Reports on Progress in Physics* **61**, 117–173 (1998).
- [25] Anna-Lena K. Hashagen and Michael M. Wolf, “Universality and optimality in the information–disturbance tradeoff,” *Annales Henri Poincaré* **20**, 219–258 (2018).
- [26] E. Brian Davies and John T. Lewis, “An operational approach to quantum probability,” *Communications in Mathematical Physics* **17**, 239–260 (1970).
- [27] John Watrous, *The Theory of Quantum Information* (Cambridge University Press, 2018).
- [28] Jerzy Neyman and Egon S. Pearson, “On the Problem of the Most Efficient Tests of Statistical Hypotheses,” *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* **231**, 289–337 (1933).
- [29] Allowing auxiliary systems, the relevant disturbance measure is the diamond norm,  $\Delta_{\diamond}(T_s) := \frac{1}{2} \sup_{\xi} \|((T_s - T_{id,d}) \otimes T_{id,d})(\xi)\|_1$ , where the state  $\xi$  includes auxiliary systems and  $\|A\|_1 := \text{tr}(\sqrt{A^\dagger A})$ . Here, for the optimal tradeoff curve, the trace norm turns out to be equal to the diamond norm distance [25].
- [30] Heng Fan, Yi-Nan Wang, Li Jing, Jie-Dong Yue, Handuo Shi, Yong-Liang Zhang, and Liang-Zhu Mu, “Quantum cloning machines and the applications,” *Physics Reports* **544**, 241–322 (2014).
- [31] Thomas Brougham, Erika Andersson, and Stephen M. Barnett, “Cloning and joint measurements of incompatible components of spin,” *Physical Review A* **73**, 062319 (2006).
- [32] G. M. D’Ariano, C. Macchiavello, and M. F. Sacchi, “Joint measurements via quantum cloning,” *Journal of Optics B: Quantum and Semiclassical Optics* **3**, 44–50 (2001).
- [33] Anna-Lena K. Hashagen, “Universal Asymmetric Quantum Cloning Revisited,” *Quant. Inf. Comp.* **17**, 0747–0778 (2017).

## APPENDIX

## I. OPTIMAL TRADEOFF RELATION

**Theorem 1** (Total variation - trace norm tradeoff). *Consider a von Neumann reference measurement given by an orthonormal basis  $\{|i\rangle \in \mathbb{C}^2\}_{i=1}^2$ , and an instrument with two corresponding outcomes. Then the worst-case total variational distance  $\delta$  and its trace-norm analogue  $\Delta$ , defined as in Eqs. (1,2), quantifying measurement error and disturbance respectively, satisfy*

$$\Delta \geq \begin{cases} \frac{1}{2} (\sqrt{1-\delta} - \sqrt{\delta})^2 & \text{if } \delta \leq \frac{1}{2}, \\ 0 & \text{if } \delta \geq \frac{1}{2}. \end{cases} \quad (\text{A1})$$

*The inequality is tight and equality is attained within the family of instruments defined by*

$$I_j(\rho) := K_j \rho K_j, \quad j = 1, 2, \quad (\text{A2})$$

with

$$K_{1,2} = \frac{1}{\sqrt{2}} \left( \sqrt{1 \pm \gamma} |1\rangle\langle 1| + \sqrt{1 \mp \gamma} |2\rangle\langle 2| \right) \quad (\text{A3})$$

with  $\gamma \in [0, 1]$ .

*Proof.* In order to derive the information-disturbance tradeoff, we need to solve the following optimization problem:

For  $\gamma \in [0, 1]$

$$\begin{aligned} & \text{minimize} && \Delta \left( T_s = \sum_{j=1}^2 I_j \right) && (\text{A4}) \\ & \text{subject to} && \delta \left( E' = \{I_j^*(\mathbb{1})\}_{j=1}^2 \right) \leq \gamma, \\ & && I_j \text{ is completely positive and} \\ & && \sum_{j=1}^2 I_j^*(\mathbb{1}) = \mathbb{1}, \end{aligned}$$

where the last two constraints ensure that  $I$  is an instrument. As discussed before, we assume that every element of the instrument can be expressed using a single Kraus operator. This agrees well with intuition, because more Kraus operators introduce more noise to the system. Furthermore, we assume that these Kraus operators can be chosen diagonal in the basis of the reference measurement,  $E = \{|j\rangle\langle j|\}_{j=1}^2$ , to reflect the symmetry of the optimization problem. These assumptions simplify the optimization problem significantly. The Kraus operators given in Eq. (4) then yield the following POVM elements of the approximate measurement

$$E'_j = (1 - b_j^2) |j\rangle\langle j| + b_j^2 (\mathbb{1} - |j\rangle\langle j|), \quad (\text{A5})$$

for  $j = 1, 2$ , where  $\bar{j} = 2$  if  $j = 1$  and  $\bar{j} = 1$  if  $j = 2$  with  $0 \leq b_1^2, b_2^2 \leq 1$ . The measurement error is thus given as

$$\begin{aligned} \delta(E') &= \sup_{\rho} \frac{1}{2} \sum_{j=1}^2 |\text{tr}(E'_j \rho) - \langle j | \rho | j \rangle| \\ &= \sup_{\rho} \frac{1}{2} \sum_{j=1}^2 \left| \text{tr} \left( (b_j^2 \mathbb{1} - (b_j^2 + b_{\bar{j}}^2) |j\rangle\langle j|) \rho \right) \right| \\ &= \sup_{\|\psi\|=1} \frac{1}{2} \sum_{j=1}^2 \left| \langle \psi | b_j^2 \mathbb{1} - (b_j^2 + b_{\bar{j}}^2) |j\rangle\langle j| | \psi \rangle \right| \\ &= \frac{1}{2} (b_1^2 + b_2^2), \end{aligned}$$

where the convexity of the  $l_1$ -norm was used. The disturbance follows from direct calculations,

$$\begin{aligned} \Delta(T_1) &= \frac{1}{2} \sup_{\rho} \|T_1(\rho) - \rho\|_1 \\ &= \frac{1}{2} \sup_{\rho} \left\| \sum_{j=1}^2 K_j \rho K_j^\dagger - \rho \right\|_1 \\ &= \frac{1}{2} \left| 1 - e^{i\beta_1} b_1 \sqrt{1 - b_2^2} - e^{i\beta_2} b_2 \sqrt{1 - b_1^2} \right|. \end{aligned}$$

Without loss of generality, we may assume that  $b_1, b_2 \geq 0$  in the optimization problem, such that an optimum is attained for  $\beta_1 = \beta_2 = 0$ . The optimization problem given in Eq. (A4) therefore simplifies:

For  $\gamma \in [0, 1]$

$$\begin{aligned} \text{minimize} \quad & \frac{1}{2} \left( 1 - b_1 \sqrt{1 - b_2^2} - b_2 \sqrt{1 - b_1^2} \right) \quad (\text{A6}) \\ \text{subject to} \quad & \frac{1}{2} (b_1^2 + b_2^2) \leq \frac{1}{2} (1 - \gamma), \\ & 0 \leq b_1, b_2 \leq 1. \end{aligned}$$

The global minimum is achieved at

$$b_1 = b_2 = \begin{cases} \sqrt{\frac{1}{2}} & \gamma \in [-1, 0] \\ \sqrt{\frac{1}{2}} \sqrt{1 - \gamma} & \gamma \in [0, 1] \end{cases}$$

and as stated in Eq. (A1).  $\square$

## II. TRADEOFF RELATION FOR OPTIMAL UNIVERSAL ASYMMETRIC CLONING

**Theorem 2** (Total variation - trace norm tradeoff using optimal universal asymmetric cloning). *Consider a von Neumann measurement given by an orthonormal basis in  $\mathbb{C}^2$  on one of the outputs of the optimal universal  $1 \rightarrow 2$  asymmetric quantum cloning channel. Then the worst-case total variational distance  $\delta$  and its trace-norm analogue  $\Delta$  satisfy*

$$\Delta = \begin{cases} \frac{1}{4} \left( \sqrt{2 - 3\delta} - \sqrt{\delta} \right)^2 & \text{if } \delta \leq \frac{1}{2}, \\ 0 & \text{if } \delta \geq \frac{1}{2}. \end{cases} \quad (\text{A7})$$

*Proof.* The marginals of the optimal cloning channel are given by

$$T_{\text{clo},i}(\rho) = a_i^2 \frac{\mathbb{1}}{2} \text{tr}(\rho) + (1 - a_i^2) \rho, \quad i = 1, 2, \quad (\text{A8})$$

with  $T_{\text{clo},1} = T_s$  and  $T_{\text{clo},2} = T_{s'}$ . The marginal quantum channel  $T_s$  describes the evolution of the quantum state and its distance to the identity channel  $T_{\text{id}}$  then quantifies the disturbance. Similarly, the marginal  $T_{s'}$ , whose output is measured by the reference measurement  $E$ , describes the measurement itself through  $E'_j = T_{s'}^*(E_j)$ . This is illustrated in Fig. 3. This yields for the disturbance

$$\begin{aligned} \Delta(T_s) &:= \frac{1}{2} \sup_{\rho} \|T_s(\rho) - \rho\|_1 \\ &= \frac{1}{2} \sup_{\rho} \left\| a_1^2 \frac{\mathbb{1}}{2} - a_1^2 \rho \right\|_1 \\ &= \frac{a_1^2}{2}. \end{aligned}$$

The measurement error turns out to be

$$\begin{aligned} \delta(E') &:= \sup_{\rho} \frac{1}{2} \sum_{j=1}^2 |\text{tr}(E'_j \rho) - \langle j | \rho | j \rangle| \\ &= \sup_{\rho} \frac{1}{2} \sum_{j=1}^2 |\text{tr}(T_{s'}^*(|j\rangle\langle j|) \rho) - \langle j | \rho | j \rangle| \\ &= \sup_{\rho} \frac{1}{2} \sum_{j=1}^2 |\text{tr}(|j\rangle\langle j| T_{s'}(\rho)) - \langle j | \rho | j \rangle| \\ &= \sup_{\rho} \frac{1}{2} \sum_{j=1}^2 \left| \langle j | a_2^2 \frac{\mathbb{1}}{2} - a_2^2 \rho | j \rangle \right| \\ &= \frac{a_2^2}{2}. \end{aligned}$$

Substituting this into the trace-preserving condition of the optimal universal asymmetric quantum cloning channel, we obtain the theorem 2.  $\square$

## III. TRADEOFF RELATION FOR COHERENT SWAP

**Theorem 3** (Total variation - trace norm tradeoff using the coherent swap). *Consider a von Neumann measurement given by an orthonormal basis in  $\mathbb{C}^2$  on one of the outputs of a coherent swap channel. Then the worst-case total variational distance  $\delta$  and its trace-norm analogue  $\Delta$  satisfy*

$$\Delta = \frac{1}{2} - \delta. \quad (\text{A9})$$

*Proof.* Using the substitution  $a_1 = a$  and  $a_2 = \sqrt{1-a^2}$  with  $a \in [0, 1]$  yields the two marginals of the coherent swap quantum channel,

$$T_s(\rho) = a^2 \tilde{\rho} + (1-a^2)\rho \quad (\text{A10})$$

and

$$T_{s'}(\rho) = (1-a^2)\tilde{\rho} + a^2\rho. \quad (\text{A11})$$

The disturbance is therefore

$$\begin{aligned} \Delta(T_s) &:= \frac{1}{2} \sup_{\rho} \|T_s(\rho) - \rho\|_1 \\ &= \frac{1}{2} a^2 \sup_{\rho} \|\tilde{\rho} - \rho\|_1. \end{aligned}$$

The optimal choice for  $\tilde{\rho}$  should clearly satisfy the points  $(\Delta(T_s) = 0, \delta(E') = 1/2)$  and  $(\Delta(T_s) = 1/2, \delta(E') = 0)$ , where again  $E' = T_{s'}^*(E)$ . For any such choice of  $\tilde{\rho}$  the disturbance thus satisfies  $\Delta(T_s) \geq a^2/2$ . The measurement error turns out to be

$$\begin{aligned} \delta(E') &:= \sup_{\rho} \frac{1}{2} \sum_{j=1}^2 |\text{tr}(E'_j \rho) - \langle j | \rho | j \rangle| \\ &= \sup_{\rho} \frac{1}{2} \sum_{j=1}^2 |\text{tr}(T_{s'}^*(|j\rangle\langle j|)\rho) - \langle j | \rho | j \rangle| \\ &= \sup_{\rho} \frac{1}{2} \sum_{j=1}^2 |\text{tr}(|j\rangle\langle j| T_{s'}(\rho)) - \langle j | \rho | j \rangle| \\ &= (1-a^2) \sup_{\rho} \frac{1}{2} \sum_{j=1}^2 |\langle j | \tilde{\rho} | j \rangle - \langle j | \rho | j \rangle|. \end{aligned}$$

Thus, an optimal choice for  $\tilde{\rho}$  that minimizes the disturbance and the measurement error is  $\tilde{\rho} = \mathbb{1}/2$ . A pure state with the same diagonal entries yields the same measurement error; it would, however, increase the disturbance caused to the system.

The disturbance is then

$$\Delta(T_s) = \frac{a^2}{2},$$

and the measurement error is

$$\delta(E') = \frac{1}{2} (1-a^2).$$

This gives the linear tradeoff curve given in theorem 3.  $\square$

#### IV. PROPERTIES OF DISTANCE MEASURES

The distance measures used throughout this manuscript to quantify the measurement error and the disturbance, denoted by  $\delta$  and  $\Delta$ , satisfy Assumption 1 and Assumption 2 of [25] respectively.

**Lemma 4.**  $\delta$  as defined in Eq. (1) satisfies the following properties:

(a)  $\delta(\{|i\rangle\langle i|\}_{i=1}^2) = 0$ ,

(b)  $\delta$  is convex,

(c)  $\delta$  is permutation invariant, i.e., for every permutation  $\pi$  and any measurement  $M$

$$\delta(\{U_{\pi}^{\dagger} M_{\pi(i)} U_{\pi}\}_{i=1}^2) = \delta(\{M_i\}_{i=1}^2),$$

where  $U_{\pi}$  is the permutation matrix that acts as  $U_{\pi} |i\rangle = |\pi(i)\rangle$ , and

(d)  $\delta$  is invariant under diagonal unitaries, i.e., that for every diagonal unitary  $D$  and any measurement  $M$

$$\delta(\{D^{\dagger} M_i D\}_{i=1}^2) = \delta(\{M_i\}_{i=1}^2).$$

*Proof.* Let  $\delta(M) := \sup_{\rho} \frac{1}{2} \sum_{i=1}^2 |\text{tr}(M_i \rho) - \langle i | \rho | i \rangle|$ . Then

(a)  $\delta(\{|i\rangle\langle i|\}_{i=1}^2) = 0$ , since

$$\delta(\{|i\rangle\langle i|\}_{i=1}^2) = \sup_{\rho} \frac{1}{2} \sum_{i=1}^2 |\langle i | \rho | i \rangle - \langle i | \rho | i \rangle| = 0,$$

(b)  $\delta$  is convex, since for any measurements  $M, M'$  and for all  $\lambda \in [0, 1]$ ,

$$\begin{aligned} &\delta(\lambda M + (1-\lambda)M') \\ &= \sup_{\rho} \frac{1}{2} \sum_{i=1}^2 |\text{tr}((\lambda M_i + (1-\lambda)M'_i)\rho) - \langle i | \rho | i \rangle| \\ &\leq \lambda \sup_{\rho} \frac{1}{2} \sum_{i=1}^2 |\text{tr}(M_i \rho) - \langle i | \rho | i \rangle| \\ &\quad + (1-\lambda) \sup_{\rho} \frac{1}{2} \sum_{i=1}^2 |\text{tr}(M'_i \rho) - \langle i | \rho | i \rangle| \\ &= \lambda \delta(M) + (1-\lambda) \delta(M'), \end{aligned}$$

(c)  $\delta$  is permutation invariant, since for every permutation  $\pi$  and any measurement  $M$

$$\begin{aligned} &\delta(\{U_{\pi}^{\dagger} M_{\pi(i)} U_{\pi}\}_{i=1}^2) \\ &= \sup_{\rho} \frac{1}{2} \sum_{i=1}^2 |\text{tr}(U_{\pi}^{\dagger} M_{\pi(i)} U_{\pi} \rho) - \langle i | \rho | i \rangle| \\ &= \sup_{\rho} \frac{1}{2} \sum_{i=1}^2 |\text{tr}(M_{\pi(i)} \rho) - \langle \pi(i) | \rho | \pi(i) \rangle| \\ &= \sup_{\rho} \frac{1}{2} \sum_{i=1}^2 |\text{tr}(M_i \rho) - \langle i | \rho | i \rangle| \\ &= \delta(\{M_i\}_{i=1}^2), \end{aligned}$$

where  $U_{\pi}$  is the permutation matrix that acts as  $U_{\pi} |i\rangle = |\pi(i)\rangle$ , and

- (d)  $\delta$  is invariant under diagonal unitaries, since for every diagonal unitary  $D$  and any measurement  $M$

$$\begin{aligned} & \delta(\{D^\dagger M_i D\}_{i=1}^2) \\ &= \sup_{\rho} \frac{1}{2} \sum_{i=1}^2 |\text{tr}(D^\dagger M_i D \rho) - \langle i | \rho | i \rangle| \\ &= \sup_{\rho} \frac{1}{2} \sum_{i=1}^2 |\text{tr}(M_i \rho) - \langle i | D^\dagger \rho D | i \rangle| \\ &= \sup_{\rho} \frac{1}{2} \sum_{i=1}^2 |\text{tr}(M_i \rho) - \langle i | \rho | i \rangle| \\ &= \delta(\{M_i\}_{i=1}^2). \end{aligned}$$

□

**Lemma 5.**  $\Delta$  as defined in Eq. (2) satisfies the following properties:

- (a)  $\Delta(T_{\text{id}}) = 0$ ,  
 (b)  $\Delta$  is convex,  
 (c)  $\Delta$  is basis-independent, i.e., for every unitary  $U$  and every quantum channel  $\Phi$

$$\Delta(U\Phi(U^\dagger \cdot U)U^\dagger) = \Delta(\Phi).$$

*Proof.* Let  $\Delta(\Phi) := \frac{1}{2} \sup_{\rho} \|\Phi(\rho) - \rho\|_1$ . Then

- (a)  $\Delta(T_{\text{id}}) = 0$ , since  $\Delta(T_{\text{id}}) = \frac{1}{2} \sup_{\rho} \|\rho - \rho\|_1 = 0$ ,  
 (b)  $\Delta$  is convex, since for any quantum channels  $\Phi, \Phi'$  and for all  $\lambda \in [0, 1]$ ,

$$\begin{aligned} & \Delta(\lambda\Phi + (1-\lambda)\Phi') \\ &= \frac{1}{2} \sup_{\rho} \|(\lambda\Phi + (1-\lambda)\Phi')(\rho) - \rho\|_1 \\ &= \frac{1}{2} \sup_{\rho} \|\lambda(\Phi(\rho) - \rho) + (1-\lambda)(\Phi'(\rho) - \rho)\|_1 \\ &\leq \lambda \frac{1}{2} \sup_{\rho} \|\Phi(\rho) - \rho\|_1 + (1-\lambda) \frac{1}{2} \sup_{\rho} \|\Phi'(\rho) - \rho\|_1 \\ &= \lambda\Delta(\Phi) + (1-\lambda)\Delta(\Phi'), \end{aligned}$$

where we have used properties of a norm and properties of a supremum of a convex functional over a convex set,

- (c)  $\Delta$  is basis-independent, i.e., for every unitary  $U$  and every quantum channel  $\Phi$

$$\begin{aligned} & \Delta(U\Phi(U^\dagger \rho U)U^\dagger) \\ &= \frac{1}{2} \sup_{\rho} \|U\Phi(U^\dagger \rho U)U^\dagger - \rho\|_1 \\ &= \frac{1}{2} \sup_{\rho} \|U\Phi(\rho)U^\dagger - U\rho U^\dagger\|_1 \\ &= \frac{1}{2} \sup_{\rho} \|\Phi(\rho) - \rho\|_1 \\ &= \Delta(\Phi), \end{aligned}$$

where we have used the fact that the trace norm is unitarily invariant. □

## V. DIFFERENT MEASURES

The optimal instruments as explained in the main text and derived in I result in optimal measurement-disturbance relations for all distance measures which satisfy the assumptions of [25]. For more details on the distance measure used in the main text, see IV.

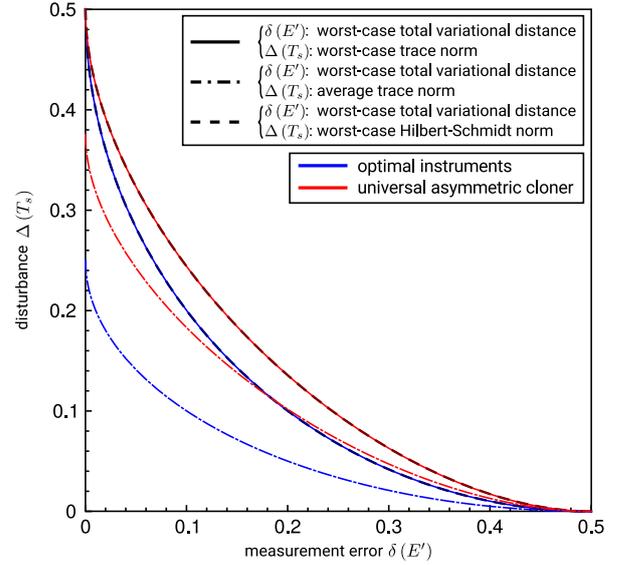


FIG. A1. Comparison of optimal quantum instruments (blue) with the optimal universal asymmetric quantum cloner (red) for different distance measures based on simulations. The tradeoff relation of the main text based on the measures of Eqs. (1) and (2) is shown (solid lines) and equivalent to a properly scaled version of the worst-case Hilbert-Schmidt norm (overlaid dashed lines) and to the worst-case infidelity (not shown). For averaging over all quantum states instead of taking the supremum of the trace norm for the disturbance, one obtains the dashdotted lines.

We here show the tradeoff relations for different choices of disturbance measures, while the measurement error is always quantified as in Eq. (1). For various meaningful measures, we observe that the optimal instruments outperform the cloner, see Fig. A1.

## VI. EXPERIMENTAL SETUP

Due to experimental and practical limitations, the actual experimental setup has been slightly different than described in the main text. However, the actual implementation is fully equivalent to the description there. In

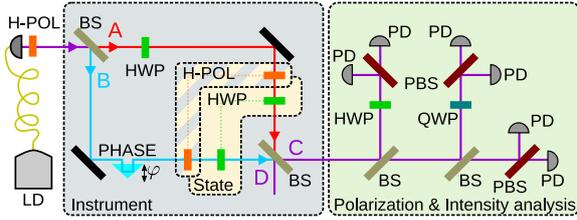


FIG. A2. Actual experimental setup. Light from a diode laser (LD) propagates through a single mode fiber and is sent through a fixed polarizer (H-POL). A beam splitter (BS) creates a spatial superposition. The attenuation of one arm can be adjusted using a half waveplate (HWP) in arm A and another H-POL. The relative phase  $\varphi$  can be varied using a piezo controlled prism. H-POLs together with variable HWPs ensure equal polarization in both arms as indicated by the dotted lines. As the H-POLs are used to vary the attenuation as well as to set the polarization state, they are part of both the instrument and the state preparation. The reflection from arm A on the second BS introduces a coupling between polarization and path. Polarization and intensity measurements are performed in output port C using waveplates (HWP and QWP), polarizing beam splitters (PBS) and photodiodes (PD). Output port D is not monitored, as for phase  $\varphi_0$  it is redundant to the output of port C at phase  $\varphi_0 + \pi$ .

order to be able to fully tune the attenuation in one of the interferometer arms, we use a half waveplate (HWP) sandwiched between two polarizers. Therefore, the polarization state  $\rho$  cannot be set before. Hence, we decided to first create the spatial superposition state  $|\phi_0\rangle$  using waveplates and polarizers and subsequently set  $\rho$  in both interferometer arms separately. With this approach, we still achieve at this stage a separable state  $\rho \otimes |\phi_0\rangle\langle\phi_0|$  within the interferometer before the interaction. As we set the polarization state directly in front of the second beam splitter of the interferometer, the reflection of beam A on the beam splitter already provides the interaction between system and auxiliary system. This reflection induces the unitary transformation  $U$  as described in the main text, enabling us to obtain the Kraus operators given in Eq. (9).

Since for a perfect beam splitter the output ports are interchanged for  $\varphi_0 \leftrightarrow \varphi_0 + \pi$ , we use only output port C to obtain data for both projections, considering the phases  $\varphi_1 = \varphi_0$  and  $\varphi_2 = \varphi_0 + \pi$ . This way, both projections are carried out with exactly the same equipment, reducing possible experimental errors.

An intuitive strategy consists of setting a specific instrument and then varying the polarization state  $\rho$ , which however requires to keep the instrument parameters ( $\alpha$  and  $\varphi$ ) stable. It turns out to be experimentally more favorable to prepare different polarization states  $\rho$  and then vary the phase  $\varphi$  for fixed  $\alpha$  and  $\rho$ . One thus associates measurements which correspond to the same state  $|\phi_0\rangle$  of the auxiliary system to the same instrument.

## VII. CHOICE OF POLARIZATION STATES

According to the parametrization  $|\psi\rangle = \cos\frac{\theta}{2}|H\rangle + \sin\frac{\theta}{2}|V\rangle$ , the experimentally prepared values for  $\theta$  were  $\{-20^\circ, -10^\circ, 0^\circ, 10^\circ, 20^\circ, 70^\circ, 80^\circ, 90^\circ, 100^\circ, 110^\circ, 160^\circ, 170^\circ, 180^\circ, 190^\circ, 200^\circ, 270^\circ\}$ . For  $\theta = 0^\circ$  and  $\theta = 180^\circ$ , the prepared state corresponds to horizontal polarization  $|H\rangle$  and vertical polarization  $|V\rangle$ , respectively. Thus, the reflection in beam A only introduces a phase, as for example the state for  $\theta = 0^\circ$  is transformed according to

$$\begin{aligned} |H\rangle \otimes (\cos\alpha|A\rangle + \sin\alpha e^{i\varphi}|B\rangle) &\rightarrow \\ |H\rangle \otimes (i\cos\alpha|A\rangle + \sin\alpha e^{i\varphi}|B\rangle), \end{aligned} \quad (\text{A12})$$

which does not change the state of the polarization. The disturbance therefore (ideally) vanishes. In contrast, for  $\theta = 90^\circ$ , we expect

$$\begin{aligned} (|H\rangle + |V\rangle) \otimes (\cos\alpha|A\rangle + \sin\alpha e^{i\varphi}|B\rangle) &\rightarrow \\ i(|H\rangle - |V\rangle) \otimes \cos\alpha|A\rangle + (|H\rangle + |V\rangle) \otimes \sin\alpha e^{i\varphi}|B\rangle, \end{aligned} \quad (\text{A13})$$

where normalization is omitted. For a given instrument characterized by  $\{\alpha, \varphi\}$ , this polarization state is expected to give the largest disturbance  $\Delta$ .

For the Kraus operators given in Eq. (9), we find for  $E'_j = K_j^\dagger K_j$  for  $j = 1, 2$ ,

$$E'_{1,2} = \frac{1}{2} \begin{pmatrix} 1 \pm \sin 2\alpha \cos \varphi & 0 \\ 0 & 1 \mp \sin 2\alpha \cos \varphi \end{pmatrix}. \quad (\text{A14})$$

Therefore, the distance of the outcome probabilities, used to obtain  $\delta$ , becomes

$$\begin{aligned} \frac{1}{2} \sum_i \left| \text{tr}(E'_i |\psi\rangle\langle\psi|) - |\langle i|\psi\rangle|^2 \right| &= \\ |\cos\theta (1 - \cos\varphi \sin 2\alpha)|, \end{aligned} \quad (\text{A15})$$

which vanishes for  $\theta = 90^\circ$  (and  $\theta = 270^\circ$ ) and can be maximal for  $\theta = 0^\circ$  (and  $\theta = 180^\circ$ ).

## VIII. ERROR ANALYSIS OF EXPERIMENTAL DATA

The statistical error of the data shown in Fig. 1 is estimated by comparing the results obtained in redundant measurements. The standard deviation of the measurement error is estimated to be around  $8.3 \cdot 10^{-5}$ , whereas the  $1\sigma$ -error bar for the estimated disturbance is approximately  $7.0 \cdot 10^{-5}$ . Those values are thus too small to be visible in Fig. 1.

Additionally to statistical errors, two different sources of systematic errors have been identified. First, the state preparation as well as the interaction are not perfectly implemented. The imperfect preparation of the initial polarization state and of the state analysis are the main

reasons that the identity channel with no disturbance at all (but high measurement error) cannot be implemented perfectly, leading to a residual disturbance, which appears as an increase of the minimal disturbance  $\Delta$  of the data in the plot. In any case, this type of error only reduces the quality of the prepared quantum instruments and does not lead to faulty conclusions.

However, as a second type of systematic error one has to ensure that the prepared polarization states are describing a great circle on the Bloch sphere and contain the states with extremal results sufficiently well. This error can be approximated by considering the data as shown in Fig. 5. By applying a parabolic model for the data points around the extrema of the probability graphs and the maxima of the trace distance graphs, the deviation of the extrema from the measured points can be estimated. This effect might cause a quantum instrument to look better than it actually is, i.e., less disturb-

ing together with smaller measurement error. Yet, for the dataset shown in Fig. 5 b), the parabolic fit results in a maximum at  $\theta \approx 89.95^\circ$  with a trace distance larger by only 0.02% compared to the trace distance at  $\theta = 90^\circ$ . The probabilities in Fig. 5 a) around  $\theta = 0^\circ$  and  $\theta = 180^\circ$  can nicely be described by parabolae, where the extrema coincide with our measured points. Thus, the systematic effect of underestimating the measurement error or the disturbance due to badly chosen measurement states is negligibly small.

In conclusion, the different sources of errors overall reduce the quality of the implemented quantum instruments and do not lead to an underestimation of disturbance and measurement error, respectively. We can thus show the implementation of instruments better than the optimal quantum cloner with high significance.

## 5.4 Conclusion

In this chapter, the concept of measurements in quantum systems has been discussed. *Weak* measurements allow to obtain a small amount of information about a quantum system, thereby causing only a little disturbance. The shift induced to the measurement device is proportional to the weak value, which has been introduced as the outcome of a weak measurement. As discussed here and in [P9], the interaction strength between the quantum system of interest and the measurement system limits the maximally achievable weak value and therefore puts a limit to the maximally obtainable amplification. Yet, in situations of relatively weak coupling, the weak amplification method can be employed for practical purposes, e.g., for aligning optical interferometers with a single position resolving detector placed at a single distance. Although for weak measurements, the coupling between system and pointer is kept as weak as possible, the interaction between those systems entangles them, which leads, upon measuring the pointer, to a disturbance to the system itself.

Similarly, as introduced by Dirac [206] and also illustrated by Heisenberg with his gedankenexperiment of a microscope [240], every measurement process introduces in general some backaction and therefore disturbance to the system which is measured. However, some measurement protocols might be more favorable in this aspect than others. Here and in Ref. [P10], different measurement protocols are compared leading finally also to the optimal measurement protocol allowing to obtain information from a quantum state at the lowest cost in terms of disturbance. This set of optimal instruments has been proposed in [234] and discussed in more detail for binary qubit measurements and experimentally investigated in [P10]. As the non-existing [245] perfect quantum cloner could be used to retrieve an arbitrary amount of information without inducing any disturbance, the approximate quantum cloning protocol [247, 248] is naively assumed to at least give an optimal tradeoff between information (or measurement error, respectively) and disturbance. Yet, the optimal measurements as derived in [P10] outperform other schemes including the optimal asymmetric cloning protocol with an experimentally significant difference between the different protocols.



# Chapter 6

## Conclusions

A strong focus of this thesis has been quantum state tomography, which is one of the most fundamental concepts to retrieve information from a quantum system. Quantum state tomography, which allows to fully describe a given quantum system in terms of its density matrix (Sec. 3.2), provides also the basis for quantum process tomography, which is an unrivaled diagnostic tool. State tomography not only provides a beneficial way to verify the successful experimental implementation and realization of a specific state or algorithm, it also allows to evaluate arbitrary functions on the state without invoking further measurements once the density matrix has been obtained. On the other hand, quantum state tomography is an expensive task as the number of measurement settings generally scales exponentially with the number of qubits, unless assumptions regarding, e.g., the symmetry of the state are made, which allow to use more efficient schemes significantly reducing the measurement effort. However, additionally to the limitation in the number of measurement settings, one has to face the issue of finite statistics arising from the finite total measurement time.

While in the limit of infinite statistics the state estimate is, at least in the absence of additional experimental systematic deviations, a perfect estimate of the true state, i.e., the state estimator is consistent, it shows a variety of different problems for typical sample sizes. This includes the occurrence of negative eigenvalues, which are often prevented using constrained optimization methods (Sec. 3.3). Those, however, may in turn lead to systematic deviations, resulting in an underestimation of the state fidelity or an overestimation of the entanglement present in the state. In this thesis, the systematic effects due to finite sample size effects together with constrained optimization have been discussed. In the publication [P1], we also provide a direct method to obtain confidence regions for various figures of merit, although those confidence regions are rather conservative. We investigated the consequences of finite statistics in state tomography in great detail in this thesis and in our work [P2], which allows to identify statistically significant contributions in a tomographically obtained estimate. Based on our findings, one can easily calculate statistical error bars for several figures of merit without needing to resort to computationally costly Monte Carlo methods. Hypothesis testing allows to identify colored noise or systematic deviations in the quantum state, providing a simple and efficient means to

benchmark quantum states and quantum processes.

As a second main pillar of this thesis, entanglement in multipartite quantum systems has been discussed. Detecting and verifying genuine multipartite entanglement, i.e., entanglement truly involving several parties, which cannot be generated using local operations and classical communication from bipartite entangled states, is still a hard task. In this thesis, a constructive and efficient tool (Sec. 4.2 and in [P3]) has been introduced and experimentally demonstrated which allows to verify entanglement in multiqubit states with only two measurements. Based on those findings, this tool has been adapted for entanglement verification for multiqutrit states (as discussed in [P4]). In future work, one could try to find a further generalization of this means to arbitrary dimensions, allowing to prove genuine multipartite entanglement in qudit states or to further improve the noise robustness of our criteria, potentially by also including more than the minimal amount of two measurements.

Even though noisy channels may make some information about a quantum state inaccessible, we have devised a method to still retrieve valuable information (Sec. 4.3 as well as [P5]). There, we have demonstrated that in a scenario without well defined local reference frames one is still able to detect entanglement and deduce several properties of the state such as the purity. This enables to significantly relax requirements on the measurement apparatus and to assess the quality of the state preparation without strong assumption on the state measurement tools. As a next step, one could try to find an efficient parametrization of a quantum state, which incorporates the symmetries under local unitary transformations. This could allow to perform some new type of tomography that respects the transformations of the noisy channels, which cannot be resolved in the experiment.

Although correlations provide a useful way to characterize a state and to describe joint measurements of several observers, the standard definition of correlations does not comply with some basic postulates as described in Sec. 4.4 and in the publications [P6] and [P7]. This has been demonstrated with the provocative experimental result of a tripartite quantum state, which violates tripartite Bell inequalities, but does not show any tripartite correlations. In context of this work, we have proposed a class of multipartite states without standard correlations, which provide some testbed for studying operationally meaningful redefinitions of multipartite quantum correlations.

Finally, quantum measurements have been reviewed in different contexts. On one hand, the scenario of weak measurements with pre- and postselected quantum particles which have weakly interacted with some measurement device has been introduced (Sec. 5.2). These weak measurements allow to perform measurements on a quantum system without a large disturbance to the latter. Yet, they provide only meaningful results in ensemble measurements as the individual outcome does not carry a significant amount of information. Albeit object of large controversies, the closely related concept of weak values can under some circumstances be used to amplify small signals, which, for example, allows to efficiently deduce the misalignment of optical interferometers. In our publication [P8], we discuss similarities and differences between eigenvalues, expectation values and weak values and consider the generalization of the notion of weak values to the case of mixed

pre- and postselection states. We applied the concept of the amplification of small signals in our work shown in [P9] to the precise measurement of beam misalignments in optical interferometers. While previous works usually consider the weak value to be constant in the regime of a small coupling strength between the pointer system and the actual system, we also describe the incoherence in the pointer system caused by the arising entanglement between both systems. This enabled us to nicely model the possibilities and limitations of the weak value amplification method.

Although the close connection of a measurement and the thereby induced disturbance to the quantum state is almost ubiquitous in quantum studies, there is no clear candidate for the *best* measurement-disturbance tradeoff relation. However, in this thesis and in the paper [P10], a new approach to measurement-disturbance relations is introduced (Sec. 5.3) achieving a new level of generality. Based on the language of quantum instruments, different measurement protocols are compared, including a promising measurement scheme based on approximate quantum cloning. But first and foremost, the optimal quantum instruments have been derived, which significantly outperform all other protocols. As we have shown also experimentally, this significant improvement does not remain a mere theoretical result, but is indeed accessible for actual quantum measurements. Those findings will allow future applications to retrieve the maximal amount of information, while reducing the disturbance thereby caused to the state to the minimal amount physically possible.



# List of Publications

## Publications discussed in this thesis

in order of appearance in thesis

[P1]

C. Schwemmer, **L. Knips**, D. Richart, T. Moroder, M. Kleinmann, O. Gühne, and H. Weinfurter,

Systematic Errors in Current Quantum State Tomography Tools

Phys. Rev. Lett. **114**, 080403 (2015).

Contribution: Conducted numerical study together with CS, prepared manuscript together with CS, DR, TM, MK, OG, HW.

[P2]

**L. Knips**, C. Schwemmer, N. Klein, J. Reuter, G. Tóth, and H. Weinfurter,

Multiqubit State Tomography with Finite Data

arXiv:1512.06866.

Contribution: Conducted numerical study, derived and proved theory, prepared manuscript together with CS, NK, GT, HW.

[P3]

**L. Knips**, C. Schwemmer, N. Klein, M. Wieśniak, and H. Weinfurter,

Multipartite entanglement detection with minimal effort

Phys. Rev. Lett. **117**, 210504 (2016).

Contribution: Designed setup together with CS and HW, built and conducted experiment together with CS and NK, evaluated data together with CS, NK, derived theory together with MW, prepared manuscript together with CS, NK, MW, HW.

[P4]

K. Rosołek, M. Wieśniak, and **L. Knips**,

Quadratic Entanglement Criteria for Qutrits

arXiv:1512.07099.

Contribution: Derived theory together with KR and MW, edited manuscript.

**[P5]**

**L. Knips**, J. Dziewior, W. Klobus, W. Laskowski, T. Paterek, P. J. Shadbolt, H. Weinfurter, and J. D. A. Meinecke

Multipartite Entanglement Analysis From Random Correlations

arXiv:1910.10732

Contribution: Conducted experiment and evaluated data together with JD and JDAM, performed simulations together with WK, derived entanglement criteria, wrote manuscript together with JD, WK, WL, JDAM, TP and HW.

**[P6]**

C. Schwemmer, **L. Knips**, M. C. Tran, A. d. Rosier, W. Laskowski, T. Paterek, and H. Weinfurter,

Genuine Multipartite Entanglement without Multipartite Correlations

Phys. Rev. Lett. **114**, 180501 (2015).

Contribution: Prepared and conducted experiment together with CS, evaluated data together with CS, prepared manuscript together with CS, MCT, AdR, WL, TP, HW.

**[P7]**

M. C. Tran, M. Zuppardo, A. d. Rosier, **L. Knips**, W. Laskowski, T. Paterek, and H. Weinfurter,

Genuine N-partite entanglement without N-partite correlation functions

Phys. Rev. A **95**, 062331 (2017).

Contribution: Conducted numerical study, edited the manuscript together with MCT, MZ, AdR, WL, TP, HW.

**[P8]**

L. Vaidman, A. Ben-Israel, J. Dziewior, **L. Knips**, M. Weißl, J. Meinecke, C. Schwemmer, R. Ber, and H. Weinfurter,

Weak value beyond conditional expectation value of the pointer readings

Phys. Rev. A **96**, 032114 (2017).

Contribution: Designed experiment together with CS and HW, conducted experiment and evaluated data together with JD, MW, prepared manuscript together with LV, ABI, JD, MW, JM, HW.

**[P9]**

J. Dziewior, **L. Knips**, D. Farfurnik, K. Senkalla, N. Benshalom, J. Efroni, J. Meinecke, S. Bar-Ad, H. Weinfurter, and L. Vaidman

Universality of weak interaction of a particle in an arm of an interferometer and its application for alignment

Proc Natl Acad Sci **116**, 2881 (2019).

Contribution: Designed experiment together with JD and HW, conducted experiment and evaluated data together with JD, KS, prepared manuscript together with JD, DF, KS, NB, JE, JM, SBA, HW, LV.

**[P10]**

**L. Knips**, J. Dziewior, A. Hashagen, J. D. A. Meinecke, M. M. Wolf, and H. Weinfurter, Preprint: Procedures for the Optimal Measurement-Disturbance Tradeoff arXiv:1808.07882 (accepted for publication in Phys. Rev. A).

Contribution: Designed experiment together with JD, conducted experiment and evaluated data together with JD, prepared manuscript together with JD, AH, JDAM, MW, HW.

---

**Further publications****[P11]**

W. Laskowski, C. Schwemmer, D. Richart, **L. Knips**, T. Paterek, and H. Weinfurter, Optimized state independent entanglement detection based on geometrical threshold criterion

Phys. Rev. A **88**, 022327 (2013).

Contribution: Edited the manuscript together with WL, CS, DR, TP, HW.

**[P12]**

A. Ben-Israel, **L. Knips**, J. Dziewior, J. Meinecke, A. Danan, H. Weinfurter, and L. Vaidman,

An Improved Experiment to Determine the ‘Past of a Particle’ in the Nested Mach-Zehnder Interferometer

Chin. Phys. Lett. **34**, 020301 (2017).

Contribution: Prepared the manuscript together with ABI, JD, JM, AD, HW, LV.

**[P13]**

X.-Y. Xu, W.-W. Pan, Q.-Q. Wang, J. Dziewior, **L. Knips**, Y. Kedem, K. Sun, J.-S. Xu, Y.-J. Han, C.-F. Li, G.-C. Guo, and L. Vaidman

Measurements of nonlocal variables and demonstration of the failure of the product rule for a pre and postselected pair of photons

Phys. Rev. Lett. **122**, 100405 (2019).

Contribution: Performed calculations together with JD, edited the manuscript together with XYX, JD, YK, KS, CFL, LV.



# Bibliography

- [1] Helge Kragh. *Quantum Generations - A History of Physics in the Twentieth Century*. Princeton University Press, Kassel, 1999. ISBN 978-0-691-01206-3.
- [2] Martin J. Klein. Max Planck and the beginnings of the quantum theory. *Archive for History of Exact Sciences*, 1(5):459–479, oct 1961. doi: 10.1007/bf00327765. URL <https://doi.org/10.1007/bf00327765>.
- [3] Max Planck and Morton Masius. *The Theory of Heat Radiation*. P. Blakiston's Son & Co. (Philadelphia), 1914.
- [4] Daniel Greenberger, Klaus Hentschel, and Friedel Weinert. *Compendium of Quantum Physics - Concepts, Experiments, History and Philosophy*. Springer Science & Business Media, Berlin Heidelberg, 2009. ISBN 978-3-540-70626-7.
- [5] Albert Einstein. Über einen die Erzeugung und Verwandlung des Lichtes betreffenden heuristischen Gesichtspunkt. *Annalen der Physik*, 17:132, 1905.
- [6] Albert Einstein. Zur Quantentheorie der Strahlung. *Mitteilungen der Physikalischen Gesellschaft Zürich*, 18:47–62, 1916.
- [7] Niels Bohr. I. On the constitution of atoms and molecules. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 26(151):1–25, jul 1913. doi: 10.1080/14786441308634955. URL <https://doi.org/10.1080/14786441308634955>.
- [8] Arthur H. Compton. A Quantum Theory of the Scattering of X-rays by Light Elements. *Physical Review*, 21(5):483–502, may 1923. doi: 10.1103/physrev.21.483. URL <https://doi.org/10.1103/physrev.21.483>.
- [9] Louis De Broglie. Recherches sur la théorie des Quanta. *Annales de Physique*, 10(3):22–128, 1925. doi: 10.1051/anphys/192510030022. URL <https://doi.org/10.1051/anphys/192510030022>.
- [10] Clinton Davisson and Lester H. Germer. Diffraction of Electrons by a Crystal of Nickel. *Physical Review*, 30(6):705–740, dec 1927. doi: 10.1103/physrev.30.705. URL <https://doi.org/10.1103/physrev.30.705>.

- [11] Werner Heisenberg. Über quantentheoretische Umdeutung kinematischer und mechanischer Beziehungen. *Zeitschrift für Physik*, 33(1):879–893, dec 1925. doi: 10.1007/bf01328377. URL <https://doi.org/10.1007/bf01328377>.
- [12] Max Born and Pascual Jordan. Zur Quantenmechanik. *Zeitschrift für Physik*, 34(1):858–888, dec 1925. doi: 10.1007/bf01328531. URL <https://doi.org/10.1007/bf01328531>.
- [13] Max Born, Werner Heisenberg, and Pascual Jordan. Zur quantenmechanik. II. *Zeitschrift für Physik*, 35(8-9):557–615, aug 1926. doi: 10.1007/bf01379806. URL <https://doi.org/10.1007/bf01379806>.
- [14] Erwin Schrödinger. Quantisierung als Eigenwertproblem. *Annalen der Physik*, 384(4):361–376, 1926. doi: 10.1002/andp.19263840404. URL <https://doi.org/10.1002/andp.19263840404>.
- [15] Werner Heisenberg. Über den anschaulichen Inhalt der quantentheoretischen Kinetik und Mechanik. *Zeitschrift für Physik*, 43(3-4):172–198, mar 1927. doi: 10.1007/bf01397280. URL <https://doi.org/10.1007/bf01397280>.
- [16] Silvia A. Camejo. Die Bohr-Einstein-Debatte. In *Skurrile Quantenwelt*, pages 87–101. Springer Berlin Heidelberg, 2006. doi: 10.1007/3-540-29721-9\_10. URL [https://doi.org/10.1007/3-540-29721-9\\_10](https://doi.org/10.1007/3-540-29721-9_10).
- [17] Albert Einstein, Boris Podolsky, and Nathan 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>.
- [18] Niels Bohr. Can Quantum-Mechanical Description of Physical Reality be Considered Complete? *Physical Review*, 48(8):696–702, oct 1935. doi: 10.1103/physrev.48.696. URL <https://doi.org/10.1103/physrev.48.696>.
- [19] Erwin Schrödinger. Discussion of Probability Relations between Separated Systems. *Mathematical Proceedings of the Cambridge Philosophical Society*, 31(04):555, oct 1935. doi: 10.1017/s0305004100013554. URL <https://doi.org/10.1017/s0305004100013554>.
- [20] Erwin Schrödinger and Paul A. M. Dirac. Probability relations between separated systems. *Mathematical Proceedings of the Cambridge Philosophical Society*, 32(03):446, oct 1936. doi: 10.1017/s0305004100019137. URL <https://doi.org/10.1017/s0305004100019137>.
- [21] John S. Bell. On the Einstein-Podolsky-Rosen Paradox. *Physics*, 1:195–200, 1964.

- [22] Stuart J. Freedman and John F. Clauser. Experimental Test of Local Hidden-Variable Theories. *Physical Review Letters*, 28(14):938–941, apr 1972. doi: 10.1103/physrevlett.28.938. URL <https://doi.org/10.1103/physrevlett.28.938>.
- [23] John F. Clauser. Experimental Investigation of a Polarization Correlation Anomaly. *Physical Review Letters*, 36(21):1223–1226, may 1976. doi: 10.1103/physrevlett.36.1223. URL <https://doi.org/10.1103/physrevlett.36.1223>.
- [24] John F. Clauser and Abner Shimony. Bell’s theorem. Experimental tests and implications. *Reports on Progress in Physics*, 41(12):1881–1927, dec 1978. doi: 10.1088/0034-4885/41/12/002. URL <https://doi.org/10.1088/0034-4885/41/12/002>.
- [25] Edward S. Fry and Randall C. Thompson. Experimental Test of Local Hidden-Variable Theories. *Physical Review Letters*, 37(8):465–468, aug 1976. doi: 10.1103/physrevlett.37.465. URL <https://doi.org/10.1103/physrevlett.37.465>.
- [26] Alain Aspect, Philippe Grangier, and Gérard Roger. Experimental Tests of Realistic Local Theories via Bell’s Theorem. *Physical Review Letters*, 47(7):460–463, aug 1981. doi: 10.1103/physrevlett.47.460. URL <https://doi.org/10.1103/physrevlett.47.460>.
- [27] 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>.
- [28] Mary A. Rowe, David Kielpinski, Volker Meyer, Charles A. Sackett, Wayne M. Itano, Christopher Monroe, and David J. Wineland. Experimental violation of a Bell’s inequality with efficient detection. *Nature*, 409(6822):791–794, feb 2001. doi: 10.1038/35057215. URL <https://doi.org/10.1038/35057215>.
- [29] Bas Hensen, Hannes Bernien, Anaïs E. Dréau, Andreas Reiserer, Norbert Kalb, Machiel S. Blok, Just Ruitenber, Raymond F. L. Vermeulen, Raymond N. Schouten, Carlos Abellán, Waldimar Amaya, V. Pruneri, Morgan W. Mitchell, Matthew Markham, Daniel J. Twitchen, David Elkouss, Stephanie Wehner, Tim H. Taminiau, and Ronald Hanson. Loophole-free Bell inequality violation using electron spins separated by 1.3 kilometres. *Nature*, 526(7575):682–686, oct 2015. doi: 10.1038/nature15759. URL <https://doi.org/10.1038/nature15759>.
- [30] Lynden K. Shalm, Evan Meyer-Scott, Bradley G. Christensen, Peter Bierhorst, Michael A. Wayne, Martin J. Stevens, Thomas Gerrits, Scott Glancy, Deny R. Hamel, Michael S. Allman, Kevin J. Coakley, Shellee D. Dyer, Carson Hodge, Adriana E. Lita, Varun B. Verma, Camilla Lambrocco, Edward Tortorici, Alan L. Migdall, Yanbao Zhang, Daniel R. Kumor, William H. Farr, Francesco Marsili, Matthew D. Shaw, Jeffrey A. Stern, Carlos Abellán, Waldimar Amaya, Valerio Pruneri, Thomas Jennewein, Morgan W. Mitchell, Paul G. Kwiat, Joshua C. Bienfang, Richard P. Mirin,

- Emanuel Knill, and Sae Woo Nam. Strong Loophole-Free Test of Local Realism. *Physical Review Letters*, 115(25), dec 2015. doi: 10.1103/physrevlett.115.250402. URL <https://doi.org/10.1103/physrevlett.115.250402>.
- [31] Marissa Giustina, Marijn A. M. Versteegh, Sören Wengerowsky, Johannes Handsteiner, Armin Hochrainer, Kevin Phelan, Fabian Steinlechner, Johannes Kofler, Jan-Åke Larsson, Carlos Abellán, Waldimar Amaya, Valerio Pruneri, Morgan W. Mitchell, Jörn Beyer, Thomas Gerrits, Adriana E. Lita, Lynden K. Shalm, Sae Woo Nam, Thomas Scheidl, Rupert Ursin, Bernhard Wittmann, and Anton Zeilinger. Significant-Loophole-Free Test of Bell’s Theorem with Entangled Photons. *Physical Review Letters*, 115(25), dec 2015. doi: 10.1103/physrevlett.115.250401. URL <https://doi.org/10.1103/physrevlett.115.250401>.
- [32] Wenjamin Rosenfeld, Daniel Burchardt, Robert Garthoff, Kai Redeker, Norbert Ortegel, Markus Rau, and Harald Weinfurter. Event-Ready Bell Test Using Entangled Atoms Simultaneously Closing Detection and Locality Loopholes. *Physical Review Letters*, 119(1), jul 2017. doi: 10.1103/physrevlett.119.010402. URL <https://doi.org/10.1103/physrevlett.119.010402>.
- [33] The BIG Bell Test Collaboration. Challenging local realism with human choices. *Nature*, 557(7704):212–216, may 2018. doi: 10.1038/s41586-018-0085-3. URL <https://doi.org/10.1038/s41586-018-0085-3>.
- [34] Artur K. Ekert. Quantum cryptography based on Bell’s theorem. *Physical Review Letters*, 67(6):661–663, aug 1991. doi: 10.1103/physrevlett.67.661. URL <https://doi.org/10.1103/physrevlett.67.661>.
- [35] Charles H. Bennett and Stephen J. Wiesner. Communication via one- and two-particle operators on Einstein-Podolsky-Rosen states. *Physical Review Letters*, 69(20):2881–2884, nov 1992. doi: 10.1103/physrevlett.69.2881. URL <https://doi.org/10.1103/physrevlett.69.2881>.
- [36] Klaus Mattle, Harald Weinfurter, Paul G. Kwiat, and Anton Zeilinger. Dense Coding in Experimental Quantum Communication. *Physical Review Letters*, 76(25):4656–4659, jun 1996. doi: 10.1103/physrevlett.76.4656. URL <https://doi.org/10.1103/physrevlett.76.4656>.
- [37] Charles H. Bennett, Gilles Brassard, Claude Crépeau, Richard Jozsa, Asher Peres, and William K. Wootters. Teleporting an unknown quantum state via dual classical and Einstein-Podolsky-Rosen channels. *Physical Review Letters*, 70(13):1895–1899, mar 1993. doi: 10.1103/physrevlett.70.1895. URL <https://doi.org/10.1103/physrevlett.70.1895>.
- [38] Marek Żukowski, Anton Zeilinger, Michael A. Horne, and Artur K. Ekert. “Event-ready-detectors” Bell experiment via entanglement swapping. *Physical Review Let-*

- ters, 71(26):4287–4290, dec 1993. doi: 10.1103/physrevlett.71.4287. URL <https://doi.org/10.1103/physrevlett.71.4287>.
- [39] Richard P. Feynman. Simulating physics with computers. *International Journal of Theoretical Physics*, 21(6-7):467–488, jun 1982. doi: 10.1007/bf02650179. URL <https://doi.org/10.1007/bf02650179>.
- [40] 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>.
- [41] 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>.
- [42] 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>.
- [43] Michael A. Nielsen and Isaac L. Chuang. *Quantum computation and quantum information*. Cambridge University Press, 2000.
- [44] Mark Fox. *Quantum Optics - An Introduction*. OUP Oxford, New York, London, 2006. ISBN 978-0-191-52425-7.
- [45] Lieven M. K. Vandersypen and Isaac L. Chuang. NMR techniques for quantum control and computation. *Reviews of Modern Physics*, 76(4):1037–1069, jan 2005. doi: 10.1103/revmodphys.76.1037. URL <https://doi.org/10.1103/revmodphys.76.1037>.
- [46] Ronald Hanson, Leo P. Kouwenhoven, Jason R. Petta, Seigo Tarucha, and Lieven M. K. Vandersypen. Spins in few-electron quantum dots. *Reviews of Modern Physics*, 79(4):1217–1265, oct 2007. doi: 10.1103/revmodphys.79.1217. URL <https://doi.org/10.1103/revmodphys.79.1217>.
- [47] Mark Saffman, Thad G. Walker, and Klaus Mølmer. Quantum information with Rydberg atoms. *Reviews of Modern Physics*, 82(3):2313–2363, aug 2010. doi: 10.1103/revmodphys.82.2313. URL <https://doi.org/10.1103/revmodphys.82.2313>.
- [48] Yuriy Makhlin, Gerd Schön, and Alexander Shnirman. Quantum-state engineering with Josephson-junction devices. *Reviews of Modern Physics*, 73(2):357–400, may 2001. doi: 10.1103/revmodphys.73.357. URL <https://doi.org/10.1103/revmodphys.73.357>.
- [49] Pieter Kok, William J. Munro, Kae Nemoto, Timothy C. Ralph, Jonathan P. Dowling, and Gerard J. Milburn. Linear optical quantum computing with photonic qubits.

- Reviews of Modern Physics*, 79(1):135–174, jan 2007. doi: 10.1103/revmodphys.79.135. URL <https://doi.org/10.1103/revmodphys.79.135>.
- [50] 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>.
- [51] Stephen Barnett. *Quantum Information*. OUP Oxford, New York, London, 2009. ISBN 978-0-198-52762-6.
- [52] Fabio Masillo, Giuseppe Sclarici, and Sandro Sozzo. proper versus improper mixtures: Toward a quaternionic quantum mechanics. *Theoretical and Mathematical Physics*, 160(1):1006–1013, jul 2009. doi: 10.1007/s11232-009-0091-6. URL <https://doi.org/10.1007/s11232-009-0091-6>.
- [53] 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>.
- [54] 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>.
- [55] He Lu, Qi Zhao, Zheng-Da Li, Xu-Fei Yin, Xiao Yuan, Jui-Chen Hung, Luo-Kan Chen, Li Li, Nai-Le Liu, Cheng-Zhi Peng, Yeong-Cherng Liang, Xiongfeng Ma, Yu-Ao Chen, and Jian-Wei Pan. Probing Multipartite Entanglement Structure with Minimal Resources. *ArXiv e-prints*, November 2017. URL <https://arxiv.org/pdf/1711.01784>.
- [56] Géza Tóth and Otfried Gühne. Entanglement detection in the stabilizer formalism. *Physical Review A*, 72(2), aug 2005. doi: 10.1103/physreva.72.022340. URL <https://doi.org/10.1103/physreva.72.022340>.
- [57] Anders S. Sørensen and Klaus Mølmer. Entanglement and Extreme Spin Squeezing. *Physical Review Letters*, 86(20):4431–4434, may 2001. doi: 10.1103/physrevlett.86.4431. URL <https://doi.org/10.1103/physrevlett.86.4431>.
- [58] John Watrous. *The Theory of Quantum Information*. Cambridge University Press, 2018. ISBN 1107180562. URL <https://cs.uwaterloo.ca/~watrous/TQI/TQI.pdf>.
- [59] Michael A. Nielsen. Conditions for a Class of Entanglement Transformations. *Physical Review Letters*, 83(2):436–439, jul 1999. doi: 10.1103/physrevlett.83.436. URL <https://doi.org/10.1103/physrevlett.83.436>.

- [60] Charles H. Bennett, Sandu Popescu, Daniel Rohrlich, John A. Smolin, and Ashish V. Thapliyal. Exact and asymptotic measures of multipartite pure-state entanglement. *Physical Review A*, 63(1), dec 2000. doi: 10.1103/physreva.63.012307. URL <https://doi.org/10.1103/physreva.63.012307>.
- [61] Yeong-Cherng Liang, Lluís Masanes, and Andrew C. Doherty. Convertibility between two-qubit states using stochastic local quantum operations assisted by classical communication. *Physical Review A*, 77(1), jan 2008. doi: 10.1103/physreva.77.012332. URL <https://doi.org/10.1103/physreva.77.012332>.
- [62] Wolfgang Dür, Guifre Vidal, and J. Ignacio 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>.
- [63] Nicolas Gisin. Hidden quantum nonlocality revealed by local filters. *Physics Letters A*, 210(3):151–156, jan 1996. doi: 10.1016/s0375-9601(96)80001-6. URL [https://doi.org/10.1016/s0375-9601\(96\)80001-6](https://doi.org/10.1016/s0375-9601(96)80001-6).
- [64] Yuan-Yuan Zhao, Markus Grassl, Bei Zeng, Guo-Yong Xiang, Chao Zhang, Chuan-Feng Li, and Guang-Can Guo. Experimental detection of entanglement polytopes via local filters. *npj Quantum Information*, 3(1), mar 2017. doi: 10.1038/s41534-017-0007-5. URL <https://doi.org/10.1038/s41534-017-0007-5>.
- [65] Michał Horodecki, Paweł Horodecki, and Ryszard Horodecki. Separability of mixed states: necessary and sufficient conditions. *Physics Letters A*, 223(1–2):1 – 8, 1996. ISSN 0375-9601. doi: [http://dx.doi.org/10.1016/S0375-9601\(96\)00706-2](http://dx.doi.org/10.1016/S0375-9601(96)00706-2). URL <http://www.sciencedirect.com/science/article/pii/S0375960196007062>.
- [66] 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>.
- [67] Barbara M. Terhal. Bell inequalities and the separability criterion. *Physics Letters A*, 271(5-6):319–326, jul 2000. doi: 10.1016/s0375-9601(00)00401-1. URL [https://doi.org/10.1016/s0375-9601\(00\)00401-1](https://doi.org/10.1016/s0375-9601(00)00401-1).
- [68] Maciej Lewenstein, Barbara Kraus, J. Ignacio Cirac, and Paweł Horodecki. Optimization of entanglement witnesses. *Physical Review A*, 62(5), oct 2000. doi: 10.1103/physreva.62.052310. URL <https://doi.org/10.1103/physreva.62.052310>.
- [69] Mohamed Bourennane, Manfred Eibl, Christian Kurtsiefer, Sascha Gaertner, Harald Weinfurter, Otfried Gühne, Philipp Hyllus, Dagmar Bruß, Maciej Lewenstein, and Anna Sanpera. Experimental Detection of Multipartite Entanglement using Witness Operators. *Physical Review Letters*, 92(8), feb 2004. doi: 10.1103/physrevlett.92.087902. URL <https://doi.org/10.1103/physrevlett.92.087902>.

- [70] Angel G. Valdenebro. Assumptions underlying Bell's inequalities. *European Journal of Physics*, 23(5):569–577, sep 2002. doi: 10.1088/0143-0807/23/5/313. URL <https://doi.org/10.1088/0143-0807/23/5/313>.
- [71] Roderich Tumulka. The Assumptions of Bell's Proof. In Mary Bell and Shan Gao, editors, *Quantum Nonlocality and Reality*, pages 79–90. Cambridge University Press, 2016. doi: 10.1017/cbo9781316219393.007. URL <https://doi.org/10.1017/cbo9781316219393.007>.
- [72] John F. Clauser, Michael A. Horne, Abner Shimony, and Richard A. Holt. Proposed Experiment to Test Local Hidden-Variable Theories. *Physical Review Letters*, 23(15):880–884, oct 1969. doi: 10.1103/physrevlett.23.880. URL <https://doi.org/10.1103/physrevlett.23.880>.
- [73] Boris S. Cirel'son. Quantum generalizations of Bell's inequality. *Letters in Mathematical Physics*, 4(2):93–100, mar 1980. doi: 10.1007/bf00417500. URL <https://doi.org/10.1007/bf00417500>.
- [74] Marco Genovese. Experimental Tests of Bell Inequalities. In Mary Bell and Shan Gao, editors, *Quantum Nonlocality and Reality*, pages 124–140. Cambridge University Press, 2016. doi: 10.1017/cbo9781316219393.009. URL <https://doi.org/10.1017/cbo9781316219393.009>.
- [75] Philipp Hyllus, Otfried Gühne, Dagmar Bruß, and Maciej Lewenstein. Relations between entanglement witnesses and Bell inequalities. *Physical Review A*, 72(1), jul 2005. doi: 10.1103/physreva.72.012321. URL <https://doi.org/10.1103/physreva.72.012321>.
- [76] Alexander Streltsov, Hermann Kampermann, Sabine Wölk, Manuel Gessner, and Dagmar Bruß. Maximal coherence and the resource theory of purity. *New Journal of Physics*, 20(5):053058, may 2018. doi: 10.1088/1367-2630/aac484. URL <https://doi.org/10.1088/1367-2630/aac484>.
- [77] Daniel M. Greenberger, Michael A. Horne, and Anton Zeilinger. Going Beyond Bell's Theorem. In *Bell's Theorem, Quantum Theory and Conceptions of the Universe*, pages 69–72. Springer Netherlands, 1989. doi: 10.1007/978-94-017-0849-4\_10. URL [https://doi.org/10.1007/978-94-017-0849-4\\_10](https://doi.org/10.1007/978-94-017-0849-4_10).
- [78] 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>.
- [79] Robert Raussendorf, Daniel E. Browne, and Hans J. Briegel. Measurement-based quantum computation on cluster states. *Physical Review A*, 68(2), aug 2003. doi: 10.1103/physreva.68.022312. URL <https://doi.org/10.1103/physreva.68.022312>.

- [80] Michael A. Nielsen. Cluster-state quantum computation. *Reports on Mathematical Physics*, 57(1):147–161, feb 2006. doi: 10.1016/s0034-4877(06)80014-5. URL [https://doi.org/10.1016/s0034-4877\(06\)80014-5](https://doi.org/10.1016/s0034-4877(06)80014-5).
- [81] Philip Walther, Kevin J. Resch, Terry Rudolph, Emmanuel Schenck, Harald Weinfurter, Vlatko Vedral, Markus Aspelmeyer, and Anton Zeilinger. Experimental One-Way Quantum Computing. *Nature*, 434(7030):169–176, 2005. URL <http://www.nature.com/nature/journal/v434/n7030/full/nature03347.html>.
- [82] Daniel Gottesman. Stabilizer Codes and Quantum Error Correction. *PhD Thesis, California Institute of Technology, Pasadena, California*, 1997. URL <https://arxiv.org/pdf/quant-ph/9705052.pdf>.
- [83] A. Robert Calderbank and Peter W. Shor. Good quantum error-correcting codes exist. *Physical Review A*, 54(2):1098–1105, aug 1996. doi: 10.1103/physreva.54.1098. URL <https://doi.org/10.1103/physreva.54.1098>.
- [84] Keisuke Fujii. Stabilizer Formalism and Its Applications. In *Quantum Computation with Topological Codes*, pages 24–55. Springer Singapore, 2015. doi: 10.1007/978-981-287-996-7\_2. URL [https://doi.org/10.1007/978-981-287-996-7\\_2](https://doi.org/10.1007/978-981-287-996-7_2).
- [85] Matteo Rossi, Marcus Huber, Dagmar Bruß, and Chiara Macchiavello. Quantum hypergraph states. *New Journal of Physics*, 15(11):113022, nov 2013. doi: 10.1088/1367-2630/15/11/113022. URL <https://doi.org/10.1088/1367-2630/15/11/113022>.
- [86] Robert 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>.
- [87] Leonard Mandel and Emil Wolf. *Optical Coherence and Quantum Optics*. Cambridge University Press, 1995. ISBN 0521417112.
- [88] Kurt Jacobs. *Quantum Measurement Theory and its Applications*. Cambridge University Press, 2014. ISBN 1107025486.
- [89] Matteo G. A. Paris. The modern tools of quantum mechanics. *The European Physical Journal Special Topics*, 203(1):61–86, apr 2012. doi: 10.1140/epjst/e2012-01535-1. URL <https://doi.org/10.1140/epjst/e2012-01535-1>.
- [90] John Neumann. *Mathematische Grundlagen der Quantenmechanik*. Springer, Berlin New York, 1955. ISBN 3540592075. URL <https://link.springer.com/book/10.1007%2F978-3-642-61409-5>.
- [91] Pier A. Mello. The von Neumann model of measurement in quantum mechanics. In *American Institute of Physics Conference Series*, volume 1575 of *American Institute*

- of Physics Conference Series*, pages 136–165, January 2014. doi: 10.1063/1.4861702. URL <https://arxiv.org/pdf/1311.7649.pdf>.
- [92] Karl Kraus. *States, Effects, and Operations: Fundamental Notions of Quantum Theory (Lecture Notes in Physics)*. Springer, 1983. ISBN 3540127321. URL <https://link.springer.com/book/10.1007%2F3-540-12732-1>.
- [93] Isaac L. Chuang and Michael A. Nielsen. Prescription for experimental determination of the dynamics of a quantum black box. *Journal of Modern Optics*, 44(11-12):2455–2467, nov 1997. doi: 10.1080/09500349708231894. URL <https://doi.org/10.1080/09500349708231894>.
- [94] Andrzej Jamiolkowski. Linear transformations which preserve trace and positive semidefiniteness of operators. *Reports on Mathematical Physics*, 3(4):275–278, dec 1972. doi: 10.1016/0034-4877(72)90011-0. URL [https://doi.org/10.1016/0034-4877\(72\)90011-0](https://doi.org/10.1016/0034-4877(72)90011-0).
- [95] Man-Duen Choi. Completely positive linear maps on complex matrices. *Linear Algebra and its Applications*, 10(3):285–290, jun 1975. doi: 10.1016/0024-3795(75)90075-0. URL [https://doi.org/10.1016/0024-3795\(75\)90075-0](https://doi.org/10.1016/0024-3795(75)90075-0).
- [96] Michael M. Wolf. Quantum Channels & Operations, Guided Tour, Lecture Notes, 2008/2009, 2012. URL <https://www-m5.ma.tum.de/foswiki/pub/M5/Allgemeines/MichaelWolf/QChannelLecture.pdf>.
- [97] Athanasios Papoulis. *Probability, Random Variables and Stochastic Processes (McGraw-Hill series in electrical engineering)*. McGraw-Hill Inc.,US, 1984. ISBN 0070484686.
- [98] Christopher C. Heyde. On a Property of the Lognormal Distribution. In *Selected Works of C.C. Heyde*, pages 16–18. Springer New York, 2010. doi: 10.1007/978-1-4419-5823-5\_6. URL [https://doi.org/10.1007/978-1-4419-5823-5\\_6](https://doi.org/10.1007/978-1-4419-5823-5_6).
- [99] Erich L. Lehmann and Joseph P. Romano. *Testing Statistical Hypotheses*. Springer New York, 2005. doi: 10.1007/0-387-27605-x. URL <https://link.springer.com/book/10.1007%2F0-387-27605-X>.
- [100] Ronald L. Wasserstein and Nicole A. Lazar. The ASA’s Statement on p-Values: Context, Process, and Purpose. *The American Statistician*, 70(2):129–133, apr 2016. doi: 10.1080/00031305.2016.1154108. URL <https://doi.org/10.1080/00031305.2016.1154108>.
- [101] Andrey Kolmogorov. Sulla Determinazione Empirica di Una Legge di Distribuzione. *Giornale dell’Istituto Italiano degli Attuari*, 4:83–91, 1933.

- [102] Nikolai Smirnov. On the estimation of the discrepancy between empirical curves of distribution for two independent samples. *Bulletin Mathématique de l'Université de Moscou*, 2, 1939.
- [103] Nikolai Smirnov. Table for Estimating the Goodness of Fit of Empirical Distributions. *The Annals of Mathematical Statistics*, 19(2):279–281, jun 1948. doi: 10.1214/aoms/1177730256. URL <https://doi.org/10.1214/aoms/1177730256>.
- [104] Donald A. Darling. The Kolmogorov-Smirnov, Cramer-von Mises Tests. *The Annals of Mathematical Statistics*, 28(4):823–838, dec 1957. doi: 10.1214/aoms/1177706788. URL <https://doi.org/10.1214/aoms/1177706788>.
- [105] Harald Cramér. On the composition of elementary errors. *Scandinavian Actuarial Journal*, 1928(1):13–74, jan 1928. doi: 10.1080/03461238.1928.10416862. URL <https://doi.org/10.1080/03461238.1928.10416862>.
- [106] Richard von Mises. *Wahrscheinlichkeit, Statistik und Wahrheit*. Julius Springer, 1928.
- [107] Erich L. Lehmann and George Casella. *Theory of Point Estimation (Springer Texts in Statistics)*. Springer, 2003. ISBN 0387985026. URL <https://link.springer.com/book/10.1007%2Fb98854>.
- [108] 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 <https://link.springer.com/book/10.1007%2Fb98673>.
- [109] Julius S. Bendat and Allan G. Piersol. *Random Data: Analysis and Measurement Procedures*. Wiley, 2010. ISBN 0470248777. URL <http://onlinelibrary.wiley.com/book/10.1002/9781118032428>.
- [110] Jerzy Neyman. Outline of a Theory of Statistical Estimation Based on the Classical Theory of Probability. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 236(767):333–380, aug 1937. doi: 10.1098/rsta.1937.0005. URL <https://doi.org/10.1098/rsta.1937.0005>.
- [111] Christopher A. Fuchs and Rüdiger Schack. 5 Unknown Quantum States and Operations, a Bayesian View. In *Quantum State Estimation*, pages 147–187. Springer Berlin Heidelberg, aug 2004. doi: 10.1007/978-3-540-44481-7\_5. URL [https://doi.org/10.1007/978-3-540-44481-7\\_5](https://doi.org/10.1007/978-3-540-44481-7_5).
- [112] Rink Hoekstra, Richard D. Morey, Jeffrey N. Rouder, and Eric-Jan Wagenmakers. Robust misinterpretation of confidence intervals. *Psychonomic Bulletin & Review*, 21(5):1157–1164, jan 2014. doi: 10.3758/s13423-013-0572-3. URL <https://doi.org/10.3758/s13423-013-0572-3>.

- [113] Edwin T. Jaynes. *Probability Theory: The Logic of Science*. Cambridge University Press, 2003. ISBN 0521592712.
- [114] Edwin T. Jaynes and Oscar Kempthorne. *Confidence Intervals vs Bayesian Intervals*, pages 175–257. Springer Netherlands, Dordrecht, 1976. ISBN 978-94-010-1436-6. doi: 10.1007/978-94-010-1436-6\_6. URL [https://doi.org/10.1007/978-94-010-1436-6\\_6](https://doi.org/10.1007/978-94-010-1436-6_6).
- [115] Ravitej Uppu, Tom A. W. Wolterink, Tristan B. H. Tentrup, and Pepijn W. H. Pinkse. Quantum optics of lossy asymmetric beam splitters. *Optics Express*, 24(15):16440, jul 2016. doi: 10.1364/oe.24.016440. URL <https://doi.org/10.1364/oe.24.016440>.
- [116] Anton Zeilinger. General properties of lossless beam splitters in interferometry. *American Journal of Physics*, 49(9):882–883, sep 1981. doi: 10.1119/1.12387. URL <https://doi.org/10.1119/1.12387>.
- [117] Bahaa E. A. Saleh and Malvin Carl Teich. *Fundamentals of Photonics*. Wiley-Interscience, 2007. ISBN 0471358320.
- [118] David N. Klyshko, Aleksey N. Penin, and Boris F. Polkovnikov. Parametric Luminescence and Light Scattering by Polaritons. *JETP Lett.*, 11:5, 1970. URL <https://doi.org/10.1103/physrevlett.25.84>.
- [119] David C. Burnham and Donald L. Weinberg. Observation of Simultaneity in Parametric Production of Optical Photon Pairs. *Physical Review Letters*, 25(2):84–87, jul 1970. doi: 10.1103/physrevlett.25.84. URL <https://doi.org/10.1103/physrevlett.25.84>.
- [120] Paul G. Kwiat, Klaus Mattle, Harald Weinfurter, Anton Zeilinger, Alexander V. Sergienko, and Yanhua Shih. New High-Intensity Source of Polarization-Entangled Photon Pairs. *Physical Review Letters*, 75(24):4337–4341, dec 1995. doi: 10.1103/physrevlett.75.4337. URL <https://doi.org/10.1103/physrevlett.75.4337>.
- [121] Christophe Couteau. Spontaneous parametric down-conversion. *Contemporary Physics*, 59(3):291–304, jul 2018. doi: 10.1080/00107514.2018.1488463. URL <https://doi.org/10.1080/00107514.2018.1488463>.
- [122] Eugene Hecht. *Optics (5th Edition)*. Addison-Wesley, 5 edition, 2002. ISBN 9783486588613.
- [123] 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>.

- [124] 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>.
- [125] Xing-Can Yao, Tian-Xiong Wang, Ping Xu, He Lu, Ge-Sheng Pan, Xiao-Hui Bao, Cheng-Zhi Peng, Chao-Yang Lu, Yu-Ao Chen, and Jian-Wei Pan. Observation of eight-photon entanglement. *Nature Photonics*, 6(4):225–228, feb 2012. doi: 10.1038/nphoton.2011.354. URL <https://doi.org/10.1038/nphoton.2011.354>.
- [126] Xi-Lin Wang, Luo-Kan Chen, W. Li, H.-L. Huang, C. Liu, C. Chen, Y.-H. Luo, Z.-E. Su, D. Wu, Z.-D. Li, H. Lu, Y. Hu, X. Jiang, C.-Z. Peng, L. Li, N.-L. Liu, Yu-Ao Chen, Chao-Yang Lu, and Jian-Wei Pan. Experimental Ten-Photon Entanglement. *Physical Review Letters*, 117(21), nov 2016. doi: 10.1103/physrevlett.117.210502. URL <https://doi.org/10.1103/physrevlett.117.210502>.
- [127] Matthias Bock, Andreas Lenhard, Christopher Chunnillall, and Christoph Becher. Highly efficient heralded single-photon source for telecom wavelengths based on a PPLN waveguide. *Optics Express*, 24(21):23992, oct 2016. doi: 10.1364/oe.24.023992. URL <https://doi.org/10.1364/oe.24.023992>.
- [128] Rolf Horn, Payam Abolghasem, Bhavin J. Bijlani, Dongpeng Kang, A. S. Helmy, and Gregor Weihs. Monolithic Source of Photon Pairs. *Physical Review Letters*, 108(15), apr 2012. doi: 10.1103/physrevlett.108.153605. URL <https://doi.org/10.1103/physrevlett.108.153605>.
- [129] Roman Schmied. Quantum state tomography of a single qubit: comparison of methods. *Journal of Modern Optics*, 63(18):1744–1758, feb 2016. doi: 10.1080/09500340.2016.1142018. URL <https://doi.org/10.1080/09500340.2016.1142018>.
- [130] Adam Miranowicz, Karol Bartkiewicz, Jan Peřina, Masato Koashi, Nobuyuki Imoto, and Franco Nori. Optimal two-qubit tomography based on local and global measurements: Maximal robustness against errors as described by condition numbers. *Physical Review A*, 90(6), dec 2014. doi: 10.1103/physreva.90.062123. URL <https://doi.org/10.1103/physreva.90.062123>.
- [131] 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](http://xqp.physik.uni-muenchen.de/publications/files/theses_phd/phd_kiesel.pdf).
- [132] Christian Schwemmer. Efficient Tomography and Entanglement Detection of Multiphoton States. *PhD Thesis, Ludwig-Maximilians-Universität München*, 2015. URL [http://xqp.physik.uni-muenchen.de/publications/files/theses\\_phd/phd\\_schwemmer.pdf](http://xqp.physik.uni-muenchen.de/publications/files/theses_phd/phd_schwemmer.pdf).

- [133] William K. Wootters and Brian D. Fields. Optimal state-determination by mutually unbiased measurements. *Annals of Physics*, 191(2):363–381, may 1989. doi: 10.1016/0003-4916(89)90322-9. URL [https://doi.org/10.1016/0003-4916\(89\)90322-9](https://doi.org/10.1016/0003-4916(89)90322-9).
- [134] 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>.
- [135] Jaroslav Řeháček, Berthold-Georg Englert, and Dagomir Kaszlikowski. Minimal qubit tomography. *Physical Review A*, 70(5), nov 2004. doi: 10.1103/physreva.70.052321. URL <https://doi.org/10.1103/physreva.70.052321>.
- [136] Joseph M. Renes, Robin Blume-Kohout, Andrew J. Scott, and Carlton M. Caves. Symmetric informationally complete quantum measurements. *Journal of Mathematical Physics*, 45(6):2171–2180, jun 2004. doi: 10.1063/1.1737053. URL <https://doi.org/10.1063/1.1737053>.
- [137] Christian Schwemmer, Géza Tóth, Alexander Niggebaum, Tobias Moroder, David Gross, Otfried Gühne, and Harald Weinfurter. Experimental Comparison of Efficient Tomography Schemes for a Six-Qubit State. *Physical Review Letters*, 113(4), jul 2014. doi: 10.1103/physrevlett.113.040503. URL <https://doi.org/10.1103/physrevlett.113.040503>.
- [138] Géza Tóth, Witłef Wieczorek, David Gross, Roland Krischek, Christian Schwemmer, and Harald Weinfurter. Permutationally Invariant Quantum Tomography. *Physical Review Letters*, 105(25), dec 2010. doi: 10.1103/physrevlett.105.250403. URL <https://doi.org/10.1103/physrevlett.105.250403>.
- [139] 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>.
- [140] Emmanuel J. Candes, Justin Romberg, and Terence Tao. Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information. *IEEE Transactions on Information Theory*, 52(2):489–509, feb 2006. doi: 10.1109/tit.2005.862083. URL <https://doi.org/10.1109/tit.2005.862083>.
- [141] 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>.
- [142] Carlos A. Riofrío, David Gross, Steven T. Flammia, Thomas Monz, Daniel Nigg, Rainer Blatt, and Jens Eisert. Experimental quantum compressed sensing for a

- seven-qubit system. *Nature Communications*, 8:15305, may 2017. doi: 10.1038/ncomms15305. URL <https://doi.org/10.1038/ncomms15305>.
- [143] Adrian Steffens, Carlos A. Riofrío, Will McCutcheon, Ingo Roth, Bryn A Bell, Alex McMillan, Mark S. Tame, John G. Rarity, and Jens Eisert. Experimentally exploring compressed sensing quantum tomography. *Quantum Science and Technology*, 2(2):025005, may 2017. doi: 10.1088/2058-9565/aa6ae2. URL <https://doi.org/10.1088/2058-9565/aa6ae2>.
- [144] Amir Kalev, Robert L. Kosut, and Ivan H. Deutsch. Quantum tomography protocols with positivity are compressed sensing protocols. *npj Quantum Information*, 1(1), dec 2015. doi: 10.1038/npjqi.2015.18. URL <https://doi.org/10.1038/npjqi.2015.18>.
- [145] Ulrich Schollwöck. The density-matrix renormalization group. *Reviews of Modern Physics*, 77(1):259–315, apr 2005. doi: 10.1103/revmodphys.77.259. URL <https://doi.org/10.1103/revmodphys.77.259>.
- [146] David Perez-Garcia, Frank Verstraete, Michael M. Wolf, and J. Ignacio Cirac. Matrix product state representations. *Quantum Information and Computation*, 7(5-6):401–430, 2007. URL [www.scopus.com](http://www.scopus.com).
- [147] Frank Verstraete, Valentin Murg, and J. Ignacio Cirac. Matrix product states, projected entangled pair states, and variational renormalization group methods for quantum spin systems. *Advances in Physics*, 57(2):143–224, mar 2008. doi: 10.1080/14789940801912366. URL <https://doi.org/10.1080/14789940801912366>.
- [148] Marcus Cramer, Martin B. Plenio, Steven T. Flammia, Rolando Somma, David Gross, Stephen D. Bartlett, Olivier Landon-Cardinal, David Poulin, and Yi-Kai Liu. Efficient quantum state tomography. *Nature Communications*, 1(9):149, dec 2010. doi: 10.1038/ncomms1147. URL <https://doi.org/10.1038/ncomms1147>.
- [149] Akio Fujiwara. Strong consistency and asymptotic efficiency for adaptive quantum estimation problems. *Journal of Physics A: Mathematical and General*, 39(40):12489–12504, sep 2006. doi: 10.1088/0305-4470/39/40/014. URL <https://doi.org/10.1088/0305-4470/39/40/014>.
- [150] Takanori Sugiyama, Peter S. Turner, and Mio Muraio. Effect of non-negativity on estimation errors in one-qubit state tomography with finite data. *New Journal of Physics*, 14(8):085005, aug 2012. doi: 10.1088/1367-2630/14/8/085005. URL <https://doi.org/10.1088/1367-2630/14/8/085005>.
- [151] Ryo Okamoto, Minako Iefuji, Satoshi Oyama, Koichi Yamagata, Hiroshi Imai, Akio Fujiwara, and Shigeki Takeuchi. Experimental Demonstration of Adaptive Quantum State Estimation. *Physical Review Letters*, 109(13), sep 2012. doi:

- 10.1103/physrevlett.109.130404. URL <https://doi.org/10.1103/physrevlett.109.130404>.
- [152] Dylan H. Mahler, Lee A. Rozema, Ardavan Darabi, Christopher Ferrie, Robin Blume-Kohout, and A. M. Steinberg. Adaptive Quantum State Tomography Improves Accuracy Quadratically. *Physical Review Letters*, 111(18), oct 2013. doi: 10.1103/physrevlett.111.183601. URL <https://doi.org/10.1103/physrevlett.111.183601>.
- [153] Konstantin S. Kravtsov, Stanislav S. Straupe, Igor V. Radchenko, Neil M. T. Houlby, Ferenc Huszár, and Sergei P. Kulik. Experimental adaptive Bayesian tomography. *Physical Review A*, 87(6), jun 2013. doi: 10.1103/physreva.87.062122. URL <https://doi.org/10.1103/physreva.87.062122>.
- [154] Daniel Lampert-Richart. Higher Dimensional Time-Energy Entanglement. *PhD Thesis, Ludwig-Maximilians-Universität München*, 2014. URL [http://xqp.physik.uni-muenchen.de/publications/files/theses\\_phd/phd\\_richart.pdf](http://xqp.physik.uni-muenchen.de/publications/files/theses_phd/phd_richart.pdf).
- [155] Gleb I. Struchalin, Ivan A. Pogorelov, Stanislav S. Straupe, Konstantin S. Kravtsov, Igor V. Radchenko, and Sergei P. Kulik. Experimental adaptive quantum tomography of two-qubit states. *Physical Review A*, 93(1), jan 2016. doi: 10.1103/physreva.93.012103. URL <https://doi.org/10.1103/physreva.93.012103>.
- [156] Qi Yin, Li Li, Xiao Xiang, Guo-Yong Xiang, Chuang-Feng Li, and Guang-Can Guo. Experimental demonstration of real-time adaptive one-qubit quantum-state tomography. *Physical Review A*, 95(1), jan 2017. doi: 10.1103/physreva.95.012129. URL <https://doi.org/10.1103/physreva.95.012129>.
- [157] Aonan Zhang, Yujie Zhang, Feixiang Xu, Long Li, and Lijian Zhang. Adaptive tomography of qubits: purity versus statistical fluctuations. *Journal of Physics A: Mathematical and Theoretical*, 51(39):395304, sep 2018. doi: 10.1088/1751-8121/aadad3. URL <https://doi.org/10.1088/1751-8121/aadad3>.
- [158] Christopher Ferrie. Self-Guided Quantum Tomography. *Physical Review Letters*, 113(19), nov 2014. doi: 10.1103/physrevlett.113.190404. URL <https://doi.org/10.1103/physrevlett.113.190404>.
- [159] Robert J. Chapman, Christopher Ferrie, and Alberto Peruzzo. Experimental Demonstration of Self-Guided Quantum Tomography. *Physical Review Letters*, 117(4), jul 2016. doi: 10.1103/physrevlett.117.040402. URL <https://doi.org/10.1103/physrevlett.117.040402>.
- [160] Michael I. Jordan and Tom M. Mitchell. Machine learning: Trends, perspectives, and prospects. *Science*, 349(6245):255–260, jul 2015. doi: 10.1126/science.aaa8415. URL <https://doi.org/10.1126/science.aaa8415>.

- [161] Francisco Albarrán-Arriagada, Juan C. Retamal, Enrique Solano, and Lucas Lamata. Measurement-based adaptation protocol with quantum reinforcement learning. *Physical Review A*, 98(4), oct 2018. doi: 10.1103/physreva.98.042315. URL <https://doi.org/10.1103/physreva.98.042315>.
- [162] Shang Yu, Francisco Albarran-Arriagada, Juan C. Retamal, Yi-Tao Wang, Wei Liu, Zhi-Jin Ke, Yu Meng, Zhi-Peng Li, Jian-Shun Tang, Enrique Solano, Lucas Lamata, Chuan-Feng Li, and Guang-Can Guo. Reconstruction of a Photonic Qubit State with Quantum Reinforcement Learning. *ArXiv e-prints*, August 2018.
- [163] Nico Klein. Multipartite entanglement detection and efficiency of state characterisation. *Master Thesis, Ludwig-Maximilians-Universität München*, 2015. URL [http://xqp.physik.uni-muenchen.de/publications/files/theses\\_master/master\\_klein.pdf](http://xqp.physik.uni-muenchen.de/publications/files/theses_master/master_klein.pdf).
- [164] Robin Blume-Kohout. Optimal, reliable estimation of quantum states. *New Journal of Physics*, 12(4):043034, apr 2010. doi: 10.1088/1367-2630/12/4/043034. URL <https://doi.org/10.1088/1367-2630/12/4/043034>.
- [165] Zdeněk Hradil. Quantum-state estimation. *Physical Review A*, 55(3):R1561–R1564, mar 1997. doi: 10.1103/physreva.55.r1561. URL <https://doi.org/10.1103/physreva.55.r1561>.
- [166] 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>.
- [167] Zhibo Hou, Han-Sen Zhong, Ye Tian, Daoyi Dong, Bo Qi, Li Li, Yuanlong Wang, Franco Nori, Guo-Yong Xiang, Chuan-Feng Li, and Guang-Can Guo. Full reconstruction of a 14-qubit state within four hours. *New Journal of Physics*, 18(8):083036, aug 2016. doi: 10.1088/1367-2630/18/8/083036. URL <https://doi.org/10.1088/1367-2630/18/8/083036>.
- [168] Jiangwei Shang, Zhengyun Zhang, and Hui Khoon Ng. Superfast maximum-likelihood reconstruction for quantum tomography. *Physical Review A*, 95(6), jun 2017. doi: 10.1103/physreva.95.062336. URL <https://doi.org/10.1103/physreva.95.062336>.
- [169] G. B. Silva, Scott Glancy, and Hilma M. Vasconcelos. Investigating bias in maximum-likelihood quantum-state tomography. *Physical Review A*, 95(2), feb 2017. doi: 10.1103/physreva.95.022107. URL <https://doi.org/10.1103/physreva.95.022107>.
- [170] Ana Predojevic, Gregor Weihs, and Maximilian Prilmüller. Analysis of (hyper-) entanglement in quantum dot systems. In Zameer U. Hasan, Philip R. Hemmer, Alan L. Migdall, and Alan E. Craig, editors, *Advances in Photonics of Quantum*

- Computing, Memory, and Communication XI*. SPIE, feb 2018. doi: 10.1117/12.2297228. URL <https://doi.org/10.1117/12.2297228>.
- [171] Adam C. Keith, Charles H. Baldwin, Scott Glancy, and E. Knill. Joint quantum-state and measurement tomography with incomplete measurements. *Physical Review A*, 98(4), oct 2018. doi: 10.1103/physreva.98.042318. URL <https://doi.org/10.1103/physreva.98.042318>.
- [172] Thomas Koshy. *Catalan Numbers with Applications*. Oxford University Press, 2008. ISBN 019533454X. URL <http://www.oxfordscholarship.com/view/10.1093/acprof:oso/9780195334548.001.0001/acprof-9780195334548-chapter-5>.
- [173] Lukas Knips, Christian Schwemmer, Nico Klein, Jonas Reuter, Geza Toth, and Harald Weinfurter. Multiqubit state tomography from a physical perspective. In *2017 Conference on Lasers and Electro-Optics Europe & European Quantum Electronics Conference (CLEO/Europe-EQEC)*. IEEE, jun 2017. doi: 10.1109/cleoe-eqec.2017.8087437. URL <https://doi.org/10.1109/cleoe-eqec.2017.8087437>.
- [174] 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](http://projecteuclid.org/DPubS/Repository/1.0/Disseminate?view=body&id=pdf_1&handle=euclid.ss/1177013815).
- [175] B. Efron and R. J. Tibshirani. *An Introduction to the Bootstrap*. 1993.
- [176] Christian Schwemmer, Lukas Knips, Daniel Richart, Harald Weinfurter, Tobias Moroder, Matthias Kleinmann, and Otfried Gühne. Systematic Errors in Current Quantum State Tomography Tools. *Physical Review Letters*, 114(8), feb 2015. doi: 10.1103/physrevlett.114.080403. URL <https://doi.org/10.1103/physrevlett.114.080403>.
- [177] Richard Kueng and Christopher Ferrie. Near-optimal quantum tomography: estimators and bounds. *New Journal of Physics*, 17(12):123013, dec 2015. doi: 10.1088/1367-2630/17/12/123013. URL <https://doi.org/10.1088/1367-2630/17/12/123013>.
- [178] Robin Blume-Kohout. Robust error bars for quantum tomography. *arXiv e-prints*, art. arXiv:1202.5270, Feb 2012.
- [179] 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>.
- [180] Philippe Faist and Renato Renner. Practical and Reliable Error Bars in Quantum Tomography. *Physical Review Letters*, 117(1), jul 2016. doi: 10.1103/physrevlett.117.010404. URL <https://doi.org/10.1103/physrevlett.117.010404>.

- [181] Jiangwei Shang, Hui Khoon Ng, Arun Sehrawat, Xikun Li, and Berthold-Georg Englert. Optimal error regions for quantum state estimation. *New Journal of Physics*, 15(12):123026, dec 2013. doi: 10.1088/1367-2630/15/12/123026. URL <https://doi.org/10.1088/1367-2630/15/12/123026>.
- [182] Xikun Li, Jiangwei Shang, Hui Khoon Ng, and Berthold-Georg Englert. Optimal error intervals for properties of the quantum state. *Physical Review A*, 94(6), dec 2016. doi: 10.1103/physreva.94.062112. URL <https://doi.org/10.1103/physreva.94.062112>.
- [183] Daniel Suess, Łukasz Rudnicki, Thiago O. Maciel, and David Gross. Error regions in quantum state tomography: computational complexity caused by geometry of quantum states. *New Journal of Physics*, 19(9):093013, sep 2017. doi: 10.1088/1367-2630/aa7ce9. URL <https://doi.org/10.1088/1367-2630/aa7ce9>.
- [184] Erik M. Gauger, Elisabeth Rieper, John J. L. Morton, Simon C. Benjamin, and Vlatko Vedral. Sustained Quantum Coherence and Entanglement in the Avian Compass. *Physical Review Letters*, 106(4), jan 2011. doi: 10.1103/physrevlett.106.040503. URL <https://doi.org/10.1103/physrevlett.106.040503>.
- [185] K. Birgitta Whaley, Mohan Sarovar, and Akihito Ishizaki. Quantum entanglement phenomena in photosynthetic light harvesting complexes. *Procedia Chemistry*, 3(1):152–164, 2011. doi: 10.1016/j.proche.2011.08.021. URL <https://doi.org/10.1016/j.proche.2011.08.021>.
- [186] Lukas Knips, Christian Schwemmer, Nico Klein, Marcin Wieśniak, and Harald Weinfurter. Multipartite Entanglement Detection with Minimal Effort. In *2015 Conference on Lasers and Electro-Optics Europe & European Quantum Electronics Conference (CLEO/Europe-EQEC)*, volume 117. IEEE, 2015.
- [187] Piotr Badziąg, Āaslav 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>.
- [188] 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>.
- [189] Marcin Wieśniak and Koji Maruyama. Package of facts and theorems for efficiently generating entanglement criteria for many qubits. *Physical Review A*, 85(6), jun 2012. doi: 10.1103/physreva.85.062315. URL <https://doi.org/10.1103/physreva.85.062315>.

- [190] Minh Cong Tran, Borivoje Dakić, Wiesław Laskowski, and Tomasz Paterek. Correlations between outcomes of random measurements. *Physical Review A*, 94(4), oct 2016. doi: 10.1103/physreva.94.042302. URL <https://doi.org/10.1103/physreva.94.042302>.
- [191] Aleksandra Dimić and Borivoje Dakić. Single-copy entanglement detection. *npj Quantum Information*, 4(1), feb 2018. doi: 10.1038/s41534-017-0055-x. URL <https://doi.org/10.1038/s41534-017-0055-x>.
- [192] Jasmin D. A. Meinecke, Pete Shadbolt, Lukas Knips, Jan Dziewior, Nicolas Brunner, Joseph Bowles, Flavien Hirsch, Jeremy L. O’Brien, and Harald Weinfurter. Noise-powered entanglement detection. In *2017 Conference on Lasers and Electro-Optics Europe & European Quantum Electronics Conference (CLEO/Europe-EQEC)*. IEEE, jun 2017. doi: 10.1109/cleoe-eqec.2017.8087356. URL <https://doi.org/10.1109/cleoe-eqec.2017.8087356>.
- [193] Julio I. de Vicente and Marcus Huber. Multipartite entanglement detection from correlation tensors. *Physical Review A*, 84(6), dec 2011. doi: 10.1103/physreva.84.062306. URL <https://doi.org/10.1103/physreva.84.062306>.
- [194] Andreas Gabriel, Łukasz Rudnicki, and Beatrix C Hiesmayr. Measurement-base-independent test for genuine multipartite entanglement. *New Journal of Physics*, 15(7):073033, jul 2013. doi: 10.1088/1367-2630/15/7/073033. URL <https://doi.org/10.1088/1367-2630/15/7/073033>.
- [195] Minh Cong Tran, Borivoje Dakić, François Arnault, Wiesław Laskowski, and Tomasz Paterek. Quantum entanglement from random measurements. *Physical Review A*, 92(5), nov 2015. doi: 10.1103/physreva.92.050301. URL <https://doi.org/10.1103/physreva.92.050301>.
- [196] Ingemar Bengtsson and Karol Zyczkowski. *Geometry of Quantum States: An Introduction to Quantum Entanglement*. Cambridge University Press, January 2008. ISBN 052189140X.
- [197] 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>.
- [198] Leah Henderson and Vlatko Vedral. Classical, quantum and total correlations. *Journal of Physics A: Mathematical and General*, 34(35):6899–6905, aug 2001. doi: 10.1088/0305-4470/34/35/315. URL <https://doi.org/10.1088/0305-4470/34/35/315>.
- [199] Charles H. Bennett, Andrzej Grudka, Michał Horodecki, Paweł Horodecki, and Ryszard Horodecki. Postulates for measures of genuine multipartite correlations.

- Physical Review A*, 83(1), jan 2011. doi: 10.1103/physreva.83.012312. URL <https://doi.org/10.1103/physreva.83.012312>.
- [200] Gian Luca Giorgi, Bruno Bellomo, Fernando Galve, and Roberta Zambrini. Genuine Quantum and Classical Correlations in Multipartite Systems. *Physical Review Letters*, 107(19), nov 2011. doi: 10.1103/physrevlett.107.190501. URL <https://doi.org/10.1103/physrevlett.107.190501>.
- [201] Vlatko Vedral. The role of relative entropy in quantum information theory. *Reviews of Modern Physics*, 74(1):197–234, mar 2002. doi: 10.1103/revmodphys.74.197. URL <https://doi.org/10.1103/revmodphys.74.197>.
- [202] Alfred Wehrl. General properties of entropy. *Reviews of Modern Physics*, 50(2): 221–260, apr 1978. doi: 10.1103/revmodphys.50.221. URL <https://doi.org/10.1103/revmodphys.50.221>.
- [203] Duanlu L. Zhou, Bei Zeng, Z. Xu, and Lixing You. Multiparty correlation measure based on the cumulant. *Physical Review A*, 74(5), nov 2006. doi: 10.1103/physreva.74.052110. URL <https://doi.org/10.1103/physreva.74.052110>.
- [204] Waldemar Klobus, Wiesław Laskowski, Tomasz Paterek, Marcin Wieśniak, and Harald Weinfurter. Higher dimensional entanglement without correlations. *ArXiv e-prints*, August 2018. URL <https://arxiv.org/abs/1808.10201>.
- [205] 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](http://xqp.physik.uni-muenchen.de/publications/files/theses_phd/phd_krischek.pdf).
- [206] Paul A. M. Dirac. *The Principles of Quantum Mechanics*. Oxford University Press, 2nd edition edition, 1935.
- [207] Yakir Aharonov, David Z. Albert, and Lev Vaidman. How the result of a measurement of a component of the spin of a spin-1/2 particle can turn out to be 100. *Physical Review Letters*, 60(14):1351–1354, apr 1988. doi: 10.1103/physrevlett.60.1351. URL <https://doi.org/10.1103/physrevlett.60.1351>.
- [208] Abraham G. Kofman, Sahel Ashhab, and Franco Nori. Nonperturbative theory of weak pre- and post-selected measurements. *Physics Reports*, 520(2):43–133, nov 2012. doi: 10.1016/j.physrep.2012.07.001. URL <https://doi.org/10.1016/j.physrep.2012.07.001>.
- [209] Jan Dziewior. The Concept of Weak Values. *Master Thesis, Ludwig-Maximilians-Universität München*, 2016. URL [http://xqp.physik.uni-muenchen.de/publications/files/theses\\_master/master\\_dziewior.pdf](http://xqp.physik.uni-muenchen.de/publications/files/theses_master/master_dziewior.pdf).

- [210] Boaz Tamir and Eliahu Cohen. Introduction to Weak Measurements and Weak Values. *Quanta*, 2(1):7, may 2013. doi: 10.12743/quanta.v2i1.14. URL <https://doi.org/10.12743/quanta.v2i1.14>.
- [211] Eliahu Cohen and Eli Pollak. Determination of weak values of quantum operators using only strong measurements. *Physical Review A*, 98(4), oct 2018. doi: 10.1103/physreva.98.042112. URL <https://doi.org/10.1103/physreva.98.042112>.
- [212] Juan P. Torres and Luis José Salazar-Serrano. Weak value amplification: a view from quantum estimation theory that highlights what it is and what isn't. *Scientific Reports*, 6(1), feb 2016. doi: 10.1038/srep19702. URL <https://doi.org/10.1038/srep19702>.
- [213] Mira Weißl. Properties of Weak Values. *Master Thesis, Ludwig-Maximilians-Universität München*, 2016. URL [http://xqp.physik.uni-muenchen.de/publications/files/theses\\_master/master\\_weissl.pdf](http://xqp.physik.uni-muenchen.de/publications/files/theses_master/master_weissl.pdf).
- [214] Katharina Senkalla. Weak value description of an Interferometer and its Application. *Master Thesis, Ludwig-Maximilians-Universität München*, 2018. URL [http://xqp.physik.uni-muenchen.de/publications/theses\\_master/master\\_senkalla.html](http://xqp.physik.uni-muenchen.de/publications/theses_master/master_senkalla.html).
- [215] Onur Hosten and Paul Kwiat. Observation of the Spin Hall Effect of Light via Weak Measurements. *Science*, 319(5864):787–790, feb 2008. doi: 10.1126/science.1152697. URL <https://doi.org/10.1126/science.1152697>.
- [216] P. Ben Dixon, David J. Starling, Andrew N. Jordan, and John C. Howell. Ultrasensitive Beam Deflection Measurement via Interferometric Weak Value Amplification. *Physical Review Letters*, 102(17), apr 2009. doi: 10.1103/physrevlett.102.173601. URL <https://doi.org/10.1103/physrevlett.102.173601>.
- [217] Jeff S. Lundeen, Brandon Sutherland, Aabid Patel, Corey Stewart, and Charles Bamber. Direct measurement of the quantum wavefunction. *Nature*, 474(7350):188–191, jun 2011. doi: 10.1038/nature10120. URL <https://doi.org/10.1038/nature10120>.
- [218] Jeff Z. Salvail, Megan Agnew, Allan S. Johnson, Eliot Bolduc, Jonathan Leach, and Robert W. Boyd. Full characterization of polarization states of light via direct measurement. *Nature Photonics*, 7(4):316–321, mar 2013. doi: 10.1038/nphoton.2013.24. URL <https://doi.org/10.1038/nphoton.2013.24>.
- [219] Mehul Malik and Robert W. Boyd. Quantum imaging technologies. *La Rivista del Nuovo Cimento*, 37(5):273–332, Jun 2014. ISSN 0393-697X. doi: 10.1393/ncr/i2014-10100-0. URL <http://doi.org/10.1393/ncr/i2014-10100-0>.

- [220] Mehul Malik, Mohammad Mirhosseini, Martin P. J. Lavery, Jonathan Leach, Miles J. Padgett, and Robert W. Boyd. Direct measurement of a 27-dimensional orbital-angular-momentum state vector. *Nature Communications*, 5(1), jan 2014. doi: 10.1038/ncomms4115. URL <https://doi.org/10.1038/ncomms4115>.
- [221] Guillaume S. Thekkadath, Lambert Giner, Yamn Chalich, Matthew J. Horton, Jash Banker, and Jeff S. Lundeen. Direct Measurement of the Density Matrix of a Quantum System. *Physical Review Letters*, 117(12), sep 2016. doi: 10.1103/physrevlett.117.120401. URL <https://doi.org/10.1103/physrevlett.117.120401>.
- [222] Lupei Qin, Luting Xu, Wei Feng, and Xin-Qi Li. Qubit state tomography in a superconducting circuit via weak measurements. *New Journal of Physics*, 19(3):033036, mar 2017. doi: 10.1088/1367-2630/aa646e. URL <https://doi.org/10.1088/1367-2630/aa646e>.
- [223] Lev Vaidman. Weak value controversy. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 375(2106):20160395, oct 2017. doi: 10.1098/rsta.2016.0395. URL <https://doi.org/10.1098/rsta.2016.0395>.
- [224] Christopher Ferrie and Joshua Combes. How the Result of a Single Coin Toss Can Turn Out to be 100 Heads. *Physical Review Letters*, 113(12), sep 2014. doi: 10.1103/physrevlett.113.120404. URL <https://doi.org/10.1103/physrevlett.113.120404>.
- [225] Aharon Brodutch. Comment on “How the Result of a Single Coin Toss Can Turn Out to be 100 Heads”. *Physical Review Letters*, 114(11), mar 2015. doi: 10.1103/physrevlett.114.118901. URL <https://doi.org/10.1103/physrevlett.114.118901>.
- [226] Christopher Ferrie and Joshua Combes. Ferrie and Combes Reply:. *Physical Review Letters*, 114(11), mar 2015. doi: 10.1103/physrevlett.114.118902. URL <https://doi.org/10.1103/physrevlett.114.118902>.
- [227] Alessandro Romito, Andrew N. Jordan, Yakir Aharonov, and Yuval Gefen. Weak values are quantum: you can bet on it. *Quantum Studies: Mathematics and Foundations*, 3(1):1–4, jan 2016. doi: 10.1007/s40509-015-0069-z. URL <https://doi.org/10.1007/s40509-015-0069-z>.
- [228] Christopher Ferrie and Joshua Combes. Weak Value Amplification is Suboptimal for Estimation and Detection. *Physical Review Letters*, 112(4), jan 2014. doi: 10.1103/physrevlett.112.040406. URL <https://doi.org/10.1103/physrevlett.112.040406>.
- [229] Xuannmin Zhu and Yu-Xiang Zhang. Influence of environmental noise on the weak value amplification. *Quantum Information Processing*, 15(8):3421–3441,

- may 2016. doi: 10.1007/s11128-016-1340-x. URL <https://doi.org/10.1007/s11128-016-1340-x>.
- [230] Jérémie Harris, Robert W. Boyd, and Jeff S. Lundeen. Weak Value Amplification Can Outperform Conventional Measurement in the Presence of Detector Saturation. *Physical Review Letters*, 118(7), feb 2017. doi: 10.1103/physrevlett.118.070802. URL <https://doi.org/10.1103/physrevlett.118.070802>.
- [231] Josiah Sinclair, Matin Hallaji, Aephraim M. Steinberg, Jeff Tollaksen, and Andrew N. Jordan. Weak-value amplification and optimal parameter estimation in the presence of correlated noise. *Physical Review A*, 96(5), nov 2017. doi: 10.1103/physreva.96.052128. URL <https://doi.org/10.1103/physreva.96.052128>.
- [232] George C. Knee, Joshua Combes, Christopher Ferrie, and Erik M. Gauger. Weak-value amplification: state of play. *Quantum Measurements and Quantum Metrology*, 3(1), jan 2016. doi: 10.1515/qmetro-2016-0006. URL <https://doi.org/10.1515/qmetro-2016-0006>.
- [233] Matin Hallaji, Amir Feizpour, Greg Dmochowski, Josiah Sinclair, and Aephraim M. Steinberg. Weak-value amplification of the nonlinear effect of a single photon. *Nature Physics*, 13:540, February 2017. URL <http://dx.doi.org/10.1038/nphys4040>.
- [234] Anna-Lena K. Hashagen and Michael M. Wolf. Universality and optimality in the information–disturbance tradeoff. *Annales Henri Poincaré*, 20(1):219–258, October 2018. doi: 10.1007/s00023-018-0724-0. URL <https://doi.org/10.1007/s00023-018-0724-0>.
- [235] Erwin Schrödinger. About Heisenberg uncertainty relation. In *Sitzungsber. Preuss. Akad. Wiss. Berlin (Math. Phys) 19, 296–303 (see also, arXiv: quant-ph/9903100)*, 1930.
- [236] E. H. Kennard. Zur Quantenmechanik einfacher Bewegungstypen. *Zeitschrift für Physik*, 44(4-5):326–352, apr 1927. doi: 10.1007/bf01391200. URL <https://doi.org/10.1007/bf01391200>.
- [237] Hermann Weyl. *Gruppentheorie und Quantenmechanik*. Leipzig, Leipzig, 1928.
- [238] Howard P. Robertson. The Uncertainty Principle. *Physical Review*, 34(1):163–164, jul 1929. doi: 10.1103/physrev.34.163. URL <https://doi.org/10.1103/physrev.34.163>.
- [239] Masanao Ozawa. Universally valid reformulation of the Heisenberg uncertainty principle on noise and disturbance in measurement. *Phys. Rev. A*, 67:042105, Apr 2003. doi: 10.1103/PhysRevA.67.042105. URL <http://link.aps.org/doi/10.1103/PhysRevA.67.042105>.

- 
- [240] Werner Heisenberg. *The Physical Principles of the Quantum Theory*. University of Chicago Press, 1930.
- [241] Stephen Boughn and Marcel Reginatto. Another look through Heisenberg's microscope. *European Journal of Physics*, 39(3):035402, mar 2018. doi: 10.1088/1361-6404/aaa33f. URL <https://doi.org/10.1088/1361-6404/aaa33f>.
- [242] Lee A. Rozema, Ardavan Darabi, Dylan H. Mahler, Alex Hayat, Yasaman Soudagar, and Aephraim M. Steinberg. Violation of Heisenberg's Measurement-Disturbance Relationship by Weak Measurements. *Physical Review Letters*, 109(10), sep 2012. doi: 10.1103/physrevlett.109.100404. URL <https://doi.org/10.1103/physrevlett.109.100404>.
- [243] Francesco Buscemi, Michael J. W. Hall, Masanao Ozawa, and Mark M. Wilde. Noise and Disturbance in Quantum Measurements: An Information-Theoretic Approach. *Phys. Rev. Lett.*, 112:050401, Feb 2014. doi: 10.1103/PhysRevLett.112.050401. URL <https://link.aps.org/doi/10.1103/PhysRevLett.112.050401>.
- [244] Cyril Branciard. Error-tradeoff and error-disturbance relations for incompatible quantum measurements. *Proceedings of the National Academy of Sciences*, 110(17):6742–6747, apr 2013. doi: 10.1073/pnas.1219331110. URL <https://doi.org/10.1073/pnas.1219331110>.
- [245] William K. Wootters and Wojciech H. Zurek. A single quantum cannot be cloned. *Nature*, 299(5886):802–803, oct 1982. doi: 10.1038/299802a0. URL <https://doi.org/10.1038/299802a0>.
- [246] Heng Fan, Yi-Nan Wang, Li Jing, Jie-Dong Yue, Han-Duo Shi, Yong-Liang Zhang, and Liang-Zhu Mu. Quantum cloning machines and the applications. *Physics Reports*, 544(3):241–322, nov 2014. doi: 10.1016/j.physrep.2014.06.004. URL <https://doi.org/10.1016/j.physrep.2014.06.004>.
- [247] Valerio Scarani, Sofyan Iblisdir, Nicolas Gisin, and Antonio Acín. Quantum cloning. *Reviews of Modern Physics*, 77(4):1225–1256, nov 2005. doi: 10.1103/revmodphys.77.1225. URL <https://doi.org/10.1103/revmodphys.77.1225>.
- [248] Anna-Lena Hashagen. Universal asymmetric quantum cloning revisited. *Quantum Information and Computation*, 17(9&10):0747–0778, 2017. doi: 10.26421/qic17.9-10. URL <http://www.rintonpress.com/xxqic17/qic-17-910/0747-0778.pdf>.
- [249] Juan Carlos Garcia-Escartin and Pedro Chamorro-Posada. A SWAP gate for qudits. *Quantum Information Processing*, 12(12):3625–3631, aug 2013. doi: 10.1007/s11128-013-0621-x. URL <https://doi.org/10.1007/s11128-013-0621-x>.



# Acknowledgments

I would like to express my gratitude to Harald Weinfurter, not only for supervising my thesis, but for investing so much time and effort in my work and our common projects. Thank you for your long-lasting support!

Most of the results presented in this thesis could only be achieved thanks to the great team work I have enjoyed during my time here. There, I want to thank foremost my (former) fellow PhD students Christian Schwemmer and Jan Dziewior, with whom both I had a fruitful time. Christian, I am very grateful for your guidance and have appreciated our common time a lot. Jan, it was a pleasure to work with you. You helped me to evolve, in so many ways. Thank you! I also want to thank Jasmin Meinecke for our scientific discussions, for the opportunity to share my thoughts with, and for providing a comfortable working atmosphere. Furthermore, I am beholden to my former colleagues Daniel Richart and Nico Klein, with whom both, on very different ways, I enjoyed to work and discuss with. That was really a nice time. Finally, I want to thank Mira Weißl, Katharina Senkalla, Maria Galli, and Leonardo Ruscio, who helped to make our group such a nice working place.

If further advice is needed, if something does not work, if some effect cannot be explained, there is one person to go. Wenjamin Rosenfeld is the one to make things happen. Thank you so much! Furthermore, I want to thank all the other former and current members of the group, Daniel Burchardt, Peter Freiwang, Robert Garthoff, Michael Krug, Tim van Leent, Lars Liebermeister, Leonhard Mayr, Sebastian Nauerth, Norbert Ortegell, Markus Rau, Kai Redeker, Daniel Schlenk, Gwenaëlle Vest, Markus Weber, Martin Zeitlmair, and Wei Zhang.

This thesis is based on multiple collaborative projects. I'd like to express my gratitude to all people who helped to obtain those results. Although I cannot list all here, I want to thank Otfried Gühne, Anna-Lena Hashagen, Matthias Kleinmann, Waldemar Kłobus, Wiesław Laskowski, Tobias Moroder, Tomasz Paterek, Jonas Reuter, Géza Tóth, Lev Vaidman, Marcin Wiesniak, and Michael Wolf. Further thanks go to Thomas Schulte-Herbrüggen for making many things possible.

I would like to thank my parents, Eva-Maria and Udo Knips, for their support and encouragement. Thank you, Alex, for your assistance in so many different levels, from proofreading to almost-instantaneous support for technical questions. And of course for all the rest.

Finally, Karen, I am thankful for all your counsel, your consolation and your love. I am deeply grateful for *you*.

# Index

- adaptive state tomography, 54
- alignment, 188
- Anderson-Darling test, 36, 85
- anomalous weak values, 184
- anticommutation relation, 106
- antistate, 148, 154, 172
  
- Bayes theorem, 37
- BBO, 45, 113, 154
- beam splitter, 43
- Bell inequality, 19, 105
- Bell state, 13, 24, 108, 133
- bias, 60, 61, 85
- Bloch sphere, 7, 56
- bootstrapping, 61, 79
- Born's rule, 29
- Bures angle, 22
- Bures distance, 22
  
- Catalan numbers, 73, 85
- Cauchy distribution, 33
- channel state duality, 30
- CHSH inequality, 20
- cluster state, 26, 113, 126, 140
- coherences, 9
- coherent swap, 195
- commutation relation, 106
- compressed sensing, 53
- condition number, 51
- confidence, 36
- confidence interval, 39
- correlation, 11, 148, 151, 154, 172
- Cramér-von Mises test, 36
- credibility interval, 40
- crossing partitions, 74
- cumulative density function, 32, 85, 154
  
- density matrix, 8, 11
- Dicke state, 26, 53, 113, 154
- direct (state) tomography, 186
  
- eigenvalue, 72
- Einstein-Podolsky-Rosen, 19
- empirical density function, 33, 85, 154
- entanglement, 13, 104, 113, 140, 148, 154, 172
- entanglement depth, 15
- entanglement intactness, 14
- entanglement witness, 17, 105, 108, 113, 140
  
- fidelity, 20, 60, 194
- Frobenius norm, 23
  
- genuine multipartite entanglement, *see* entanglement
- Greenberger-Horne-Zeilinger state, 25, 27, 53, 106, 110, 113, 126, 140
  
- Haar measure, 133
- hedged maximum likelihood estimation, 57
- Heisenberg microscope, 190
- Heisenberg's uncertainty relation, 189, 198
- hidden variables, 19
- Hilbert-Schmidt norm, 22
- HWP, *see* waveplate
- hypothesis test, 34
  
- Kolmogorov-Smirnov test, 35, 154
- kurtosis, 33
  
- likelihood, 36, 56, 61
- local interaction, 188
- LOCC, 16
  
- machine learning, 54

- Matrix Product State tomography, 54  
maximally mixed state, 24, 72, 109, 110  
maximum likelihood, 38, 56, 61, 85  
measurement-disturbance relation, 191  
measurement-disturbance tradeoff, 198  
moment generating function, 34, 85  
Monte Carlo, 80  
multiqubit system, 10, 124
- no-cloning theorem, 192  
no-correlation, 148, 154, 172  
non-crossing partitions, 74
- one-way quantum computing, 25  
overcomplete Pauli tomography, 50  
Ozawa uncertainty relation, 191
- p value, 34, 85  
partial trace, 12, 198  
Pauli matrix, 8  
phase matching, 44  
PI tomography, 53  
Poincaré sphere, 8  
point estimator, 38, 61  
pointer, 182  
POM, *see* POVM  
populations, 9  
positive partial transpose, 17, 105  
post-measurement state, 30, 198  
postselection, 183  
POVM, 30, 198  
predictability, 150  
probability density function, 32  
producibility, 14  
product state, 10, 133  
purification, 13  
purity, 135
- quantum channel, 30, 198  
quantum cloning, 192  
quantum instrument, 31, 198  
quantum measurement, 28, 198  
quantum state tomography, 48, 61, 113, 198  
quantum teleportation, 24
- qubit, 6  
qudit, 11, 27, 124  
ququad, 11  
ququart, *see* ququad  
qutrit, 11, 27, 124, 126
- random measurements, 131, 140  
rebit, 59  
Robertson uncertainty relation, 189
- Schatten-1 norm, 23  
Schatten-2 norm, 23  
Schur norm, 23  
self-guided tomography, 54  
separability, 14  
shaking hands, 74  
SIC POVM, 52  
significance, 36  
skewness, 33  
SLOCC, 16  
smallest credibility region, 40  
spatial projection, 188  
spont. param. down conversion, 44, 113, 154  
statistical moments, 33, 85, 131  
super-radiance, 27
- tomography, *see* quantum state tomography  
trace distance, 23, 198
- uncertainty relation, 190  
unphysical estimate, 55
- W state, 154  
waveplate, 41  
weak amplification, 186, 188  
weak coupling, 183  
weak measurements, 183  
weak value, 183, 188  
white noise, *see* maximally mixed state  
Wigner semicircle distribution, 72, 85  
worst-case total variational distance, 192  
worst-case trace norm distance, 192
- YVO<sub>4</sub>, 43