
Efficient Tomography and Entanglement Detection of Multiphoton States

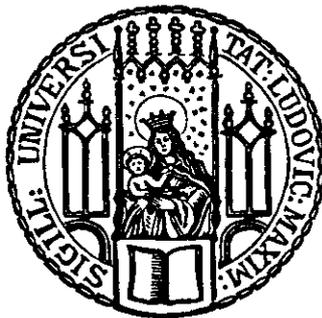
Christian Schwemmer



München, August 2015

Efficient Tomography and Entanglement Detection of Multiphoton States

Christian Schwemmer



München, August 2015

Efficient Tomography and Entanglement Detection of Multiphoton States

Christian Schwemmer

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

vorgelegt von
Christian Michael Schwemmer
aus Förrenbach

München, August 2015

Erstgutachter: Prof. Dr. Weinfurter
Zweitgutachter: Prof. Dr. Belén Paredes
Tag der mündlichen Prüfung: 30.9.2015

Für meine Mutter.



Landscape with Blind Orion Searching for the Sun, Nicolas Poussin, 1658.

Ist der Zwerg auf den Schultern des Riesen nicht
immer größer, als der Riese selbst?

Johann Gottfried von Herder,
Abhandlung über den Ursprung der Sprache

Contents

Abstract	xvii
Zusammenfassung	xix
1. Introduction	1
2. Fundamentals of quantum information theory	5
2.1. Single qubits	6
2.1.1. From classical bits to quantum bits	6
2.1.2. Single-qubit manipulation and discrimination	9
2.2. Multiqubit states and entanglement	10
2.2.1. Multiple qubits	10
2.2.2. Classification of multiqubit states	12
2.2.3. Entanglement detection	17
2.2.4. Entanglement measures	21
2.3. Experimental realization of photonic multiqubit states	25
2.3.1. Spontaneous parametric down conversion	25
2.3.2. Linear optical setups	29
3. Entanglement detection by correlation measurements	35
3.1. Optimized state-independent entanglement detection	36
3.1.1. Experimentally friendly entanglement criterion	36
3.1.2. Experimental Schmidt decomposition	38
3.1.3. Decision tree	40
3.2. No correlation states	42
3.2.1. States and antistates	43
3.2.2. Family of genuine entangled no correlation states	44
3.3. Publications	48
P3.1. Experimental Schmidt Decomposition and State Independent Entanglement Detection	50
P3.2. Optimized state-independent entanglement detection based on a geometrical threshold criterion	57
P3.3. Genuine Multipartite Entanglement without Multipartite Correlations	68

Contents

4. Quantum state tomography	85
4.1. Standard tomography	86
4.1.1. Projector-based complete tomography	87
4.1.2. Pauli tomography scheme	89
4.1.3. Physicality constraint	90
4.1.4. Comparison of algorithms	93
4.2. Partial tomography schemes	95
4.2.1. Compressed sensing	95
4.2.2. Permutationally invariant tomography	96
4.2.3. Compressed sensing in the permutationally invariant subspace	102
4.3. Publications	102
P4.1. Permutationally Invariant Quantum Tomography	104
P4.2. Experimental Comparison of Efficient Tomography Schemes for a Six-Qubit State	112
P4.3. Permutationally invariant state reconstruction	120
5. Systematic errors of standard quantum state estimation	145
5.1. Origin of unphysical density matrices	146
5.2. Constrained state estimation	149
5.2.1. Flaws of constrained state estimation	149
5.2.2. Biasedness and physicality	152
5.3. Publications	153
P5.1. Systematic Errors in Current Quantum State Tomography Tools	155
6. Conclusions	163
A. Appendix	167
A.1. Additional remarks on convex optimization	167
A.1.1. Maximum likelihood	167
A.1.2. Least squares	168
A.1.3. Barrier term	169
A.1.4. Implementation	169
A.2. Measurement settings for PI tomography	169
Publication list	173
Bibliography	177
Acknowledgments	201
Curriculum vitae	203

List of Figures

2.1. Bloch sphere	7
2.2. SLOCC classes for three qubits	14
2.3. Graph states	16
2.4. Four-qubit symmetric Dicke states	17
2.5. Entanglement witnesses	20
2.6. Geometric measure of entanglement	24
2.7. Spontaneous parametric down conversion	27
2.8. Photon source based on a UV enhancement cavity	29
2.9. Building blocks of linear setups	32
3.1. Schmidt decomposition	39
3.2. Schmidt decomposition in practice	40
3.3. Two-qubit decision tree	41
3.4. Efficiency of the two-qubit decision tree	42
3.5. Entanglement of no-correlation states	45
3.6. Vanishing correlations and predictability	46
4.1. Convergence of fit algorithms	94
4.2. Convergence of compressed sensing	96
4.3. Number of PI correlations	97
4.4. Measurement operators for PI tomography	99
4.5. Block diagonal decomposition of PI operators	101
5.1. Unphysical result due to an odd number of counts	146
5.2. Unphysicality and statistics	148
5.3. Biased fidelity estimation	150
5.4. Overestimation of negativity	151
5.5. Effect of fitting	151
5.6. Biasedness of constrained state estimation	152

List of Figures

List of Tables

2.1. Angle settings for Pauli measurements	33
A.1. Optimized settings for four-qubit PI tomography	170
A.2. Randomly chosen settings for four-qubit PI tomography	170
A.3. Optimized settings for six-qubit PI tomography	171

List of Tables

Abbreviations

APD	avalanche photo diode
BBO	β -barium-borate
BS	beam splitter
CS	compressed sensing
FPGA	field programmable gate array
FWHM	full width at half maximum
GHZ	Greenberger-Horne-Zeilinger
LBO	lithium-triborate
LHV	local hidden variable
LOCC	local operations and classical communication
LU	local unitary
HWP	half-wave plate
PBS	polarizing beam splitter
PI	permutationally invariant
POVM	positive operator valued measure
PPT	positive partial transposition
PVM	projector valued measure
QWP	quarter-wave plate
SHG	second harmonic generation
SLOCC	statistical local operations and classical communication
SPDC	spontaneous parametric down-conversion
Ti:Sa	titanium:sapphire
UV	ultraviolet
YVO	yttrium-vanadate

Abbreviations

Abstract

Quantum mechanics is without doubt one of the most successful theories in the history of science. Yet, due to its counter intuitive predictions, it still challenges many physicists. In this context, as a consequence of the principle of superposition, especially the phenomenon of entanglement has to be named. Although entanglement lacks any classical interpretation, it is nonetheless at the heart of novel methods for data processing like quantum communication or quantum computation. In an interdisciplinary effort of mathematics, computer science and physics, researchers from the respective fields try to better understand and characterize multipartite entanglement and to pave the way for its practical application.

In this work, the focus is on the development of appropriate tools to analyze and characterize multipartite entangled states as well as their experimental implementation by means of polarization encoded multiphoton states. More precisely, Bell states, and especially symmetric Dicke states with up to six photons are investigated. In order to generate these states, photon sources based on the process of spontaneous parametric down-conversion are used.

In the first part of this work, two schemes that allow to detect entanglement of two- and multiqubit states with possibly few correlation measurements are presented and experimentally implemented. In the first scheme, the state is transformed into the Schmidt basis such that entanglement can be verified with at most three correlation measurements. The second scheme makes use of the principle of correlation complementarity in order to perform only those measurements which are most advantageous for entanglement detection. The scheme can be represented in a compact form called decision tree which determines the next measurement to be carried out depending on the result of the previous measurement. In another experiment, it is investigated what kind of correlations are typical for genuine N -partite entanglement. Interestingly, it can be shown that N -partite entanglement does not necessitate the presence of correlations between all parties. This is demonstrated on the example of photonic three- and five-qubit states.

In the second part of this work, permutationally invariant tomography is presented which is a novel partial tomography scheme for the efficient analysis of symmetric multiqubit states. In the experimental implementation of this scheme, the emphasis was to increase the count rates thus far that full tomography of six-photon states became possible. In experiments with four- and six-photon symmetric Dicke states, various tomography schemes, like full tomography and permutationally invariant tomography, were compared against each other and it

Abstract

was shown that the results of the respective schemes are compatible. In this context, it is of utmost importance that not only the measurement scheme itself is scalable but also subsequent data analysis. This can be achieved by a specially adopted algorithm which, on the one hand, utilizes the symmetry of the state to be analyzed, and, on the other hand, resorts to highly efficient methods from convex optimization. It was also observed that data analysis itself can cause systematic errors of the final result. As it turned out, for the measurement statistics that are typical of today's multiqubit experiments, the systematic errors are of the same magnitude as the statistical errors and can therefore not be ignored.

Zusammenfassung

Die Quantenmechanik zählt zweifelsohne zu einer der erfolgreichsten wissenschaftlichen Theorien, die aber aufgrund ihrer kontraintuitiven Vorhersagen viele Physiker nach wie vor herausfordert. In diesem Zusammenhang ist insbesondere das aus dem Superpositionsprinzips folgende Phänomen der Verschränkung zu nennen. Auch wenn sich Verschränkung klassisch nicht verstehen lässt, so bildet sie doch das Herzstück für neuartige Methoden der Datenverarbeitung wie Quantenkommunikation oder Quantenrechnen. In einer interdisziplinären Anstrengung von Mathematik, Informatik und Physik versuchen Forscher Vielteilchenverschränkung besser zu verstehen und zu charakterisieren, sowie für die praktische Anwendung nutzbar zu machen.

Die vorliegende Arbeit beschäftigt sich mit der Entwicklung von geeigneten Methoden zur Analyse und Charakterisierung von vielteilchenverschränkten Zuständen sowie deren experimentellen Realisierung anhand von verschiedenen polarisationskodierten Mehrphotonenzuständen. Hierbei werden, neben Bell-Zuständen, insbesondere symmetrische Dicke-Zustände mit bis zu sechs Photonen untersucht. Als Photonenquelle wird dabei der Prozess der spontanen parametrischen Fluoreszenz in verschiedenen Varianten verwendet.

Im ersten Teil dieser Arbeit werden zwei Methoden vorgestellt und experimentell implementiert, die es erlauben die Verschränkung von Zwei- und Mehrqubit-Zuständen mit möglichst wenigen Korrelationsmessungen nachzuweisen. Hierbei beruht die erste Methode auf einer Transformation des Zustandes in die Schmidt-basis, wodurch sich im Anschluss Verschränkung mit maximal drei Korrelationsmessungen nachweisen lässt. Die zweite Methode verwendet das Prinzip der Korrelationskomplementarität, um nur die Messungen durchzuführen, die am vorteilhaftesten sind um Verschränkung möglichst schnell nachzuweisen. Das Schema lässt sich kompakt in Form eines Entscheidungsbaumes ausdrücken, der die als nächstes durchzuführende Messung in Abhängigkeit des vorherigen Messergebnisses festlegt. In einem weiteren Experiment wird untersucht, welche Arten von Korrelationen für genuine N -Teilchenverschränkung notwendig sind. Dabei kann gezeigt werden, dass für N -Teilchenverschränkung interessanterweise keine N -Parteien-Korrelationen zwingend erforderlich sind. Dies wird exemplarisch für photonische Zustände mit drei und fünf Qubits demonstriert.

Im zweiten Teil dieser Arbeit wird mit permutationsinvarianter Tomographie eine neuartige partielle Tomographiemethode vorgestellt, die eine effiziente Untersuchung von symmetrischen Multiqubitzuständen erlaubt. In der experimentellen

Zusammenfassung

Umsetzung liegt der Schwerpunkt hierbei insbesondere darauf, durch viele Detailverbesserungen die Zählraten soweit zu erhöhen, dass eine vollständige Zustandsanalyse auch für sechs Photonen möglich ist. In Experimenten mit symmetrischen Dicke-Zuständen mit vier und sechs Photonen werden verschiedene tomographischen Methoden, darunter vollständige Tomographie und permutationsinvariante Tomographie, miteinander verglichen. Dabei konnte gezeigt werden, dass alle Methoden kompatible Resultate liefern. Hier ist es insbesondere wichtig, dass nicht nur das Messschema selbst skalierbar ist, sondern auch die anschließende Auswertung der Daten. Dies wird durch einen speziell angepassten Algorithmus erreicht, der einerseits die Symmetrie des zu analysierenden Zustandes ausnutzt und andererseits auf hocheffiziente Methoden der konvexen Optimierung zurückgreift. In diesem Zusammenhang wurde auch untersucht, in wie weit die Datenauswertung selbst zu systematischen Fehlern führen kann. Wie sich herausstellte, sind bei für Multiqubitexperimenten typischen Messstatistiken die systematischen Fehler von der gleichen Größenordnung wie die statistischen und können folglich nicht vernachlässigt werden.

1. Introduction

The nineteenth century witnessed an unprecedented deepening in the understanding of the laws of physics culminating in the theories of electrodynamics and thermodynamics. Together with the theory of classical mechanics, these theories form the core of what is today called *classical* physics. However, at the turn of the century, there were, due to refined experimental techniques, more and more new findings that could not be understood within the framework of classical physics. Especially, black body radiation, the stability of atoms and the spectra of atoms lacked a satisfactory explanation. In 1900, when Max Planck [1] derived the radiation law, the first of these puzzles could be resolved, thereby constituting the birth of quantum mechanics. In the first decades of the twentieth century, quantum mechanics developed further and comprehensive explanations for all of the above mentioned problems could be given. Despite its tremendous success, there was no consensus about the correct interpretation of the new theory, as shown by the historical debate between Nils Bohr and Albert Einstein [2, 3].

Especially, the probabilistic nature of quantum mechanics lead to a strong feeling of unease in Albert Einstein. In 1935, he made his concerns explicit by publishing a seminal paper [4] together with Boris Podolski and Nathan Rosen. In this paper, the authors consider a state that was later called “entangled” by Erwin Schrödinger [5]. Entangled states allow for situations where, although the system as a whole is perfectly known, the knowledge about its individual subsystems can be minimal. In such cases, the system can only be treated as a whole, even if its subsystems are arbitrarily far separated. This lead Einstein and his coworkers to the conclusion that quantum mechanics must be incomplete and they postulated the existence of additional, so-called hidden variables. For almost three decades, entanglement was mainly considered as a strange peculiarity of the theory which is, at best, of philosophical interest but has no practical relevance.

This situation changed completely, when John Bell reformulated the original argument of Einstein, Podolski and Rosen in a way that allowed to deduce experimentally testable predictions which differ from those of quantum mechanics [6]. Over the years, these tests, which are now known under the term “Bell tests”, were further refined, but already the first experimental tests made a clear statement in favor of quantum mechanics [7–9]. Nonetheless, hidden variable theories could not be completely eradicated since all Bell tests that were performed up to now have some “loopholes” which allow, at least in principle, for a classical interpretation. While certain loopholes have been closed separately [10, 11], at the

1. Introduction

time of writing, an experiment where all the loopholes are closed simultaneously is still missing. However, it is more than likely that such an experiment will be carried out in the near future [12–14].

Additionally to these fundamental aspects, it was realized that quantum mechanics also allows for novel approaches with respect to information processing [15, 16]. Examples for potential new applications with entanglement as a key resource comprise, to name just few, dense coding [17], quantum cryptography [18], quantum enhanced metrology [19] or quantum teleportation [20]. For some computational tasks, like, searching in huge databases [21] or factoring large numbers [22], it could be shown that they can, at least in principle, be performed exponentially faster than classically.

Central to quantum information is the quantum bit, or shortly qubit [23], which is the quantum mechanical analog of the classical bit. The main difference between qubits and their classical counterparts is that they cannot only be in the ground or the excited state of a two-level system but in any superposition of the two. Experimentally, qubits can be implemented with various physical systems, like, e.g., trapped atoms [24, 25], trapped ions [26], superconducting qubits [27–29], NV centers in diamond [30], and photons [31]. Each of the different approaches has its advantages but also its specific drawbacks. For quantum computational problems, localized, well controllable systems like atoms, ions or superconducting qubits seem to be preferable, whereas for quantum communication tasks where qubits have to be exchanged over large distances, photons are the system of choice. In the last years, due to refined experimental techniques, the number of entangled qubits was constantly increasing [32–35]. However, despite of this enormous progress, it remains an open question whether it will be possible to control hundreds of qubits one day and to really harness the power that lies within quantum computation. In this context, it is still an open issue how large a quantum system can grow before inherent decoherence, even arbitrarily small, induces a transition from quantum to classical [36, 37].

In this thesis, the polarization degree of freedom of photons was used for the experimental implementation of qubits and to investigate various quantum information tasks. Photons are an ideal test system for proof-of-principle realizations of many quantum information protocols [31] as they experience almost no decoherence. However, the major advantage of photons is at the same time their worst drawback, namely, that they practically do not interact with each other. Consequently, two-qubit gates as they are required for quantum computation cannot be realized in a simple manner. This restriction could be overcome in 2001, when Emanuel Knill, Raymond Laflamme, and Gerard Milburn presented a scheme for optical quantum computing that requires only single-photon sources, linear optical components and photon number resolving detectors [38]. The original scheme was later on further refined and could be reduced to the deterministic preparation of specific multipartite entangled quantum states together with suitable single-qubit

operations [39–43]. At the time of writing, the main drawback of optical quantum computation is still the lack of practical single-photon sources but the technical development proceeds fast [44–48]. Until such sources are available, one has to resort, as also in this work, to the process of spontaneous parametric down conversion [49, 50] together with linear optical networks in order to prepare photonic multiqubit states.

Although the number of qubits is still moderate in today’s multiqubit experiments, up to six in this work, full quantum state tomography, i.e., a complete characterization of the state, becomes already a great challenge. The limit where full tomography becomes infeasible due to experimental restrictions, like set-up stability, drifts, measurement time etc. will soon be reached. Thus, it is of utmost importance to find simple and experimentally friendly tools where relevant quantities about a quantum state, like, e.g., its overlap with respect to a target state, can be inferred from partial information. The research focus of this work is twofold. On the one hand, it is on the development and application of tools for entanglement detection from possibly few correlation measurements [51–53] (the corresponding publications are included in this work as P3.1 and P3.2). In this context, also the connection between genuine N -partite entanglement and the existence of N -partite correlations is investigated [54] (see publication P3.3). On the other hand, the focus is on the increase of rates such that full tomography of a six-qubit state becomes possible and can be compared against highly efficient tomography protocols which require only partial information [55–58] (see publications P4.1, P4.2 and P4.3). There, it was also observed that the standard analysis tools for tomography can lead to systematic errors [59] (see publication P5.1).

For many entanglement criteria, it is required that the state to be characterized is properly aligned with respect to the local measurement bases [60, 61], which does not pose a problem, when all observers share a common reference frame. If, however, a common reference frame cannot be guaranteed, a situation that may likely occur in quantum information protocols where the observers are far separated from each other, entanglement might therefore not be detected, although it is present. In order to overcome this problem, two different schemes are discussed (see P3.1 and P3.2). The first one, which is designed for pure two-qubit states, effectively performs a Schmidt decomposition [62] of the state. Then, after a redefinition of the local measurement bases, entanglement can easily be verified with a simple criterion [51] which requires at maximum three correlation measurements. The second scheme is more general and is based on the principle of correlation complementarity [63], where the results of the performed correlation measurements determine which is the next measurement to be carried out.

In order to detect entanglement between N parties, one normally resorts to entanglement criteria which depend on correlations between all parties [51, 60, 64–71]. In this context, correlations are defined by the standard correlation function as used in classical statistics and signal analysis [72, 73]. However, as first shown

1. Introduction

by Kaszlikowski et al. in [74], quantum mechanics allows for states that are genuinely N -partite entangled, but lack N -partite correlations. At no surprise, these findings triggered a vivid discussion about the distinction between classical and quantum correlations [75, 76]. In this work, the result by Kaszlikowski et al. is generalized and a constructive scheme is presented which allows to generate an infinite family of genuinely N -partite entangled states with vanishing N -partite correlations. In three- and five-qubit experiments, such peculiar quantum states were prepared for the first time [54] (see also P3.3).

As already mentioned above, full quantum state tomography is not feasible for larger systems. However, many prominent and highly relevant quantum states belong to subclasses of states which can be represented by few parameters. For such states, efficient tomography schemes exist, like, e.g., tomography of matrix product states [77], compressed sensing [55] or permutationally invariant tomography [56]. As these schemes are to be applied in situations where full tomography cannot be performed, it is absolutely necessary to test these schemes against full tomography for systems where this is still possible. In this work, full tomography of a symmetric six-photon Dicke state [78, 79] was performed and the result was compared with the results from compressed sensing, permutationally invariant tomography and a combination of the two [58] (see publications P4.1 and P4.2). For scalable tomography schemes, it is also crucial to have highly efficient algorithms for data analysis at hand [57] (see P4.3).

In every experimental situation one has to perform a thorough and careful analysis of all the errors that might influence the final result. Here, one has to distinguish between unavoidable statistical errors [80] and systematic errors, like, e.g., drifts of the experimental apparatus [81]. Interestingly, also data analysis itself can lead to systematic errors if additional constraints, like the physicality of an estimated quantum state, are applied. In this context, the standard procedures to analyze tomography data were investigated. There, it was observed that, for typical count statistics of today's multiqubit experiments, the deviations can be considerable [59] (see publication P5.1).

The thesis is structured as follows: In chapter 2, a short review of the foundations of quantum information is given and the notations used throughout this work are introduced. More precisely, the theoretical concepts of two- and multiqubit entanglement are explained together with the most relevant tools to detect and quantify entanglement. Additionally, the basic building blocks of the experiments that were performed in the course of this work are discussed. Then, in chapter 3, both theoretical and experimental results how entanglement can be detected in an experimentally efficient way are presented. In chapter 4, novel methods for efficient state tomography are presented and compared against the standard method. Finally, in chapter 5, the standard tools for data analysis are reviewed, especially with respect to their reliability. The respective chapters also contain the relevant research papers that were written in the course of this work.

2. Fundamentals of quantum information theory

In this chapter, the basic concepts and tools required to describe and analyze single and multiqubit systems are introduced. Section 2.1 starts with a formal definition of what a qubit is [23] and discusses both the similarities and differences between a classical and a quantum bit. Then, in section 2.2, systems consisting of more than a single qubit are considered. As a consequence of the superposition principle, such systems can be in a state without classical analog, a so-called entangled state. After a formal definition of entanglement [82], several possibilities how to classify entangled states are presented. More precisely, entangled states with more than two qubits feature a rich hierarchical structure, i.e., they can be entangled in more than just one way [83].

If entanglement is to be used as a resource for practical applications, it is vital to detect and quantify its amount in an experimentally prepared state. Therefore, in section 2.2.3, several tools that allow to detect entanglement [84, 85] of a state are discussed and in section 2.2.4, some well-known entanglement measures [86, 87] are listed.

In section 2.3, the last part of this chapter, the focus lies on the practical realization of multiphoton entangled states. It begins with a review of photon sources based on the process of spontaneous parametric down-conversion (SPDC) [49, 50] as such photon sources are still the workhorse of today's multiphoton experiments. Then, the basic building blocks to further process the generated photons and to prepare the desired state are discussed. In detail, these building blocks are beam splitters (BS), both polarizing and non-polarizing, phase shifters such as half (HWP) and quarter wave plates (QWP) and combinations thereof, e.g. for polarization analysis.

In summary, this introductory chapter can be considered as a short review of the most relevant concepts and tools required to carry out experiments with multiphoton entangled states. The discussion goes largely along the same line as in [88] and the introductory chapters of [89–91]. The chapter lays the necessary foundations to understand the experiments described in chapters 3 to 5.

2.1. Single qubits

2.1.1. From classical bits to quantum bits

In the following, the basic notation and concepts necessary to describe single two-level quantum systems are introduced. Such systems, called quantum bit or, for short, *qubit* [23], are the quantum mechanical analog of the classical bit. Contrary to a classical bit, which can either be in the ground state 0 or in the excited state 1, a quantum bit can be in any superposition of these two states. Formally, a pure single-qubit state is given by [62, 92–95]

$$|\psi\rangle = \frac{1}{\sqrt{|\alpha|^2 + |\beta|^2}}(\alpha|0\rangle + \beta|1\rangle) \quad (2.1)$$

where $|0\rangle$ and $|1\rangle$ designate two orthogonal basis states of the Hilbert space $\mathcal{H} = \mathbb{C}^{\otimes 2}$ and with complex numbers α and β . Here, and later on in this thesis, the *Dirac notation* $|\cdot\rangle$ will be used to discriminate quantum states from classical states. Since the state $|\psi\rangle$ has to be normalized, i.e. $|\langle\psi|\psi\rangle|^2 = 1$, it can be rewritten, up to a negligible global phase, as

$$|\psi(\theta, \phi)\rangle = \cos\frac{\theta}{2}|0\rangle + e^{i\phi}\sin\frac{\theta}{2}|1\rangle. \quad (2.2)$$

with $\theta \in [0, \pi]$ and $\phi \in [0, 2\pi[$. Written in the form of Eq. 2.2, it is obvious that $|\psi\rangle$ can be represented as a point on a three dimensional sphere called *Bloch sphere*, see Fig. 2.1.

Alternatively, a pure single-qubit state $|\psi(\theta, \phi)\rangle$ can also be interpreted as the eigenvector of an observable $\sigma(\theta, \phi)$ with eigenvalue $+1$. The operator $\sigma(\theta, \phi)$ can be expressed in terms of any operator basis, like, e.g., the frequently used Pauli basis

$$\sigma(\theta, \phi) = \sin\theta\cos\phi\sigma_x + \sin\theta\sin\phi\sigma_y + \cos\theta\sigma_z \quad (2.3)$$

where $\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, $\sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$, and $\sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ are the Pauli spin matrices. The second eigenvector of $\sigma(\theta, \phi)$ with eigenvalue -1 is labeled $|\psi(\theta, \phi)\rangle_{\perp}$ and is perpendicular to $|\psi(\theta, \phi)\rangle$, i.e., $\langle\psi(\theta, \phi)|\psi(\theta, \phi)\rangle_{\perp} = 0$. Each of the Pauli matrices has two eigenvectors, one with eigenvalue $+1$ and the other one with eigenvalue -1 , i.e.,

$$\sigma_x|x^{\pm}\rangle = \pm|x^{\pm}\rangle, \quad (2.4)$$

$$\sigma_y|y^{\pm}\rangle = \pm|y^{\pm}\rangle, \quad (2.5)$$

$$\sigma_z|z^{\pm}\rangle = \pm|z^{\pm}\rangle \quad (2.6)$$

with $|x^{\pm}\rangle = 1/\sqrt{2}(|0\rangle \pm |1\rangle)$, $|y^{\pm}\rangle = 1/\sqrt{2}(|0\rangle \pm i|1\rangle)$, $|z^+\rangle = |0\rangle$, and $|z^-\rangle = |1\rangle$. The eigenvectors define the x -, y - and z - axes of the Bloch sphere, see Fig. 2.1.

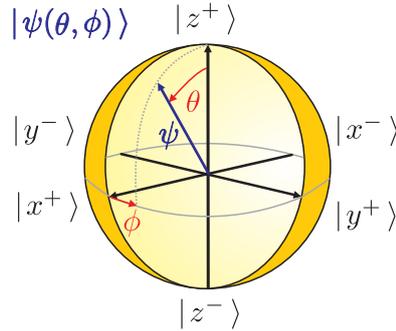


Figure 2.1.: A pure quantum state $|\psi(\theta, \phi)\rangle$ can be represented as a point on the surface of the Bloch sphere. The mixed states lie inside the sphere with the totally mixed state $\mathbb{1}/2$ at the origin of the sphere. The x -, y - and z -axes of the coordinate system are defined by the eigenvectors of the Pauli matrices σ_x , σ_y and σ_z .

Therefore, from now on, “measuring in the σ_x basis” will be used as a synonym for performing projection measurements on the pair of orthogonal vectors $|x^+\rangle$ and $|x^-\rangle$. Correspondingly, the expressions “measuring in the σ_y basis” and “measuring in the σ_z basis” will be used.

As already mentioned, the points on the surface of the Bloch sphere can be interpreted as pure quantum states. In addition, also the points lying inside the sphere can be associated with quantum states, the so-called *mixed* states, which are states with a classical statistical distribution over pure states. A mixed state cannot be described by a state vector $|\psi\rangle$ anymore but, instead, by a density matrix

$$\varrho = \sum_{i=0}^1 p_i |\psi(\theta_i, \phi_i)\rangle \langle \psi(\theta_i, \phi_i)| \quad (2.7)$$

with probabilities p_i which have to fulfill the constraints $p_i \geq 0$ and $\sum_i p_i = 1$. Of course, also every pure state $|\psi(\theta, \phi)\rangle$ can be expressed by means of a density matrix. Then, the sum Eq. 2.7 collapses since all p_i except one vanish, i.e.,

$$|\psi(\theta, \phi)\rangle \longrightarrow \varrho_{|\psi(\theta, \phi)\rangle} = |\psi(\theta, \phi)\rangle \langle \psi(\theta, \phi)|. \quad (2.8)$$

The complete information about a quantum state ϱ is contained in its density matrix. Hence, in order to uniquely specify a quantum state, it is sufficient to define all entries of the corresponding density matrix. However, especially for theoretical arguments, expressing ϱ by means of Pauli matrices turns out to be advantageous. Then, ϱ is given by

$$\varrho = \frac{1}{2}(T_0 \mathbb{1} + T_x \sigma_x + T_y \sigma_y + T_z \sigma_z) \quad (2.9)$$

2. Fundamentals of quantum information theory

with the coefficients called Bloch vector, $\vec{T} = (T_0, T_x, T_y, T_z)$. The elements of the Bloch vector are the expectation values of the Pauli matrices and are called correlations. They can be determined as

$$T_x = \text{Tr}[\rho\sigma_x], \quad (2.10)$$

$$T_y = \text{Tr}[\rho\sigma_y], \quad (2.11)$$

$$T_z = \text{Tr}[\rho\sigma_z] \quad (2.12)$$

which simplifies to

$$T_x = \langle \psi | \sigma_x | \psi \rangle, \quad (2.13)$$

$$T_y = \langle \psi | \sigma_y | \psi \rangle, \quad (2.14)$$

$$T_z = \langle \psi | \sigma_z | \psi \rangle \quad (2.15)$$

for pure states. The correlation T_0 is always 1, independent of the state, due to normalization. The actual physical meaning of the Bloch vector can best be understood when it is compared to the Stokes vector $\vec{S} = (S_0, S_1, S_2, S_3)$ which describes the polarization state of light. The components of \vec{S} are given by

$$S_0 = P_{|z^+\rangle} + P_{|z^-\rangle}, \quad (2.16)$$

$$S_1 = P_{|z^+\rangle} - P_{|z^-\rangle}, \quad (2.17)$$

$$S_2 = P_{|x^+\rangle} - P_{|x^-\rangle}, \quad (2.18)$$

$$S_3 = P_{|y^+\rangle} - P_{|y^-\rangle} \quad (2.19)$$

where $P_{|z^+\rangle}$ is the power observed behind a polarizer oriented along $|z^+\rangle$ and $P_{|z^-\rangle}$ etc. defined correspondingly. Usually, the Stokes vector is normalized to the total intensity, i.e., $\vec{S}_N = \frac{1}{S_0}\vec{S}$. By expressing the Pauli matrices in Eq. 2.10-2.12 in terms of their corresponding eigenvectors, i.e., $\sigma_x = |x^+\rangle\langle x^+| - |x^-\rangle\langle x^-|$ etc., one obtains

$$T_x = \text{Tr}[\rho |x^+\rangle\langle x^+|] - \text{Tr}[\rho |x^-\rangle\langle x^-|], \quad (2.20)$$

$$T_y = \text{Tr}[\rho |y^+\rangle\langle y^+|] - \text{Tr}[\rho |y^-\rangle\langle y^-|], \quad (2.21)$$

$$T_z = \text{Tr}[\rho |z^+\rangle\langle z^+|] - \text{Tr}[\rho |z^-\rangle\langle z^-|]. \quad (2.22)$$

Since $\text{Tr}[\rho |x^+\rangle\langle x^+|]$ is the probability to successfully project ρ on $|x^+\rangle$ and thus proportional to the intensity of a classical light field, the Bloch vector can be seen as the quantum mechanical analog of the Stokes vector. However, the ordering of the entries is swapped, S_1 corresponds to T_z , S_2 corresponds to T_x , and S_3 corresponds to T_y .

2.1.2. Single-qubit manipulation and discrimination

Unitary transformations

Any two points on the surface of the Bloch sphere, i.e., any two pure single-qubit states, can be mapped onto each other by an appropriate reversible transformation U . In order to be reversible, such transformations, called *unitary*, have to fulfill the constraints $U^\dagger U = U U^\dagger = \mathbb{1}$ and $\det(U) = 1$. For many quantum information protocols unitary single-qubit manipulations are required. In setups using polarized photons, these transformations are implemented with $\lambda/4$ wave plates (QWP), $\lambda/2$ wave plates (HWP) or, for a pure phase shift, by tilting a birefringent crystal around its optical axis which is aligned parallel to vertical polarization. These components will be discussed in detail in section 2.3.2.

Single-qubit characterization

So far, the basic concepts to describe pure and mixed single-qubit states as well as possible manipulations have been discussed. However, what is still missing, are appropriate tools to distinguish and quantify different states. Therefore, in the following, the most important quantities for the characterization and discrimination of single-qubit states will be introduced. Please note that these quantities will also be used later on in order to characterize multiqubit states. Then the summation indices have to be adopted correspondingly to reflect the larger system size.

Fidelity Probably the most popular measure to quantify how close two quantum states ϱ and σ are is the fidelity. Formally, the fidelity \mathcal{F} is given by [96, 97]

$$\begin{aligned} \mathcal{F}(\varrho, \sigma) &= \text{Tr} \left[\sqrt{\sqrt{\sigma} \varrho \sqrt{\sigma}} \right]^2 \\ &= \text{Tr} \left[\sqrt{\sqrt{\varrho} \sigma \sqrt{\varrho}} \right]^2 = \mathcal{F}(\sigma, \varrho). \end{aligned} \quad (2.23)$$

If, one of the states is pure, the expression for the fidelity simplifies to

$$\mathcal{F}(\varrho, \sigma) = \text{Tr}[\varrho \sigma] \quad (2.24)$$

as can easily be shown by direct calculation.

Hilbert-Schmidt norm Another widely spread measure to quantify the closeness of two quantum states is the Hilbert-Schmidt norm. It is defined as

$$\|\varrho\|_2 = \sqrt{\sum_{i=0}^1 \sum_{j=0}^1 |\varrho_{i,j}|^2} \quad (2.25)$$

2. Fundamentals of quantum information theory

for a single quantum state ϱ . Based on this definition, the difference between ϱ and σ can then be quantified by

$$\|\varrho - \sigma\|_2 = \sqrt{\sum_{i=0}^1 \sum_{j=0}^1 |\varrho_{i,j} - \sigma_{i,j}|^2}. \quad (2.26)$$

Trace distance The trace distance between two states ϱ and σ is given by the sum of the absolute value of the (real) eigenvalues of their difference, i.e.,

$$\|\varrho - \sigma\|_1 = \sum_{i=1}^4 |\lambda_i| \quad (2.27)$$

with eigenvalues λ_i .

Purity In order to quantify how pure a state ϱ is, i.e., how close it lies to the surface of the Bloch sphere, the purity $\mathcal{P}(\varrho)$ is used. It is given by

$$\mathcal{P}(\varrho) = \text{Tr}[\varrho^2] \quad (2.28)$$

which can, in terms of the Bloch vector elements, be rewritten as

$$\begin{aligned} \mathcal{P}(\varrho) &= \frac{1}{2}(T_0^2 + T_x^2 + T_y^2 + T_z^2) \\ &= \frac{1}{2}(1 + T_x^2 + T_y^2 + T_z^2). \end{aligned} \quad (2.29)$$

From Eq. 2.29, it can be directly seen that the purity is proportional to the square of the Bloch vector's length. For a pure state, the purity is 1 whereas for the totally mixed state $\mathbb{1}/2$ it is $1/2$.

2.2. Multiqubit states and entanglement

2.2.1. Multiple qubits

In the previous section, the conceptual basics necessary to describe single-qubit states were discussed. In the following, these concepts will be generalized to systems consisting of multiple qubits. One possible choice of basis states of a two-qubit Hilbert space $\mathcal{H}_2 = \mathbb{C}^4$ is given by the tensor products of the basis states of the single-qubit Hilbert space $\mathcal{H}_1 = \mathbb{C}^2$ as introduced in section 2.1,

$$|00\rangle = |0\rangle \otimes |0\rangle \quad (2.30)$$

$$|01\rangle = |0\rangle \otimes |1\rangle \quad (2.31)$$

$$|10\rangle = |1\rangle \otimes |0\rangle \quad (2.32)$$

$$|11\rangle = |1\rangle \otimes |1\rangle. \quad (2.33)$$

2.2. Multiqubit states and entanglement

A general pure two-qubit state is then given by an arbitrary *superposition* of the basis states

$$|\psi\rangle_{1,2} = c_{0,0}|00\rangle + c_{0,1}|01\rangle + c_{1,0}|10\rangle + c_{1,1}|11\rangle \quad (2.34)$$

with complex coefficients $c_{i,j}$ which have to fulfill, due to normalization of $|\psi\rangle_{1,2}$, the constraint $|c_{0,0}|^2 + |c_{0,1}|^2 + |c_{1,0}|^2 + |c_{1,1}|^2 = 1$. Analogously, for N qubits, the basis states are

$$|00\dots 0\rangle = |0\rangle^{\otimes N} = |0\rangle_1 \otimes |0\rangle_2 \otimes \dots \otimes |0\rangle_N, \quad (2.35)$$

$$|00\dots 1\rangle = |0\rangle_1 \otimes |0\rangle_2 \otimes \dots \otimes |1\rangle_N, \quad (2.36)$$

⋮

$$|11\dots 1\rangle = |1\rangle^{\otimes N} = |1\rangle_1 \otimes |1\rangle_2 \otimes \dots \otimes |1\rangle_N, \quad (2.37)$$

and correspondingly a general N -qubit pure state is given by

$$|\psi\rangle_{1,2,\dots,N} = \sum_{\substack{i_1, i_2, \dots, i_N \\ \in \{0,1\}}} c_{i_1, i_2, \dots, i_N} |i_1 i_2 \dots i_N\rangle \quad (2.38)$$

with $c_{i_1, i_2, \dots, i_N} \in \mathbb{C}$ and $\sum_{i_1, i_2, \dots, i_N} |c_{i_1, i_2, \dots, i_N}|^2 = 1$. A mixed N -qubit state ϱ is, analogously to the single-qubit case, described by an incoherent sum of pairwise orthogonal pure states $|\psi_1\rangle, \dots, |\psi_{2^N}\rangle$,

$$\varrho = \sum_{i=1}^{2^N} p_i |\psi_i\rangle \langle \psi_i| \quad (2.39)$$

with $p_i \geq 0$ and $\sum_{i=1}^{2^N} p_i = 1$. Please note, that ϱ can also be expressed in terms of another set of pure states forming a basis. As in the single-qubit case, it is often favorable to express a mixed multiqubit state in the Pauli basis

$$\varrho = \frac{1}{2^N} \sum_{\substack{j_1, j_2, \dots, j_N \\ \in \{0, x, y, z\}}} T_{j_1, j_2, \dots, j_N} \sigma_{j_1} \otimes \sigma_{j_2} \otimes \dots \otimes \sigma_{j_N} \quad (2.40)$$

where T_{j_1, j_2, \dots, j_N} is called correlation tensor and is a generalization of the Bloch vector from Eq. 2.9. The entries of T_{j_1, j_2, \dots, j_N} are called correlations again and can be inferred from ϱ as

$$T_{j_1, j_2, \dots, j_N} = \text{Tr}[\varrho \sigma_{j_1} \otimes \sigma_{j_2} \otimes \dots \otimes \sigma_{j_N}]. \quad (2.41)$$

2. Fundamentals of quantum information theory

Entanglement

For two and more qubits, as a consequence of the superposition principle, there exist states which cannot be expressed as a simple tensor product of individual subsystems [82]

$$|\psi\rangle_{1,2} \neq |\phi\rangle_1 \otimes |\zeta\rangle_2 \quad (2.42)$$

and, correspondingly, for mixed states

$$\varrho \neq \sum_i p_i (\varrho_1)_i \otimes (\varrho_2)_i. \quad (2.43)$$

Such states are called nonseparable or entangled, an expression which was coined by Schrödinger [5]. There are four two-qubit states which are specific with respect to several properties. These states, called Bell states, are defined as

$$|\phi^+\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle), \quad (2.44)$$

$$|\phi^-\rangle = \frac{1}{\sqrt{2}}(|00\rangle - |11\rangle), \quad (2.45)$$

$$|\psi^+\rangle = \frac{1}{\sqrt{2}}(|01\rangle + |10\rangle), \quad (2.46)$$

$$|\psi^-\rangle = \frac{1}{\sqrt{2}}(|01\rangle - |10\rangle), \quad (2.47)$$

and play an important role in quantum information theory [17, 18, 20]. They are, for example, maximally entangled with respect to several entanglement measures such as the negativity or the concurrence. They also maximally violate the CHSH-inequality (for details see section 2.2.3). In the course of this work, the Bell states Eq. 2.44-2.47 are of high importance as they are in the focus of publications P3.1 and P3.2 where a scheme to prove entanglement from possibly few measurement results is developed.

2.2.2. Classification of multiqubit states

Separability

For more than two qubits, there are various different levels of separability. The state may be fully separable or, e.g., only separable with respect to certain subgroups of particles or not separable at all. According to [85, 98], an N -qubit state ϱ is called k -separable if a decomposition of the form

$$\varrho = \sum_i p_i \otimes_{n=1}^k (\varrho_{S_n})_i \quad (2.48)$$

exists where $\otimes_{n=1}^k (\varrho_{S_n})_i$ denotes the tensor product of k density matrices for a chosen partition $\{S_1, \dots, S_k\}$ into k disjoint non-empty subsets with $k \leq N$. The

2.2. Multiqubit states and entanglement

various different levels of separability define an hierarchical ordering of quantum states. For $k = 1$, the states are called genuinely N -partite entangled where $\varrho \neq \sum_i p_i \otimes_{n=1}^k (\varrho_{S_n})_i$ for any $k > 1$. Next, for $k = 2$, come the biseparable states which can be factored in two subsystems. Finally, at the lowest level of the hierarchy, there are the fully separable or classical states. Please note that, based on this ordering of quantum states, one cannot make a statement about which of two states with the same separability is more entangled. In order to quantify entanglement, entanglement measures have been introduced (see section 2.2.4).

Stochastic local operation and classical communication

A sole distinction between entangled and non-entangled states turned out to be not sufficient to properly handle entanglement. Therefore, a finer classification of entanglement was required. For this purpose, the following approach turned out to be most fruitful. Consider a multiqubit state $|\psi\rangle$ of which each particle is distributed to one party. Now, the task for all the recipients is to transform the state $|\psi\rangle$ into a target state $|\phi\rangle$ and back again. For both transformations the recipients have to use the same set of operations. If this task can be achieved, the states $|\psi\rangle$ and $|\phi\rangle$ belong to the same equivalence class with respect to the chosen set of operations.

Obviously, the classification depends on the set of operations that are allowed. As entanglement cannot be created by local manipulation, one usually restricts the set of operations to local ones only. Depending on the different kinds of local transformations one can distinguish between several different cases.

- Local unitary (LU) operations, i.e., operations that can be expressed as $(U_1 \otimes \dots \otimes U_N |\psi\rangle) \longrightarrow |\tilde{\psi}\rangle$, with $U_i^\dagger U_i = U_i U_i^\dagger = \mathbb{1}$ for $i = 1, \dots, N$ are the most simple transformations that can be conceived. However, as they only correspond to rotations of the local coordinate systems, they are rather unsuitable to classify quantum states.
- Another, less restricted set of operations, is defined by the so-called local operations and classical communication (LOCC) [83, 99]. There, local operations, like measurements are allowed and, additionally, a classical channel between the different parties is established by which they can exchange information about, e.g., the result of a local measurement *before* another local measurement is performed. LOCC was conceived in the context of entanglement measures (section 2.2.4) with the aim to identify states with the same amount of entanglement [99]. However, also LOCC turned out to be too restrictive and to result in a classification of states that is still too fine.
- An even broader class of operations, stochastic LOCC (SLOCC), turned out to deliver the most appropriate classification of quantum states. Two states

2. Fundamentals of quantum information theory

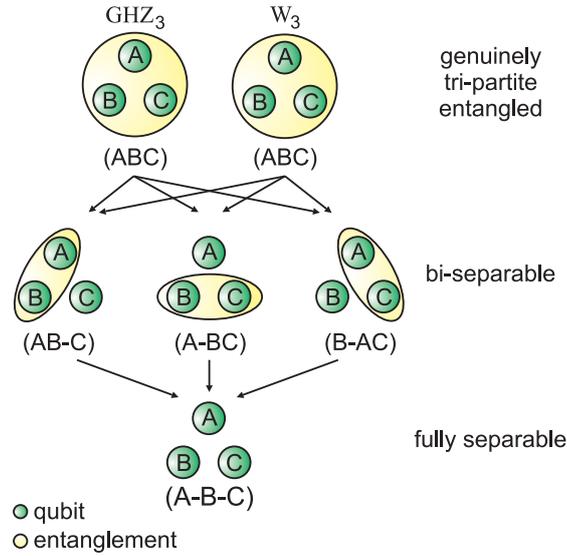


Figure 2.2.: The classification according to SLOCC operations of three qubits yields the fully-entangled GHZ and W classes, the biseparable states and the fully separable states [83].

are defined as equivalent with respect to SLOCC if the *probability* to convert one state into another and back is non-zero.

The application of SLOCC to the classification of pure three-qubit states was first treated by Dür, Vidal and Cirac [83]. They discovered six different SLOCC classes, the fully separable states (denoted by A-B-C in Fig. 2.2), three different types of bipartite states (AB-C, A-BC and B-AC) and genuinely tripartite entangled states (ABC). Astonishingly, they found that the state space separates into two inequivalent classes of genuinely tripartite entangled states, the GHZ-class and the W-class. They are both named after their most well-known representatives, namely, the Greenberger-Horne-Zeilinger (GHZ) state and the W state which are defined as

$$|\text{GHZ}_3\rangle = \frac{1}{\sqrt{2}}(|000\rangle + |111\rangle) \quad (2.49)$$

$$|\text{W}_3\rangle = \frac{1}{\sqrt{3}}(|001\rangle + |010\rangle + |100\rangle). \quad (2.50)$$

In detail, this means that no member of the GHZ-class can be converted by means of SLOCC into a state belonging to the W-class and vice versa. Even if a shared quantum resource of bi- or tripartite entangled states is at disposal, a local transformation between the two classes is not possible [100]. The classification via SLOCC for three-qubit states was extended to mixed states by Acín et al. in [101].

2.2. Multiqubit states and entanglement

For multiple qubits, a classification of entanglement by means of SLOCC turned out to be far more complex. In [102] for example, Verstraete et al., describe nine representative groups of four-qubit pure states that are equivalent under SLOCC with each group parametrized by up to four independent complex parameters. An alternative ansatz to classify four-qubit quantum states was proposed by Lamata and coworkers [103] which allows to inductively construct SLOCC classes for any number of qubits. With this ansatz, the six SLOCC classes for three qubits are recovered [104], but, for four qubits one yields, in contrast to the scheme introduced by Verstraete et al., only eight SLOCC classes. For further schemes how to classify multiqubit states, the interested reader may also want to consider [105] or [106].

Operator eigenstates

Apart from the classification schemes presented so far, another way to structure entangled states is according to characteristic properties. In the following, two groups of states shall be discussed that are eigenstates of certain operators and, at the same time, are highly relevant in quantum information and for this thesis. First, graph states [107, 108] will be considered, including GHZ and cluster states which are the most prominent examples for this group of states. Then, Dicke states [78, 79] will be introduced as they play a key role in the experiments described in this work. Of course, there are far more groups of states than can be discussed within the scope of this thesis and the interested reader is therefore referred to the respective literature, like e.g., [109–112] for multiqubit/qudit singlet states, [113–115] for bound entangled states or [116–118] for matrix product states.

A graph state $|g\rangle$, as the name already suggests, is defined by a mathematical graph, where each vertex V of the graph $g(V, E)$ represents a qubit and the edges E represent Ising-type nearest neighbor interactions [107, 108, 119]. Equivalently, a graph state can also be specified by a set of operators $\{\mathcal{S}_i\}$ with common eigenvector $|g\rangle$ and corresponding eigenvalue 1, i.e.,

$$\mathcal{S}_i |g\rangle = |g\rangle \quad \forall i. \quad (2.51)$$

The set of operators $\{\mathcal{S}_i\}$ is called stabilizer and its elements \mathcal{S}_i are named stabilizing operators [107, 119]. Please note that different graph states can be equivalent with respect to the aforementioned LU transformations or graph transformation rules as described in [107]. Both for the two- and three-qubit case, this ansatz delivers a single graph state class, with $|\phi^+\rangle$ and $|\text{GHZ}_3\rangle$ as representative states. Accordingly, the N -partite GHZ states are given by

$$|\text{GHZ}_N\rangle = \frac{1}{\sqrt{2}}(|0\rangle^{\otimes N} + |1\rangle^{\otimes N}) \quad (2.52)$$

with starlike graphs as shown in Fig. 2.3. Graph states have been extensively investigated as they allow for a wide range of applications in quantum informa-

2. Fundamentals of quantum information theory

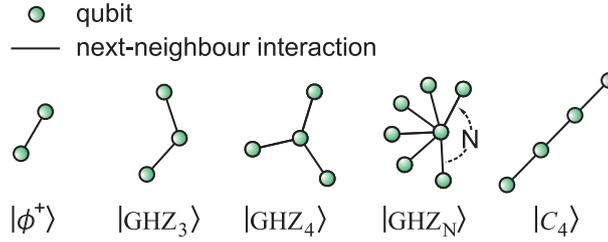


Figure 2.3.: Several prominent graph states together with the corresponding graphs.

tion science, like quantum enhanced metrology [19], dense-coding [120], secret-sharing [121] or one way quantum computation [39, 40, 122].

The second group of operator eigenstates discussed in this thesis are Dicke states which have first been studied by R. H. Dicke in the context of light emission from clouds of atoms [78]. There, it has been observed that atoms in a highly correlated state, a so-called Dicke state, emit light more strongly in comparison to atoms which are independent with respect to each other. Formally, Dicke states are eigenstates of the square of the total spin operator J_N^2 and, at the same time, of the spin operator component in the z -direction $J_{N,z}$

$$J_N^2 |D_N^{(k)}\rangle = j(j+1) |D_N^{(k)}\rangle \quad (2.53)$$

$$J_{N,z} |D_N^{(k)}\rangle = m |D_N^{(k)}\rangle \quad (2.54)$$

with $j \in \{0, 1, \dots, N/2\}$, $m \in \{-j, -j+1, \dots, j-1, j\}$ and $k = m + j$ for spin $1/2$ particles such as qubits. The operators J_N^2 and $J_{N,i}$ are defined as

$$J_N^2 = J_{N,x}^2 + J_{N,y}^2 + J_{N,z}^2 \quad \text{and} \quad (2.55)$$

$$J_{N,i} = \frac{1}{2} \sum \sigma_i^{(k)} \quad (2.56)$$

with $\sigma_i^{(3)} = \mathbb{1} \otimes \mathbb{1} \otimes \sigma_i \otimes \dots \otimes \mathbb{1}$ and $i \in \{x, y, z\}$. Most commonly, Dicke states with maximal J_N^2 are considered, see Fig. 2.4. These states are invariant under permutation of particles and can be used as a basis of the symmetric subspace (for more details see section 4). The most relevant examples of symmetric Dicke states are the W-states $|W_N\rangle$ ($j = N/2, m = -N/2 + 1$) and the states $|D_N^{(N/2)}\rangle$ ($j = N/2, m = 0$). For more details about Dicke states, the reader is referred to the special literature, like, e.g., [79, 123–127] for a theoretical treatment of Dicke states and [32, 58, 128–131] for their experimental realization.

2.2. Multiqubit states and entanglement

Dicke states $ D_N^{(k)}\rangle$ with $N=4$	
 $k=0$	$ D_4^{(0)}\rangle = 0000\rangle$
 $k=1$	$ D_4^{(1)}\rangle = W_4\rangle = (0001\rangle + 0010\rangle + 0100\rangle + 1000\rangle)/2$
 $k=2$	$ D_4^{(2)}\rangle = (0011\rangle + 0101\rangle + 0110\rangle + 1001\rangle + 1010\rangle + 1100\rangle)/\sqrt{6}$
 $k=3$	$ D_4^{(3)}\rangle = \bar{W}_4\rangle = (1110\rangle + 1101\rangle + 1011\rangle + 0111\rangle)/2$
 $k=4$	$ D_4^{(4)}\rangle = 1111\rangle$
 qubit in the ground state symmetrization qubit in the excited state	

Figure 2.4.: The four-qubit symmetric Dicke states $|D_4^{(k)}\rangle$.

2.2.3. Entanglement detection

Although the formal definition of entanglement looks simple [82], testing whether a given quantum state is entangled can be fairly complicated. However, for performing experiments, it is of utmost importance to have simple tools at hand that allow to decide whether a prepared state is entangled or not. In the following, a short overview of some of the most relevant entanglement criteria will be given. For a more detailed discussion of entanglement criteria, the reader is referred to the excellent review articles [84, 85, 132–134].

Bell inequalities

In the development of quantum information theory, Bell inequalities were the first criterion to detect entanglement [6], although the main motivation to formulate a Bell inequality was not to develop an efficient tool for entanglement detection. The high relevance of Bell inequalities comes mostly from the possibility to experimentally test the EPR paradox [4]. More precisely, a Bell inequality allows one to decide whether a local hidden variable (LHV) theory is consistent with quantum mechanics or whether such a theory can be refuted. In this context, LHV theories represent an entire class of models which rely on two assumptions. The first one, locality, phrases that observations on two space-like separated systems are independent from each other. The second assumption, realism, expresses that if measurement outcomes can be predicted with certainty, then there must be an underlying set of variables which determine the outcomes. From all these (hypothetical) models, predictions for the statistics of measurement outcomes can be deduced which are bounded by inequalities, the aforementioned Bell inequalities. These inequalities can then be tested in experiments. They are violated by

2. Fundamentals of quantum information theory

many entangled states whereas any separable state does not lead to a violation. Hence, Bell inequalities are a unique tool to investigate the discrepant predictions of quantum mechanics and LHV theories. Please note that not all entangled states necessarily lead to the violation of a certain Bell inequality [82, 135]. Nevertheless, Bell inequalities are still widely used as an entanglement criterion, although they are not optimal with respect to entanglement detection [136]. The most well-know Bell inequality in this context is probably the Clauser-Horne-Shimony-Holt inequality [68]. For an extensive review of Bell inequalities see [137–141].

Positive partial transpose

A popular and widely used criterion to detect entanglement between two qubits is based on the partial transpose of the density matrix [142]. In order to determine the partial transpose of a two-qubit density matrix ρ , parameterizing the entries of ρ with binary numbers is most favorable

$$\rho = \sum_{i,j \in \{0,1\}} \sum_{k,l \in \{0,1\}} \rho_{ij,kl} |i\rangle\langle j| \otimes |k\rangle\langle l|. \quad (2.57)$$

The partial transpose of ρ with respect to the first qubit, ρ^{T_1} , is then defined as

$$\rho^{T_1} = \sum_{i,j \in \{0,1\}} \sum_{k,l \in \{0,1\}} \rho_{ji,kl} |i\rangle\langle j| \otimes |k\rangle\langle l| \quad (2.58)$$

with a corresponding definition for ρ^{T_2} . A state is denoted to have a positive partial transpose (PPT), or just to be PPT, if it does not possess any negative eigenvalues after partial transposition. If however, $\rho^{T_1} \not\geq 0$ or $\rho^{T_2} \not\geq 0$, i.e., partial transposition leads to negative eigenvalues, then the state is entangled. In [143] it was shown that the partial transposition is a positive but not a completely positive map and that this property is crucial to use the PPT criterion for entanglement detection. Please note that the PPT criterion is sufficient and necessary only for $2 \otimes 2$ and $2 \otimes 3$ systems. The main drawback of the PPT criterion is probably that it requires complete information about a quantum state which has to be obtained via quantum state tomography. As will be discussed in detail in chapter 4, the experimental effort of quantum state tomography scales exponentially with the number of qubits and is thus unfeasible for larger numbers of qubits.

Correlation criteria

As explained in subsection 2.2.1, there is a one-to-one mapping between a quantum state and its correlation tensor, i.e., a quantum state is uniquely defined by its correlations. As the correlations are directly accessible by measurements, it is desirable to find entanglement criteria which are based, at best, on a subset of

2.2. Multiqubit states and entanglement

correlations. Using a special scalar product defined for correlation tensors T^{ϱ_1} and T^{ϱ_2} of two quantum states ϱ_1 and ϱ_2

$$(T^{\varrho_1}, T^{\varrho_2}) = \sum_{i_1, \dots, i_N \in \{x, y, z\}} T_{i_1, \dots, i_N}^{\varrho_1} T_{i_1, \dots, i_N}^{\varrho_2}, \quad (2.59)$$

Badziąg et al. [51] deduced a simple inequality which can be used for entanglement detection. The inequality reads

$$T^{\max} := \max_{T^{\text{prod}}} (T^{\text{prod}}, T^e) < (T^e, T^e) \quad (2.60)$$

where T^{prod} represents the correlation tensor of any product state. Eq. 2.60 can only be fulfilled if ϱ is entangled. It was further shown that $T^{\max} \leq 1$ [51] which is a very positive result for experiments. More precisely, this means that if the task is to prove that an experimentally prepared state is entangled, one keeps on measuring different correlations and stops as soon as the threshold is surpassed. The criterion can be used to formulate an efficient scheme to detect entanglement called *decision tree*, for details see chapter 3.1 and [52, 53]. The criterion can be refined to detect genuine N -partite entanglement as discussed in [65, 67].

Entanglement witnesses

Another widely used procedure to detect entanglement is the application of entanglement witnesses [143, 144]. They can be used to exclude separability and to even prove genuine multipartite entanglement. Linear witness are an experimentally friendly and extremely powerful tool which does not, in contrast to the PPT criterion, require state tomography. For certain classes of states, like, e.g., for Dicke states [61] or for certain graph states [60], there exist entanglement witnesses which require only two or three measurement settings, independent of the number of qubits. The main idea behind entanglement witnesses is to utilize the fact, that the separable states form a convex subset within the state space, as shown in Fig. 2.5. A witness is an operator \mathcal{W} which defines a hyperplane dividing the state space into two half-spaces, one containing all separable states and the other containing only genuinely multipartite entangled states, see Fig. 2.5. Hence, in order to prove that a state is genuinely multipartite entangled, it is sufficient to show that it lies in the corresponding half-space. Witness operators are constructed such that $\text{Tr}[\mathcal{W}\varrho_{\text{sep}}] > 0$ for all separable states ϱ_{sep} [143]. If, however, $\text{Tr}[\mathcal{W}\varrho] < 0$, then ϱ is proven to be entangled. A simple and widely used type of entanglement witnesses are the projector based witnesses,

$$\begin{aligned} \mathcal{W} &= \alpha \mathbb{1}^{\otimes N} - |\chi\rangle\langle\chi| \\ &\text{with } \alpha = \max_{|\psi_{\text{sep}}\rangle} |\langle\psi_{\text{sep}} | \chi\rangle|^2 \end{aligned} \quad (2.61)$$

2. Fundamentals of quantum information theory

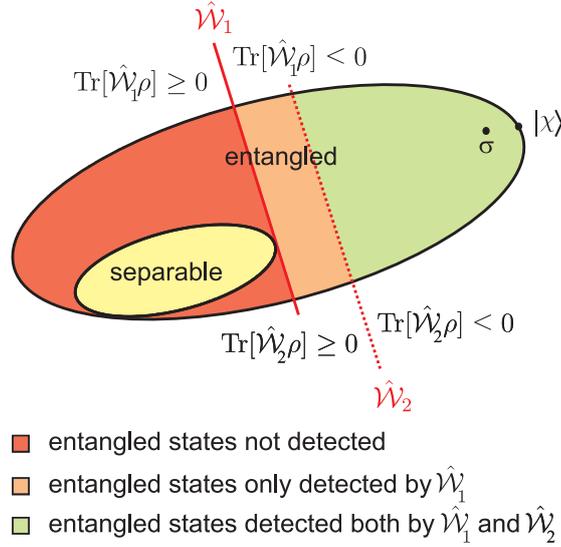


Figure 2.5.: Entanglement witnesses are hyperplanes in the state space. The witnesses \mathcal{W}_1 and \mathcal{W}_2 have been constructed to detect the entanglement of the pure state $|\chi\rangle$. Both will detect the entanglement of the mixed state σ . The witness \mathcal{W}_1 is finer than \mathcal{W}_2 as it detects all states that are detected by \mathcal{W}_2 and, additionally, states that are not detected by \mathcal{W}_2 .

which can be constructed for any given entangled state $|\chi\rangle$. However, it has to be noted that from $\text{Tr}[\mathcal{W}\rho] > 0$ one cannot conclude that ρ is separable. There might be a more sensitive witnesses \mathcal{W}' which can detect entanglement of ρ , i.e., $\text{Tr}[\mathcal{W}'\rho] < 0$. Witness \mathcal{W}' is said to be finer than witness \mathcal{W} if every state that is detected to be entangled by \mathcal{W} is also detected by \mathcal{W}' but not vice versa. This also allows to define optimality of a linear witness. A witness is optimal if no finer witness exists.

Other criteria

There are far more approaches to detect entanglement than can be discussed within the scope of this work. Therefore, a few more criteria, together with the respective references, shall at least be named. The quantum Fisher information is not only a measure to quantify the suitability of a certain quantum state for metrological applications [19, 145, 146] (for an overview of the current status of the field of quantum metrology see [147–152]). Furthermore, it can also be used to derive bounds for entanglement detection [153–155]. Spin-squeezing is another metrological concept [156, 157] that allows to deduce entanglement criteria, so called spin squeezing inequalities, from collective angular momentum operators of a spin system [61, 126, 158, 159]. The density matrix element criterion delivers inequalities which depend on diagonal and off-diagonal elements of the density

matrix. A violation of one of the various inequalities signals entanglement [160]. For further criteria, like, e.g., the range criterion, the matrix realignment criterion, the reduction criterion, or the majorization criterion, please see the excellent review articles [84, 85, 132–134].

2.2.4. Entanglement measures

Apart from a sole detection of entanglement, as discussed in the previous subsection, in many situations it is also of interest to quantify entanglement. Therefore, in the following, a short summary of some important entanglement measures shall be given, starting with a formal definition of the term *entanglement measure*. For a more complete overview of entanglement measures, refer to the review articles [86, 87].

A quantity can be called an entanglement measure $E(\varrho)$ only if it fulfills certain requirements as listed, e.g., in the overview article by Gühne and Tóth [84],

1. $E(\varrho_{\text{sep}}) = 0$ for all separable states ϱ_{sep} ,
2. $E(\varrho)$ cannot be increased by LOCC operations Λ^{LOCC} , i.e., $E(\Lambda^{\text{LOCC}}(\varrho)) \leq E(\varrho)$ and,
3. $E(\varrho)$ does not change under local unitary operations (LU).

The first property does not require further explanation, as any value $E(\varrho_{\text{sep}}) > 0$ indicates entanglement. Also the second requirement is natural as it results from the fact that entanglement cannot be created by the application of LOCC operations. The third requirement is also obvious, as LU operations only correspond to a redefinition of the local measurement bases. Sometimes, further properties are required which are not necessary but often advantageous, like,

4. convexity: $E(\sum_k p_k \varrho_k) \leq \sum_k p_k E(\varrho_k)$ and
5. additivity: $E(\varrho^{\otimes n}) = nE(\varrho)$ where $\varrho^{\otimes n}$ denotes n copies of ϱ .

Please note that for mixed states, the quantification of entanglement can pose a serious problem. In principle, every entanglement measure $E(\phi)$ defined for pure states ϕ can be generalized for mixed states via the convex-roof extension [161]

$$E(\varrho) = \inf_{p_k, |\phi_k\rangle} \sum_k p_k E(|\phi_k\rangle) \quad (2.62)$$

with $\varrho = \sum_k p_k |\phi_k\rangle\langle\phi_k|$. However, an analytical solution to the optimization over all possible decompositions of ϱ in Eq. 2.62, can only be given for a few measures. A numerical search over all decompositions will generally lead to an upper bound, i.e., an overestimation of entanglement. Recently, a technique based on semidefinite programming was proposed that provides a lower bound [162].

2. Fundamentals of quantum information theory

In the following, the formal definition and the most important properties of several widely used entanglement measures will be given. The discussion goes mainly along the same line as in [88, 91].

Entanglement of formation

For pure states $\varrho = |\psi\rangle\langle\psi|$, the entanglement of formation [163] is defined as the von Neumann entropy of the reduced state ϱ_1

$$E_F = S(\varrho_1) = -\text{Tr}[\varrho_1 \log_2 \varrho_1] = -\sum_i \lambda_i^{\varrho_1} \log_2 \lambda_i^{\varrho_1} \quad (2.63)$$

where

$$\begin{aligned} \varrho_1 &= \text{Tr}_2[\varrho] \\ &= \sum_{m \in \{0,1\}} \sum_{i,j \in \{0,1\}} \sum_{k,l \in \{0,1\}} \varrho_{ij,kl} |i\rangle\langle j| \otimes |m\rangle\langle k| \langle l|m\rangle \end{aligned} \quad (2.64)$$

denotes partial trace over the second qubit and $\lambda_i^{\varrho_1}$ are the eigenvalues of ϱ_1 . The definition of Eq. 2.63 is motivated by the observation that tracing out one qubit of a Bell state (see section 2.2.1) leads to a maximally mixed state. In contrast, a pure separable state remains pure when one qubit is traced out. Hence, the entanglement of formation is a well suited measure as it quantifies the local mixedness of a state.

For mixed states, the entanglement of formation can be calculated via the convex-roof extension

$$E_F(\varrho) = \inf_{p_k, |\phi_k\rangle} \sum_k p_k S((\varrho_1)_k) \quad (2.65)$$

which is the least probable von Neumann entropy of any ensemble of pure states realizing ϱ . Interestingly, as will be discussed below, for bipartite mixed states, the entanglement of formation can be determined analytically via the concurrence.

Negativity

The negativity is closely related to the PPT criterion as it quantifies by how much the PPT criterion is violated. Formally, it is defined as

$$\mathcal{N}(\varrho) = (\|\varrho^{T_1} - 1\|)/2 = \sum_{\lambda_i < 0} |\lambda_i| \quad (2.66)$$

where $\|A\| = \text{Tr}[\sqrt{A^\dagger A}]$ is the trace norm of A and λ_i are the eigenvalues of ϱ^{T_1} . With the negativity determined, one can also calculate the logarithmic negativity which is defined as [164]

$$E_{\mathcal{N}}(\varrho) = \log_2 \|\varrho^{T_1}\|_1 = \log_2(2\mathcal{N}(\varrho) + 1). \quad (2.67)$$

2.2. Multiqubit states and entanglement

The negativity is convex but not additive, whereas, in contrast, the logarithmic negativity is additive but not convex. Both quantities fulfill the first three requirements of an entanglement measure and thus allow to quantify entanglement of an experimentally prepared state, however, as partial transposition is required, only at the price of full state tomography.

The negativity can also be generalized to the three-qubit case. The tripartite negativity is given by the geometric mean of the negativities of all bipartitions [165, 166]

$$\mathcal{N}_3(\varrho) = (\mathcal{N}_{1,23}(\varrho)\mathcal{N}_{2,13}(\varrho)\mathcal{N}_{3,12}(\varrho))^{1/3} \quad (2.68)$$

and allows to quantify the amount of W-type entanglement as it solely depends on the residual bipartite entanglement.

Concurrence

Another widely used entanglement measure is the concurrence. For pure states $|\psi\rangle$, it is formally defined as

$$C(|\psi\rangle) = \sqrt{2(1 - \text{Tr}[\varrho_1^2])} = \langle\psi|\Theta|\psi\rangle \quad (2.69)$$

where ϱ_1 denotes the reduced state after tracing out the second qubit (see section 2.2.4) and where Θ is an operator generating the anti-unitary transformation $\Theta\psi = (\sigma_y \otimes \sigma_y)\psi^*$ with the asterisk $*$ denoting complex conjugation [167]. From a physical point of view, $\langle\psi|\Theta|\psi\rangle$ can be interpreted as the overlap between ψ and its universal spin-flipped counterpart ψ^* . The generalization of the concurrence to mixed states is given in [168] as

$$C(\varrho) = \max(0, \lambda_1 - \lambda_2 - \lambda_3 - \lambda_4) \quad (2.70)$$

with λ_i the eigenvalues of the matrix $\sqrt{\sqrt{\varrho}(\sigma_y \otimes \sigma_y)\varrho^*(\sigma_y \otimes \sigma_y)\sqrt{\varrho}}$ in decreasing order. Interestingly, for pure two-qubit systems, the concurrence can be directly measured when using two copies of the same two-qubit state [169–172]. In case of mixed states, this approach allows to determine a lower bound on the concurrence [173, 174].

Despite of its very different origin, the concurrence is closely related to the entanglement of formation which can be expressed as a function of the concurrence

$$E_F(\varrho) = h((1 + \sqrt{1 - (C(\varrho))^2})/2) \quad (2.71)$$

with the binary entropy function given by $h(x) = x \log x - (1 - x) \log(1 - x)$. Please note that in the literature, the term $(C(\varrho))^2$ is also known as tangle $\tau(\varrho)$.

Geometric measure of entanglement

As already explained, the set of fully separable states forms a convex subset within the state space. The geometric measure of entanglement $E_G(|\chi\rangle)$ [175–177] quantifies the minimal distance between an entangled state $|\chi\rangle$ and the fully separable states, see Fig. 2.6. Formally, it is given by

$$E_G(|\psi\rangle) = 1 - \sup_{|\phi\rangle_{\text{sep}}} |\langle\phi_{\text{sep}}|\psi\rangle|^2. \quad (2.72)$$

For mixed states, as usually observed in experiments, the geometric measure of entanglement can be bounded from below, e.g., by the expectation value of a linear witness [178–180].

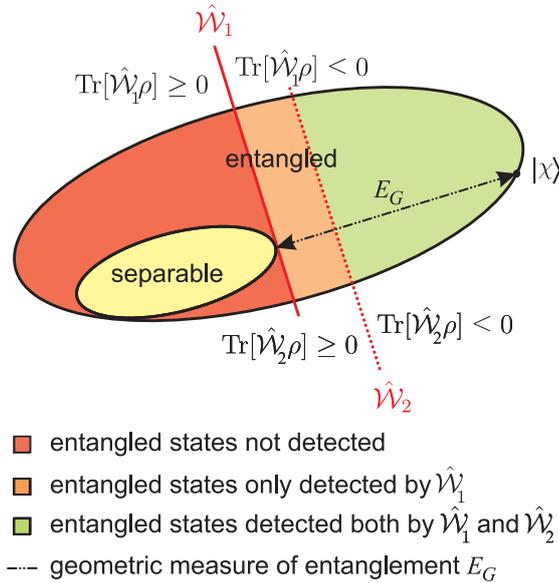


Figure 2.6.: The geometric measure of entanglement $E_G(|\chi\rangle)$ is defined as the minimal distance between an entangled state $|\chi\rangle$ and the set of separable states, here shown by the dashed line. It is possible to give an estimate for the geometric measure of entanglement via the expectation value of an entanglement witness.

Other measures

There are far more entanglement measures than can be discussed in this thesis. Therefore, a few more entanglement measures shall at least be named. The robustness [181] quantifies by how much a state can be mixed with a separable state such that the overall state is still entangled. The robustness plays a vital role in the noise tolerance of various entanglement detection criteria. The 3-tangle [182, 183] is another entanglement measure for pure three-qubit states

which allows to quantify the amount of GHZ-type entanglement. Together with the local entropy, it allows to distinguish between states belonging to the GHZ-class of tripartite entangled states and those belonging to the W-class. For a more elaborate treatment of entanglement, please refer to the excellent review articles [84–87, 132–134, 184, 185].

2.3. Experimental realization of photonic multiqubit states

In the following, a basic description of the process of SPDC will be given. Only the most relevant aspects which are necessary to understand the experiments described in chapters 3–5 will be discussed. For a more detailed treatment of the subject, the reader is referred to the rich special literature on SPDC in general [186–188] and [189–191] for pulsed sources. The subsequent discussion goes along the same line as [91, 192].

2.3.1. Spontaneous parametric down conversion

Spontaneous parametric down conversion is a nonlinear process mediated by non-inversion symmetric crystals where one pump photon converts into a pair of down conversion photons, called signal and idler. In the following, a simple model describing this process shall be given. In an anisotropic medium, an electric field \mathbf{E} induces a polarization \mathbf{P} , whose components can be expressed by a power series in terms of the electric field components [193],

$$P_i(\mathbf{E}) = \epsilon_0 \left(\sum_j \chi_i^{(1)j} E_j + \sum_{j,k} \chi_i^{(2)j,k} E_j E_k + \sum_{j,k,l} \chi_i^{(3)j,k,l} E_j E_k E_l \right), \quad (2.73)$$

with $i, j, k, l \in \{x, y, z\}$, ϵ_0 the dielectric constant and $\chi^{(m)}$ the electric susceptibility of order m . The size of $\chi^{(m)}$ decreases rapidly with larger m , $\chi^{(1)} = 1$, $\chi^{(2)} \approx 10^{-10} \text{cm/V}$ and $\chi^{(3)} \approx 10^{-17} \text{cm}^2/\text{V}^2$, and thus the coupling of the electric field to the crystal is dominated by the first terms of the power series in Eq. 2.73. For example, the term proportional to $\chi^{(2)}$ gives rise to the well-known three-wave mixing, the term proportional to $\chi^{(3)}$ is responsible for four-wave mixing etc. From now on, we will restrict our discussion to the process of SPDC which is described by $\chi^{(2)}$. Please note that in order to properly describe SPDC, a quantization of the respective fields is necessary [188].

Analogously to classical mechanics, both energy and linear momentum have to be conserved. Expressed in terms of the interacting waves, this yields

$$\hbar\omega_p = \hbar\omega_s + \hbar\omega_i + \Delta E \quad \text{and} \quad \hbar\mathbf{k}_p = \hbar\mathbf{k}_s + \hbar\mathbf{k}_i + \hbar\Delta\mathbf{k} \quad (2.74)$$

2. Fundamentals of quantum information theory

with $\Delta E = 0$ and $\Delta \mathbf{k} = 0$. The frequencies of the pump, signal and idler photon are labeled by ω_p , ω_s and ω_i . The absolute value of the wave vectors $|\mathbf{k}_j|$ is given by $|\mathbf{k}_j| = \omega_j n(\omega_j)/c$ where $n(\omega_j)$ denotes the refractive index at frequency ω_j with $j \in \{p, s, i\}$ and c is the speed of light. Eq. 2.74 is also known under the term phase-matching condition which reflects the property that only for $\Delta \mathbf{k} = 0$ considerable emission takes place. The most natural way to achieve phase-matching is to suitably choose \mathbf{k}_p , $\hbar\omega_p$ and the direction of the optical axis of a birefringent crystal. Light that is polarized parallel to the optical axis of a birefringent crystal is called extraordinary (abbreviated as *e*) and light that is polarized perpendicular to the optical axis is denoted ordinary (abbreviated as *o*). In detail, for phase-matching one utilizes the fact that for ordinary light the refractive index n_o is independent of the propagation direction, whereas for extraordinary light the refractive index n_e depends on the propagation direction. Hence, n_e becomes a function of the tilting angle Θ of the optical axis relative to \mathbf{k}_p , i.e. $n_e = n_e(\Theta)$. Then, for a given wave vector of the pump beam \mathbf{k}_p and, e.g., \mathbf{k}_s and \mathbf{k}_i lying in a plane parallel to the optical table as in the experiments performed in this work, one can achieve $\Delta \mathbf{k} = 0$ for a suitable choice of θ . Depending on the birefringent crystal, either of the refractive indices n_e and n_o can be larger than the other one. Crystals with $n_e > n_o$ are called positive and those with $n_e < n_o$ are called negative. Generally, the dependence of the refractive index from the wavelength is expressed by the so-called Sellmeier equation. The respective Sellmeier coefficients can be found, e.g., in [194]. Now, one can distinguish between two different types of phase-matching, namely, type I and type II. In both cases, the pump photon is *e*-polarized whereas the down-conversion photons are *o*-polarized for type I and *e*- and *o*-polarized for type II, respectively. Additionally, one distinguishes between collinear down-conversion where $\mathbf{k}_p \parallel \mathbf{k}_s \parallel \mathbf{k}_i$ and noncollinear down-conversion where all choices of \mathbf{k}_p , \mathbf{k}_s and \mathbf{k}_i are allowed that fulfill $\Delta \mathbf{k} = 0$. In the course of this work, noncollinear type I down conversion at a pump wavelength of 405 nm and collinear type II down-conversion at a pumpwavelength of 390 nm was used. In both cases, β -barium-borate (BBO) crystals were applied and only the degenerate case, i.e., signal and idler photons are emitted at the same wavelength, was considered.

As multiphoton states are in the focus of chapter 4 of this work, it is necessary to extend the above model and to consider higher order terms in the process of SPDC. Therefore, we will use the creation operator a^\dagger and, correspondingly, the annihilation operator a which have the properties

$$a^\dagger|n\rangle = \sqrt{n+1}|n+1\rangle \text{ and } a|n\rangle = \sqrt{n}|n-1\rangle \quad (2.75)$$

and where $n = \langle a^\dagger a \rangle$ is the photon number. Since all multiphoton experiments described in this thesis are based on a collinear type II down-conversion source, we

2.3. Experimental realization of photonic multiqubit states

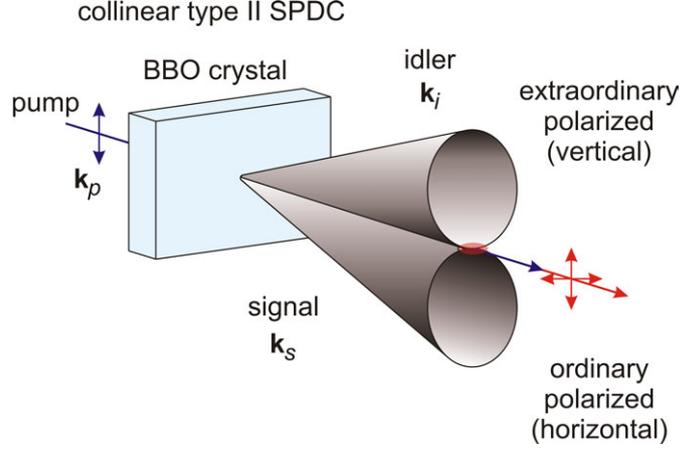


Figure 2.7.: In collinear type II down-conversion, the photons are emitted in two cones that intersect at just one line. Spatial filtering of the down-conversion photons is achieved by coupling into a single mode fiber

will restrict the following treatment to this particular case. There, the interaction Hamiltonian is given by [49]

$$\mathcal{H}_{\text{Int,col}} = i\hbar\kappa_{\text{col}}a_H^\dagger a_V^\dagger + h.c. \quad (2.76)$$

with κ_{col} proportional to the second order susceptibility $\chi^{(2)}$ and the absolute value of the pump field $|E_p|$. In type II collinear down conversion, the photons are emitted in two cones that have to be aligned by the aforementioned tilting of the crystal such that they intersect at just one line, see Fig. 2.7. Then the photons are collected by a single mode fiber which defines a spatial mode (spatial filtering). The state vector of the down converted fields is given by

$$|\psi\rangle_{\text{col}} = e^{-\frac{i}{\hbar} \int dt \mathcal{H}_{\text{Int,col}} t} |\text{vac}\rangle \quad (2.77)$$

where $|\text{vac}\rangle$ is the vacuum state. A Taylor expansion of Eq. 2.77 yields [91]

$$\begin{aligned} |\psi\rangle_{\text{col}} &= \frac{1}{\cosh \tau_{\text{col}}} \sum_{n=0}^{\infty} \frac{\tanh^n \tau_{\text{col}}}{n} (a_H^\dagger a_V^\dagger)^n |\text{vac}\rangle \\ &= \frac{1}{\cosh \tau_{\text{col}}} \sum_{n=0}^{\infty} \tanh^n \tau_{\text{col}} |n_H, n_V\rangle_a \end{aligned} \quad (2.78)$$

with $\tau_{\text{col}} \propto \kappa_{\text{col}} t$ and the twin Fock state $|n_H, n_V\rangle_a$ consisting of n_H horizontally polarized and n_V vertically polarized photons in mode a .

As the focus of this work is on the generation and characterization of multi-photon states, the dependence of the expected count rate from the pump power P_p shall be discussed. As can be seen from Eq. 2.78, the probability to generate $2n$ down-conversion photons is $|\langle \psi_{\text{col}} | n_H, n_V \rangle_a|^2$ and thus proportional to

2. Fundamentals of quantum information theory

$\tanh^{2n} \tau_{\text{col}}$. Using the approximation $\tanh \tau_{\text{col}} \approx \tau_{\text{col}}$ which is valid for $\tau_{\text{col}} \ll 1$ one obtains the count rate is proportional to

$$\tau_{\text{col}}^{2n} \propto |E_p|^{2n} \propto P_p^n. \quad (2.79)$$

Hence, the count rates for the emission of n photon pairs increase polynomially with order n with the pump power. Nevertheless, due to the small nonlinearities of crystals like BBO, the likelihood to generate n photon pairs is very low. Thus, one has to concentrate the pump power in short pulses which leads to a significantly increased creation probability per pulse.

Experimental implementation

As explained, in order to obtain reasonable count rates in multiphoton experiments based on SPDC, ultrashort laser pulses and high pump power are required. In the following, the photon source that was used for the experiments in publications P3.3, P4.1 and P4.2 and which fulfills both these requirements shall be described. More precisely, it will be discussed how SPDC photons are generated that are spatially, spectrally and temporarily indistinguishable, which is a prerequisite to use them for quantum information processing. A schematic drawing of the whole laser system including the SPDC source can be seen in Fig. 2.8. In a first step, a continuous wave solid state laser at 1064 nm pumps a neodymium yttrium vanadate (Nd:YVO₄) laser where up-conversion to 532 nm by second harmonic generation (SHG) takes place. The Nd:YVO₄ is stabilized to lase at a constant output power of 10 W and is used to pump a titanium:sapphire (Ti:Sa) oscillator. The Ti:Sa oscillator delivers short pulses with a duration of approximately 130 fs at a repetition rate of 80 MHz. The center wavelength is 780 nm and the total output power approximately 2W¹. In the next step, up-conversion to 390 nm in a 3 mm thick lithium-triborate (LBO) crystal is performed. The average power is approximately 800 mW, but for multiphoton experiments with up to six photons as performed in this work (see chapter 4), an even higher pump power is advantageous. Therefore, the light is coupled into an ultraviolet (UV) enhancement cavity [91, 195, 196] where a maximal UV pump power of up to 10 W can be achieved. For long-term stability of the cavity, two of its mirrors are mounted on piezo crystals, which allow for active stabilization using the Hänsch-Couillaud locking scheme [197]. Inside the bow-tie enhancement cavity, the light is focused on a 1 mm thick anti-reflection coated BBO crystal that is cut for collinear type II down-conversion. One of the cavity mirrors has a dichroic coating with high reflectivity for the pump photons at 390 nm and high transmission for the down-conversion photons at 780 nm. It is known that due to the birefringence of the

¹The three stages of the laser system are a TP80[®] power supply from Spectra Physics[®], a 10 W Millennia[®]XS from Spectra Physics[®], and a Tsunami[®] Ti:Sa oscillator from Spectra Physics[®]. The anti-reflection coated 1 mm thick BBO inside the cavity was obtained from Newlight Photonics[®].

2.3. Experimental realization of photonic multiqubit states

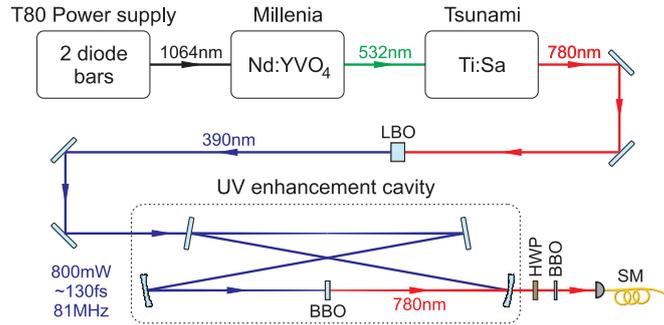


Figure 2.8.: Schematic drawing of the SPDC source used throughout this work. Its main components are a Ti:Sa oscillator and an UV enhancement cavity with a BBO crystal inside.

BBO, there is a transversal walk-off between signal and idler photons [191, 198]. The effect is partly compensated by a HWP outside the cavity which rotates horizontal to vertical polarization and vice versa and, subsequently, by sending the photons through a second BBO crystal with only half the thickness of the first one [50, 199]. Spatial filtering is achieved by coupling the down-conversion photons into a single mode fiber [200] which guides the photons to the linear optical setup where the actual experiment takes place. Additionally, since the spectra of the signal and idler photons are not identical [201–203], spectral filtering with an interference filter is performed². Please note that spectral filtering also increases the coherence length of the photons and thus washes out information about their creation time. For a more detailed description of the photon source utilized throughout this work, see [192, 195, 196].

2.3.2. Linear optical setups

In order to perform quantum information experiments with multiple photons one needs, additionally to the above described SPDC photon source, a so-called linear optical setup where the photons are further processed. In this context, the term linear means that only passive optical elements like wave plates, BSs or phase shifters are used. More precisely, this implies that photons may interfere with each other but a direct photon-photon interaction is excluded. However, this limitation does not lead to any restriction with respect to the experiments performed in the course of this work. In order to observe a multiphoton state, the photons are split up into certain well defined spatial modes by BSs. Due to the probabilistic nature of the BSs, the probability to observe a single photon in every spatial mode is below 1 and therefore a conditional detection scheme is applied. In detail, this means that only events where a single photon was detected in every mode are

²The interference filter used throughout this work is a MaxLine[®] filter from Semrock[®] with a center wavelength of 780 nm, a FWHM bandwidth of 3 nm, and a transmission of > 90%.

2. Fundamentals of quantum information theory

considered. Generally, every multiphoton experiment requires, depending on the state one wants to prepare, a different linear optical setup and, therefore, only the basic building blocks shall be discussed here. For the specific setups used for this thesis, refer to chapters 3 and 4 and the corresponding publications at the end of these chapters.

Birefringent phase shifters

The simplest components of linear optical setups with only a single input and output mode are birefringent phase shifters. They perform unitary transformations like, e.g., a rotation of the polarization or adding a phase between two orthogonal polarizations. In photonic setups, the polarization is commonly manipulated with QWPs and HWPs. Such devices can be implemented with uniaxial birefringent crystals like, e.g., quartz. Commonly, a wave plate consists of *two* crystals that are cut parallel to their optical axis and then optically contacted (or cemented) with their optical axis crossed. Depending on the different thickness of the two crystals, any phase shift can be implemented *exactly*, i.e., the phase shift is, e.g., $\pi/2$ and not $\pi/2 + 2n\pi$, $n \in \mathbb{N}$. The main advantage of using two crystals in a configuration with $n = 0$, also known as zero-order configuration, is its reduced sensitivity with respect to shifts of the wavelength or temperature changes in comparison to a wave plate consisting of a single crystal only. For $|z^+\rangle$ and $|z^-\rangle$ chosen as basis states and θ defined as the angle between the optical axis and the $|z^+\rangle$ direction, the corresponding transformations for a HWP and a QWP are given by [204]

$$\text{HWP}(\theta) = \begin{pmatrix} \cos 2\theta & \sin 2\theta \\ \sin 2\theta & -\cos 2\theta \end{pmatrix} = \sin 2\theta \sigma_x + \cos 2\theta \sigma_z \quad (2.80)$$

$$\begin{aligned} \text{QWP}(\theta) &= \begin{pmatrix} \cos^2 \theta - i \sin^2 \theta & (1+i) \cos \theta \sin \theta \\ (1+i) \cos \theta \sin \theta & -i \cos^2 \theta + \sin^2 \theta \end{pmatrix} \\ &= \frac{1}{2}((1-i)\mathbb{1} + 2(1+i) \cos \theta \sin \theta \sigma_x + (1+i) \cos 2\theta \sigma_z). \end{aligned} \quad (2.81)$$

In general, an arbitrary unitary transformation cannot be realized with a single wave plate. However, a combination of three wave plates, two QWPs and one HWP allows for arbitrary unitary transformations. In particular, such a transformation is given by $U(\theta_1, \theta_2, \theta_3) = \text{QWP}(\theta_1)\text{HWP}(\theta_2)\text{QWP}(\theta_3)$ with properly chosen angles θ_1 , θ_2 , and θ_3 . Please note that in order to realize a certain unitary transformation, the choice of angles θ_1 , θ_2 and θ_3 is not unique and, in many cases, these angles can only be found numerically.

In some applications, like, e.g., the compensation of unwanted phase shifts as induced by the Goos-Hänchen effect [205], it is necessary to shift a phase between $|z^+\rangle$ and $|z^-\rangle$. This can be realized by rotating a birefringent crystal with its

2.3. Experimental realization of photonic multiqubit states

optical axis parallel to $|z^-\rangle$ around $|z^-\rangle$ and thus varying the thickness of the crystal. In particular, the transformation can be expressed as

$$\begin{aligned} \text{PS}(\phi) &= \begin{pmatrix} i \cos \frac{\phi}{2} + \sin \frac{\phi}{2} & 0 \\ 0 & i \cos \frac{\phi}{2} - \sin \frac{\phi}{2} \end{pmatrix} \\ &= i \cos \frac{\phi}{2} \mathbb{1} + \sin \frac{\phi}{2} \sigma_z \end{aligned} \quad (2.82)$$

where ϕ designates the angle by how much the crystal is rotated. Experimentally, yttrium-vanadate (YVO₄) crystals are applied as phase shifter both as single crystals [191, 206] and in pairs with the optical axis of the crystals crossed [131, 207], analogously to the zero-order configuration common for wave plates.

Beam splitters

Beam splitters are four port devices with two input and two output ports and are used to split up or overlap spatial modes. They are characterized by their reflectivity R_H (R_V) and transmittance T_H (T_V) for the respective polarization. For a lossless BS with real transmittance and reflectivity, $R_H + T_H = 1$ and $R_V + T_V = 1$ holds. With the corresponding input and output modes labeled as in Fig. 2.9a), the transformation of a BS can be expressed by means of creation and annihilation operators

$$a_H^\dagger \longrightarrow \frac{1}{\sqrt{2}}(\sqrt{T_H}c_H^\dagger + i\sqrt{R_H}d_H^\dagger) \quad (2.83)$$

$$a_V^\dagger \longrightarrow \frac{1}{\sqrt{2}}(\sqrt{T_V}c_V^\dagger + i\sqrt{R_V}d_V^\dagger) \quad (2.84)$$

$$b_H^\dagger \longrightarrow \frac{1}{\sqrt{2}}(\sqrt{T_H}d_H^\dagger + i\sqrt{R_H}c_H^\dagger) \quad (2.85)$$

$$b_V^\dagger \longrightarrow \frac{1}{\sqrt{2}}(\sqrt{T_V}d_V^\dagger + i\sqrt{R_V}c_V^\dagger). \quad (2.86)$$

Note that a BS introduces a phase shift of $e^{i\pi/2}$ which is required to guarantee unitarity [208].

Projection measurements in arbitrary bases

In order to analyze quantum states, e.g., by means of state tomography as will be discussed in chapter 4, it is necessary to perform projection measurements on various different basis states. However, experimentally, it is often only possible to perform σ_z measurements, i.e, to project on $|z^+\rangle$ and $|z^-\rangle$. Therefore, the natural question arises how a measurement in an arbitrary basis like $\sigma(\theta, \phi)$ from Eq. 2.3 can be performed. More precisely, a general procedure has to be found such that one can project on the eigenvectors $|\psi(\theta, \phi)\rangle$ and $|\psi(\theta, \phi)\rangle_\perp$ of $\sigma(\theta, \phi)$.

2. Fundamentals of quantum information theory

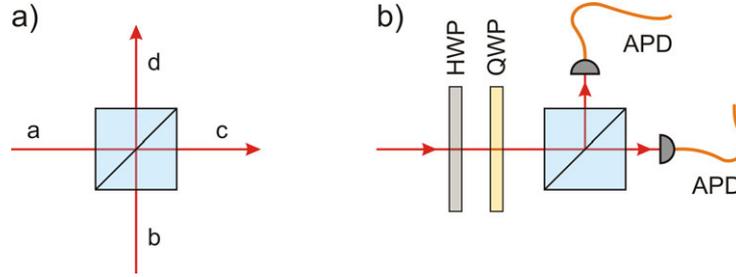


Figure 2.9.: The most important building blocks of linear optical setups are a) BSs and b) polarization analyses. A BS is defined by the transformations it performs on its input modes. A combination of a polarizing BS (PBS), i.e., a BS that transmits horizontally polarized light and reflects vertically polarized light, and a HWP and QWP, as shown in b), can be used to separate photons with two orthogonal eigenstates. Afterwards the photons are sent on APDs for detection.

The solution to this problem is to perform an appropriate unitary operation on the state to be analyzed before the projection measurements. One possibility is to first apply a $\text{HWP}(\theta_1)$, then a $\text{QWP}(\theta_2)$ and finally measure σ_z with a polarizing BS, see Fig. 2.9b). The angles θ_1 and θ_2 have to be chosen such that the relations

$$\text{QWP}(\theta_2)\text{HWP}(\theta_1)|\psi(\theta, \phi)\rangle = |z^+\rangle \quad (2.87)$$

$$\text{QWP}(\theta_2)\text{HWP}(\theta_1)|\psi(\theta, \phi)\rangle_{\perp} = |z^-\rangle \quad (2.88)$$

are fulfilled, or, equivalently,

$$\begin{aligned} \sigma_z &= |z^+\rangle\langle z^+| - |z^-\rangle\langle z^-| \\ &= (\text{QWP}(\theta_2)\text{HWP}(\theta_1))\sigma(\theta, \phi)(\text{QWP}(\theta_2)\text{HWP}(\theta_1))^{\dagger}. \end{aligned} \quad (2.89)$$

In many cases the angles θ_1 and θ_2 can only be determined numerically. However, for the most common basis settings, an analytic solution can be given as shown in Tab. 2.1.

Data acquisition

At every output of the polarization analysis, the photons are coupled into multimode fibers that are connected to actively quenched avalanche photo diodes³ (APD) which have a detection efficiency of 55% – 65% at 780 nm. The photon counting modules are connected to a coincidence logic based on a field programmable gate array (FPGA)⁴ [91], which allows for conditional detection. The pulse rate of the laser is measured with a photo diode and fed into the external

³Throughout this work, photon counting modules SPCM-AQ4C from Perkin Elmer[®] were used.

⁴The FPGA was a Virtex[™]-4 XC4VLX25-FF668-10 produced by Xilinx[®].

2.3. Experimental realization of photonic multiqubit states

measurement setting	angle θ_1 of HWP	angle θ_2 of QWP
σ_x	22.5°	0
σ_y	0	45°
σ_z	0	0

Table 2.1.: In many experimental realizations, only measurements in the σ_z basis can be directly carried out. Therefore, in order to perform measurements in other bases, like σ_x or σ_y , one first has to apply a unitary transformation. One possible implementation is a combination of a HWP and QWP at appropriate angles θ_1 and θ_2 in addition to a subsequent measurement in the σ_z basis.

clock of the FPGA. Consequently, the FPGA has a maximal detection rate of up to 81 MHz. The FPGA is capable of processing the detection signals of 12 photon counters which allows to perform multiphoton experiments with up to six photons. In order to carry out measurements in arbitrary bases, the HWP and QWP of every polarization analysis can be rotated with stepper motors⁵.

⁵All rotation stages are DRTM 40-D25-0SM440-(K) manufactured by OWIS®.

2. Fundamentals of quantum information theory

3. Entanglement detection by correlation measurements

In every experiment with entangled quantum states, it is crucial to first prove that the prepared states are actually entangled by means of suitable entanglement criteria (see section 2.2.3). In general, one can distinguish between entanglement criteria that are state independent but often require complete knowledge of the density matrix and criteria that are state dependent like, e.g., entanglement witnesses and require only few measurements. Obviously, one would like to have entanglement criteria at hand that are, at the same time, as general as possible but require only partial knowledge of the state.

In this chapter, such entanglement criteria will be presented. They are based on correlation functions which are directly accessible in experiments and which do not require to determine the complete density matrix. The chapter is divided into two parts. In the first part, a simple entanglement criterion proposed by Badziąg et al. [51] will be presented which is utilized to develop two schemes to detect entanglement at minimal experimental effort. The first scheme is based on Schmidt decomposition and requires at maximum three correlation measurements to prove entanglement of pure two-qubit states. The second scheme is based on the principle of correlation complementarity and allows to detect entanglement in an N -qubit state with an exponential speedup in comparison to random correlation measurements [53]. The description of these schemes stresses especially aspects that could not be discussed in publications P3.1 and P3.2.

As almost all entanglement criteria based on correlations rely on N -partite correlation functions [51, 60, 64–71], one might expect that the existence of N -partite correlations is a prerequisite for genuine N -partite entanglement. However, this assumption turns out to be wrong, as first shown by Kaszlikowski et al. [74]. In the second part of this chapter, a general construction principle for N -partite entangled states without N -partite correlations is presented. Interestingly, such states can even violate Bell inequalities which means, in consequence, that there is no local realistic model describing their properties as shown in publication P3.3.

3.1. Optimized state-independent entanglement detection

In the following section, it will be discussed how correlations can be used as an efficient tool to detect entanglement. Therefore, at first, it is necessary to explain how correlations are measured in practice. In section 2.2.1, it was explained that every quantum state ρ can be expressed by its correlations T_{i_1, i_2, \dots, i_N} which can formally be determined as

$$T_{i_1, i_2, \dots, i_N} = \text{Tr}[\rho \sigma_{i_1} \otimes \sigma_{i_2} \otimes \dots \otimes \sigma_{i_N}] \quad (3.1)$$

with $i_1, i_2, \dots, i_N \in \{0, x, y, z\}$. Experimentally, correlations are inferred from the expectation values of projection measurements. Therefore, the Pauli matrices in Eq. 3.1 are expressed in their eigenbases, i.e.,

$$\begin{aligned} \sigma_0 &= |x^+\rangle\langle x^+| + |x^-\rangle\langle x^-| \\ &= |y^+\rangle\langle y^+| + |y^-\rangle\langle y^-| \\ &= |z^+\rangle\langle z^+| + |z^-\rangle\langle z^-| \end{aligned} \quad (3.2)$$

$$\sigma_x = |x^+\rangle\langle x^+| - |x^-\rangle\langle x^-| \quad (3.3)$$

$$\sigma_y = |y^+\rangle\langle y^+| - |y^-\rangle\langle y^-| \quad (3.4)$$

$$\sigma_z = |z^+\rangle\langle z^+| - |z^-\rangle\langle z^-|. \quad (3.5)$$

For example, the correlation $T_{z,z}$ of a two-qubit state, can then be inferred from the following projection measurements

$$\begin{aligned} T_{z,z} &= \text{Tr}[\rho(\sigma_z \otimes \sigma_z)] \\ &= \text{Tr}[\rho(|z^+\rangle\langle z^+| - |z^-\rangle\langle z^-|) \otimes (|z^+\rangle\langle z^+| - |z^-\rangle\langle z^-|)] \\ &= \text{Tr}[\rho |z^+\rangle\langle z^+| \otimes |z^+\rangle\langle z^+|] - \text{Tr}[\rho |z^+\rangle\langle z^+| \otimes |z^-\rangle\langle z^-|] \\ &\quad - \text{Tr}[\rho |z^-\rangle\langle z^-| \otimes |z^+\rangle\langle z^+|] + \text{Tr}[\rho |z^-\rangle\langle z^-| \otimes |z^-\rangle\langle z^-|]. \end{aligned} \quad (3.6)$$

All other correlations can be determined correspondingly.

3.1.1. Experimentally friendly entanglement criterion

The entanglement criterion used throughout this chapter was originally proposed by Badziąg et al. [51] and originates from the following geometrical observation. Consider two real unit vectors \vec{s} and \vec{e} together with their scalar product: If the relation $\vec{s} \cdot \vec{e} < \vec{e} \cdot \vec{e}$ holds, then the two vectors cannot be identical, i.e., $\vec{s} \neq \vec{e}$. In a few steps, this simple geometrical fact will be transformed into a practical entanglement criterion [51]. For this purpose, a scalar product (\cdot, \cdot) for correlation tensors has to be introduced

$$(T^{\rho A}, T^{\rho B}) = \sum_{\substack{i_1, \dots, i_N \\ \in \{0, x, y, z\}}} T_{i_1, \dots, i_N}^{\rho A} T_{i_1, \dots, i_N}^{\rho B} \quad (3.7)$$

3.1. Optimized state-independent entanglement detection

where T^{ϱ_A} and T^{ϱ_B} are the correlation tensors of two states ϱ_A and ϱ_B . As explained in chapter 2, a state ϱ_{sep} is fully separable if it can be written as a convex sum of the form

$$\varrho_{\text{sep}} = \sum_i p_i (\varrho_1)_i \otimes \dots \otimes (\varrho_N)_i \quad (3.8)$$

with $p_i > 0$ and $\sum_i p_i = 1$. For such states also the correlation tensor $T^{\varrho_{\text{sep}}}$ becomes separable

$$T^{\varrho_{\text{sep}}} = \sum_i p_i T_i^{\text{prod}} \quad \text{with} \quad T_i^{\text{prod}} = T_i^{(1)} \cdot T_i^{(2)} \cdot \dots \cdot T_i^{(N)} \quad (3.9)$$

and where the $T_i^{(k)}$ ($k = 1, \dots, N$) describe single-qubit correlation tensors. Combining the geometrical observation mentioned above with Eq. 3.7 and Eq. 3.9, one obtains the following entanglement criterion:

If a state ϱ with correlation tensor T^ϱ is fully separable, then there must be a pure product state associated with the correlation tensor T^{prod} such that the inequality $(T^\varrho, T^{\text{prod}}) \geq (T^\varrho, T^\varrho)$ holds.

The validity of this inequality can easily be proven by assuming that the contrary is true, i.e., $(T^\varrho, T^{\text{prod}}) < (T^\varrho, T^\varrho)$. This would imply, due to the separability of T^ϱ , that

$$(T^\varrho, T^\varrho) = \sum_i p_i (T^\varrho, T_i^{\text{prod}}) < \sum_i p_i (T^\varrho, T^\varrho) = (T^\varrho, T^\varrho) \quad (3.10)$$

which is a contradiction. Hence, if the inequality

$$\max_{T^{\text{prod}}} (T^\varrho, T^{\text{prod}}) < (T^\varrho, T^\varrho) \quad (3.11)$$

is true, ϱ must be entangled. If one now restricts the indices i_1, \dots, i_N in Eq. 3.7 to $\{x, y, z\}$, i.e.,

$$(T^{\varrho_A}, T^{\varrho_B}) = \sum_{\substack{i_1, \dots, i_N \\ \in \{x, y, z\}}} T_{i_1, \dots, i_N}^{\varrho_A} T_{i_1, \dots, i_N}^{\varrho_B}, \quad (3.12)$$

the maximal attainable value of the left-hand side of Eq. 3.11 is given by [51]

$$T^{\text{max}} = \max_{\vec{m}_1 \otimes \dots \otimes \vec{m}_N} (T^\varrho, \vec{m}_1 \otimes \dots \otimes \vec{m}_N) \quad (3.13)$$

where the $\vec{m}_k = (T_x^{(k)}, T_y^{(k)}, T_z^{(k)})$ are unit vectors and can be associated with a pure state of the k th party. More precisely, T^{max} corresponds to the expectation value of a correlation measurement, where $\sigma^{(k)} = T_x^{(k)} \sigma_x + T_y^{(k)} \sigma_y + T_z^{(k)} \sigma_z$ is chosen as the local measurement basis for the k th qubit. Since the absolute value of a

3. Entanglement detection by correlation measurements

correlation can be 1 at maximum, this finally results in the desired entanglement criterion: A quantum state ρ is entangled if

$$\sum_{\substack{i_1, \dots, i_N \\ \in \{x, y, z\}}} (T_{i_1, \dots, i_N}^\rho)^2 > 1 \quad (3.14)$$

holds. From the experimental point of view, this is a remarkably positive result because it means that one can start measuring correlations according to Eq. 3.14 and, as soon as the threshold of Eq. 3.14 is surpassed, entanglement is detected and one can stop measuring. The challenge now is to find strategies that allow to break the threshold with as little experimental effort as possible.

3.1.2. Experimental Schmidt decomposition

The first strategy to surpass the limit of 1 in Eq. 3.14 is tailored to pure or close to pure two-qubit states and is based on a redefinition of the local measurement directions. Every pure two-qubit state can be expressed in a particular form [62, 209]

$$|\psi_S\rangle = \cos \theta |a\rangle |b\rangle + \sin \theta |a_\perp\rangle |b_\perp\rangle \quad \theta \in [0, \pi/4] \quad (3.15)$$

called *Schmidt decomposition* with local bases $|a\rangle/|a_\perp\rangle$ and $|b\rangle/|b_\perp\rangle$. The local bases are denoted as Schmidt bases and allow both parties to define their respective local measurements. For the first party, from now on called Alice, one obtains

$$\begin{aligned} \sigma_{x'} &= |a_\perp\rangle\langle a| + |a\rangle\langle a_\perp| \\ \sigma_{y'} &= i|a_\perp\rangle\langle a| - i|a\rangle\langle a_\perp| \\ \sigma_{z'} &= |a\rangle\langle a| - |a_\perp\rangle\langle a_\perp| \end{aligned} \quad (3.16)$$

with corresponding definitions of $\sigma_{x''}$, $\sigma_{y''}$ and $\sigma_{z''}$ for the second party called Bob. The main advantage of the Schmidt decomposition with respect to fast entanglement detection is that it leads to a ‘‘concentration’’ of the correlations. More precisely, it results in a diagonalization of the correlation tensor which can be interpreted as a 4×4 matrix in the two-qubit case, see Fig. 3.1. Then, the entanglement criterion Eq. 3.14 allows to detect entanglement of all entangled pure two-qubit states (i.e. $\theta \neq 0$) from measuring $T_{z',z''}^2$ and $T_{y',y''}^2$ since $T_{z',z''}^2 + T_{y',y''}^2 = 1 + T_{y',y''}^2 > 1$. However, generally, the computational basis does not coincide with the Schmidt basis and, therefore, a simple procedure to determine the Schmidt basis is required. In order to find such a procedure, in a first step, the basis vectors of the respective Schmidt bases are expressed in terms of the computational basis

$$\begin{aligned} |a\rangle &= \cos \xi_A |0\rangle + e^{i\varphi_A} \sin \xi_A |1\rangle \\ |a_\perp\rangle &= \sin \xi_A |0\rangle - e^{i\varphi_A} \cos \xi_A |1\rangle \\ |b\rangle &= \cos \xi_B |0\rangle + e^{i\varphi_B} \sin \xi_B |1\rangle \\ |b_\perp\rangle &= e^{i\delta} (\cos \xi_B |0\rangle - e^{i\varphi_B} \cos \xi_B |1\rangle). \end{aligned} \quad (3.17)$$

3.1. Optimized state-independent entanglement detection

		x''	y''	z''
		0	0	$1-2\sin^2\theta$
A \ B	B	0	0	$1-2\sin^2\theta$
x'	0	$\sin 2\theta$	0	0
y'	0	0	$-\sin 2\theta$	0
z'	$1-2\sin^2\theta$	0	0	1

Figure 3.1.: When expressing a pure two-qubit state in the Schmidt basis, entanglement can be detected from measuring the correlations $T_{z',z''}$ and $T_{y',y''}$. The local Bloch vectors of Alice ($T_{x',0}, T_{y',0}, T_{z',0}$) and Bob ($T_{0,x'}, T_{0,y'}, T_{0,z'}$) are represented by the blue shaded boxes.

For states that are not maximally entangled, the coefficients ξ_A and φ_A (ξ_B and φ_B correspondingly) can directly be determined from local measurements

$$\vec{\alpha} := \frac{(T_{x,0}, T_{y,0}, T_{z,0})}{\sqrt{T_{x,0}^2 + T_{y,0}^2 + T_{z,0}^2}} = (\sin 2\xi_A \cos \varphi_A, \sin 2\xi_A \sin \varphi_A, \cos 2\xi_A) \quad (3.18)$$

as can be proven by direct calculation. From Fig. 3.1, it can be seen that in the Schmidt basis only the local Bloch vector elements $T_{z',0}$ and $T_{0,z''}$ are non-zero. The easiest way for Alice to achieve this is to redefine her z direction according to

$$\sigma_z \longrightarrow \sigma_{z'} = \vec{\alpha} \cdot \vec{\sigma} = \alpha_x \sigma_x + \alpha_y \sigma_y + \alpha_z \sigma_z \quad (3.19)$$

where α_x , α_y , and α_z are the components of $\vec{\alpha}$. Then, she redefines the x and y direction by choosing two directions which are orthogonal with respect to $\vec{\alpha}$ and with respect to each other. One possible choice for example is

$$\begin{aligned} \sigma_x &\longrightarrow \sigma_{x'} = \frac{-\alpha_x \alpha_x \sigma_x - \alpha_y \alpha_z \sigma_y + (1 - \alpha_z^2) \alpha_z}{\sqrt{1 - \alpha_z^2}} \\ \sigma_y &\longrightarrow \sigma_{y'} = \frac{\alpha_y \sigma_x - \alpha_x \sigma_y}{\sqrt{1 - \alpha_z^2}}. \end{aligned} \quad (3.20)$$

If Bob redefines his local basis correspondingly, the Schmidt basis is established up to a phase δ [52, 53] which cannot be determined from local measurements alone and would require one correlation measurement involving both parties. However, from the experimental point of view, this does not really pose a restriction. As can be seen from Fig. 3.2, entanglement can still be detected since $T_{z',z''}^2 + T_{y',y''}^2 = 1 + (\sin 2\theta \cos \delta)^2 > 1$ for $\delta \neq \pi/2$ and $\delta \neq 3\pi/2$. In the case $T_{y',y''} = 0$, it is sufficient to measure $T_{x',y''}$ or $T_{y',x''}$ to surpass the limit.

For maximally entangled states $|\psi_{\max\text{-ent}}\rangle$, this scheme has to be adopted since maximally entangled states have vanishing local Bloch vectors. As maximally

3. Entanglement detection by correlation measurements

		x''	y''	z''
		B		
A		0	0	1-2sin ² θ
x'	0	sin2θcosδ	sin2θsinδ	0
y'	0	sin2θsinδ	-sin2θcosδ	0
z'	1-2sin ² θ	0	0	1

Figure 3.2.: When only local measurement are performed by Alice and Bob, their local Schmidt bases can only be determined up to a relative phase δ . The local Bloch vectors of Alice and Bob are represented by the blue shaded boxes.

entangled states admit infinity many Bloch decompositions [67], one of the two parties, like, e.g. Bob, can therefore freely chose a local coordinate system. Let us assume that he decides to stick with his basis definition. Then Bob implies a filter operation of the form $F = |0\rangle\langle 0| + \epsilon|1\rangle\langle 1|$ ($0 < \epsilon < 1$) which, if it is successful¹, results in a state that is *not* maximally entangled

$$(\mathbb{1} \otimes F)|\psi_{\max\text{-ent}}\rangle = \frac{1}{\sqrt{1 + \epsilon^2}}(|a\rangle|0\rangle + \epsilon|a_{\perp}\rangle|1\rangle) \quad (3.21)$$

and thus leads to an emerging local Bloch vector with Alice. Hence, Alice can redefine her local measurement directions according to Eq. 3.19 and Eq. 3.20. After Bob removes the filter, entanglement can again be detected by measuring $T_{z',z''}$ and $T_{y',y''}$. Please note that the filter operation can also be seen as a test which allows to distinguish between maximally entangled states and the maximally mixed state because for the maximally mixed state, the local Bloch vectors would not emerge after filtering.

In summary, the presented scheme allows to detect entanglement of all pure entangled states with six local measurements, if necessary, filtering by Bob together with Alice measuring her local Bloch vector a second time, and, at most, three correlation measurements. For more details concerning the performance and an experimental implementation of the scheme, see P3.1 and P3.2.

3.1.3. Decision tree

The second strategy to break the threshold of Eq. 3.14 is based on the principle of correlation complementarity [60, 63]. The term correlation complementarity denotes the fact that for a set of dichotomic, mutually anticommuting operators $\{\sigma_1, \sigma_2, \dots, \sigma_N\}$ the trade-off relation

$$T_{\sigma_1}^2 + T_{\sigma_2}^2 + \dots + T_{\sigma_N}^2 \leq 1 \quad (3.22)$$

¹Since Bob has to inform Alice whether the filtering operation was successful or not, this measurement is *not* local.

3.1. Optimized state-independent entanglement detection

holds where T_{σ_k} ($k = 1, \dots, N$) is the expectation value of σ_k . Consequently, this means that as soon as one of the expectation values is extremal, i.e., $T_{\sigma_j} = \pm 1$, then all T_{σ_k} with $k \neq j$ have to vanish. As only non-zero correlations contribute to the criterion of Eq. 3.14, the vanishing correlations do not have to be determined. Hence, correlation complementarity can be used to design a measurement scheme that reduces the experimental effort considerable.

In the two-qubit case, such a scheme can be given in a compact form, called decision tree. Alice and Bob chose independently from each other three orthogonal directions and agree to measure correlations along these directions. For simplicity, the measurement directions are abbreviated as x, y, z for both Alice and Bob. The

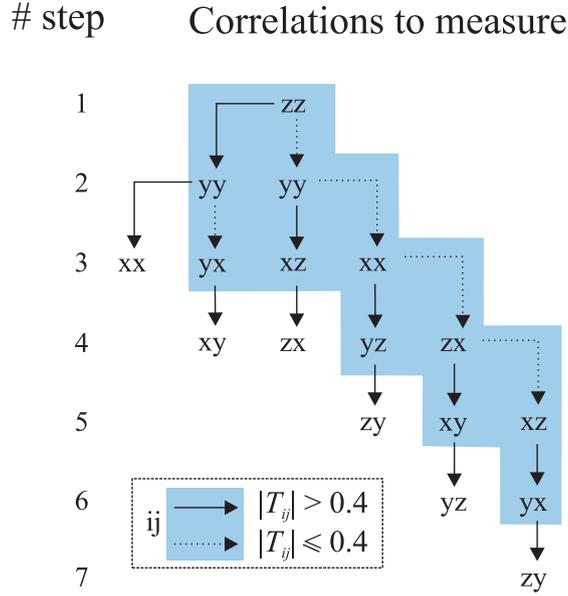


Figure 3.3.: The decision tree for fast entanglement detection of two qubits. Alice and Bob chose their local measurement directions independently from each other and agree to perform correlation measurements along these directions, i.e., no shared reference frame is required. The tree starts with a measurement of $T_{z,z}$ and for each following step one proceeds along the continuous arrow if the absolute value of the correlation measurement is > 0.4 . The scheme succeeds as soon as $\sum_{i,j \in \{x,y,z\}} T_{i,j}^2 > 1$. Every maximally entangled pure state with its Schmidt bases aligned along x, y , or z is detected as entangled within the blue shaded area.

sequence of the correlation measurements is given by the decision tree in Fig. 3.3 and the results of the corresponding measurements. It starts with a measurement of $T_{z,z}$ and continues along the continuous arrow if $|T_{z,z}| > 0.4$ and along the dashed one otherwise. At every following step, one measures the respective correlation and proceeds along the continuous or the dashed arrow depending on whether the absolute value of the result is above or below 0.4. The scheme ends

3. Entanglement detection by correlation measurements

as soon as entanglement is detected, i.e., as soon as $\sum_{i,j \in \{x,y,z\}} T_{i,j}^2 > 1$. The optimal value for the decision parameter was determined by numerical simulations showing that the best choice is 0.4.

In order to evaluate the performance of the decision tree, it has to be compared against randomly chosen correlation measurements. As random sampling, in con-

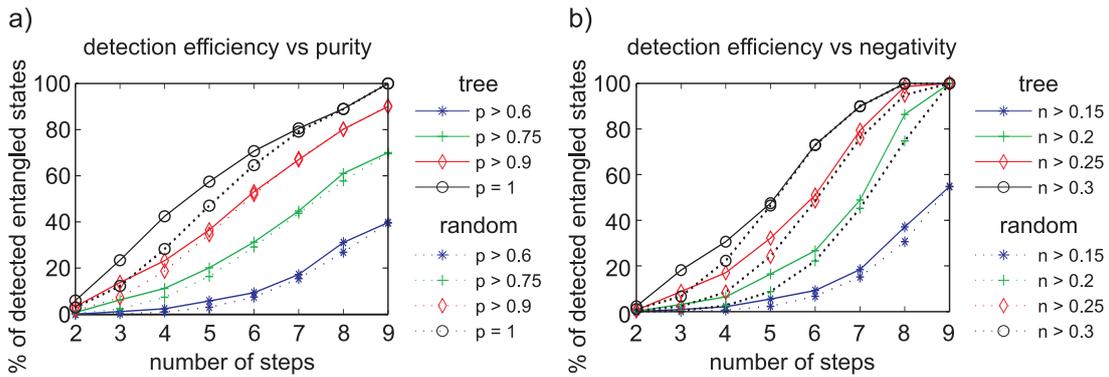


Figure 3.4.: Efficiency of the two-qubit decision for randomly mixed states according to the Haar measure. The efficiency of the tree in comparison with random measurements increases both with a) the purity of the state as well as with b) the entanglement as quantified by the negativity.

trast to the decision tree, does not take into account correlation complementarity, one expects a better performance for the tree. That this is actually the case can be seen in Fig. 3.4a) where the number of states detected by the decision tree is compared against random sampling for different purities. As expected, all pure entangled states are detected by both the tree and random sampling, however, the tree succeeds faster. The advantage of the tree is more pronounced for pure states. A similar behavior is observed for the negativity as shown in Fig. 3.4b). All states with a negativity > 0.2 are detected as entangled and the tree succeeds faster than random measurements. For more details on the performance of the decision tree and how it can be generalized to more than two qubits, see publication P3.2.

3.2. No correlation states

In this part, a particular class of states will be discussed which has vanishing N -partite correlations but still exhibits N -partite entanglement. The first member state of this class was discussed by Kaszlikowski et al. [74]. Here, the result of Kaszlikowski will be generalized and a constructive scheme that allows to construct such particular states will be given. Therefore, at first the concept of antistates has to be introduced, and second, the entanglement criterion of Eq. 3.11 has to be adopted such that genuine N -partite entanglement for vanishing N -partite correlations can be detected.

3.2.1. States and antistates

For every pure state $|\psi\rangle$ with an odd number of qubits, there exists an antistate $|\bar{\psi}\rangle$, such that, by evenly mixing $|\psi\rangle$ and $|\bar{\psi}\rangle$

$$\varrho_\psi^{nc} = \frac{1}{2} |\psi\rangle\langle\psi| + \frac{1}{2} |\bar{\psi}\rangle\langle\bar{\psi}| \quad (3.23)$$

all N -partite correlations of ϱ_ψ^{nc} vanish. In detail, the antistate has reversed correlations with respect to $|\psi\rangle$ for every odd number of parties whereas all other correlations are unchanged. In order to find the antistate, let us first consider $|\psi\rangle$ expressed in the computational basis

$$|\psi\rangle = \sum_{k_1, \dots, k_N=0}^1 \alpha_{k_1, \dots, k_N} |k_1, \dots, k_N\rangle \quad (3.24)$$

with normalized coefficients $\alpha_{k_1, \dots, k_N} \in \mathbb{C}$. Mathematically, $|\bar{\psi}\rangle$ can be obtained by applying local universal-not gates [210] $G_{\text{UNOT}} = (\sigma_z \sigma_x K)$ with K denoting complex conjugation in the computational basis. Direct calculation shows that $G_{\text{UNOT}} \sigma_x G_{\text{UNOT}}^\dagger = -\sigma_x$, $G_{\text{UNOT}} \sigma_y G_{\text{UNOT}}^\dagger = -\sigma_y$ and $G_{\text{UNOT}} \sigma_z G_{\text{UNOT}}^\dagger = -\sigma_z$ and thus only correlations for an odd number of observers are inverted. Application of G_{UNOT} to all the subsystems yields

$$|\bar{\psi}\rangle = \sum_{k_1, \dots, k_N=0}^1 (-1)^{k_1 + \dots + k_N} \alpha_{1-k_1, \dots, 1-k_N}^* |k_1, \dots, k_N\rangle \quad (3.25)$$

where the asterisk* stands for complex conjugation. Please note that although G_{UNOT} is an antiunitary operation, i.e., $G_{\text{UNOT}}^\dagger = -G_{\text{UNOT}}$, the antistate $|\bar{\psi}\rangle$ is always a well-defined quantum mechanical state. This construction principle can easily be generalized to mixed states by defining $\bar{\varrho} = G_{\text{UNOT}}^{\otimes N} \varrho (G_{\text{UNOT}}^{\otimes N})^\dagger$ which transforms every pure state in the decomposition of ϱ into its antistate. For states with all $\alpha_{k_1, \dots, k_N} \in \mathbb{R}$, complex conjugation is unnecessary and no-correlation states can be obtained by local operations alone.

One might now wonder if a similar construction principle exists also for states with an even number of qubits. Interestingly, the answer turns out to be negative as there is a counterexample where it can be shown that an antistate cannot exist. Let us consider the N -qubit GHZ state $|\text{GHZ}_N\rangle = 1/\sqrt{2}(|0\rangle^{\otimes N} + |1\rangle^{\otimes N})$ with 2^{N-1} nonvanishing full correlations in the xy plane, non-zero $T_{z, \dots, z}$, and $2^{N-1} - 1$ nonvanishing correlations between a smaller number of subsystems. All these correlations have absolute value 1. Let us now assume that the GHZ state has an antistate $|\overline{\text{GHZ}}_N\rangle$ with inverted correlations between all N parties (with no constraints applied on correlations for a smaller number of parties). Then, the

3. Entanglement detection by correlation measurements

fidelity between the GHZ state and its antistate, expressed by the corresponding correlation tensors $T_{i_1, \dots, i_N}^{\text{GHZ}_N}$ and $T_{i_1, \dots, i_N}^{\overline{\text{GHZ}_N}}$

$$\mathcal{F} = |\langle \text{GHZ}_N | \overline{\text{GHZ}_N} \rangle|^2 = \frac{1}{2^N} \sum_{\substack{i_1, \dots, i_N \\ \in \{0, x, y, z\}}} T_{i_1, \dots, i_N}^{\text{GHZ}_N} T_{i_1, \dots, i_N}^{\overline{\text{GHZ}_N}} < 0 \quad (3.26)$$

is negative because more than half of the correlations have opposite sign. Hence, for GHZ states with an even number of qubits, there are no antistates. Thus a general construction principle to generate antistates with an even number of qubits cannot exist.

3.2.2. Family of genuine entangled no correlation states

In this section, a whole family of entangled three-qubit states with vanishing tripartite correlations will be presented. As a starting point, let us consider the three-qubit pure state

$$|\phi(\alpha, \beta)\rangle = \sin \beta \cos \alpha |001\rangle + \sin \beta \sin \alpha |010\rangle + \cos \beta |100\rangle \quad (3.27)$$

with $\alpha, \beta \in (0, \pi/2)$. Note that $|\phi(\alpha, \beta)\rangle$ (with any local unitaries applied thereon) defines a whole subspace of tripartite entangled states within the three-qubit Hilbert space. Mixing $|\phi\rangle$ with its antistate

$$|\bar{\phi}(\alpha, \beta)\rangle = \sin \beta \cos \alpha |110\rangle + \sin \beta \sin \alpha |101\rangle + \cos \beta |011\rangle \quad (3.28)$$

defines a family of no-correlation states

$$\varrho_\phi^{nc}(\alpha, \beta) = 0.5 |\phi(\alpha, \beta)\rangle\langle\phi(\alpha, \beta)| + 0.5 |\bar{\phi}(\alpha, \beta)\rangle\langle\bar{\phi}(\alpha, \beta)|. \quad (3.29)$$

In the following, it will be shown that, despite of vanishing tripartite correlations, all member states of the family $\varrho_\phi^{nc}(\alpha, \beta)$ are genuinely tripartite entangled. Therefore, the entanglement criterion of Eq. 3.11 is adopted to

$$\max_{T^{\text{bi-sep}}} (T^e, T^{\text{bi-sep}}) < (T^e, T^e) \quad (3.30)$$

where optimization is performed over all biseparable pure states $\varrho_{\text{bi-sep}}$. If inequality 3.30 holds, then ϱ cannot be biseparable, i.e., it must be genuinely tripartite entangled. Please note that here the scalar product for correlations tensors as defined in Eq. 3.7 is used. As can be shown by direct calculation, the right-hand side of Eq. 3.30 is 4 for all member states $\varrho_\phi^{nc}(\alpha, \beta)$ of the family [54]. Hence, in order to decide for which values of α and β , the state $\varrho_\phi^{nc}(\alpha, \beta)$ is genuinely tripartite entangled, the left-hand side of Eq. 3.30 has to be determined by performing an optimization over all biseparable states. Since the optimization process is rather

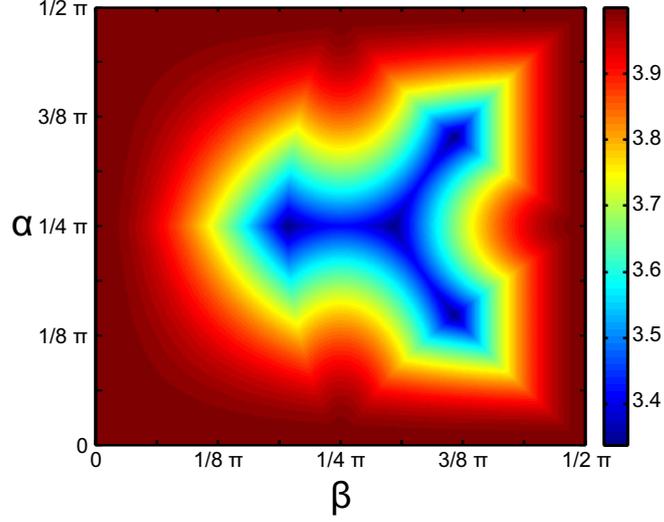


Figure 3.5.: Contour plot of the left-hand side of the entanglement criterion of Eq. 3.30 for states of the family $\varrho_\phi^{nc}(\alpha, \beta)$. If the value is below 4, the state is genuinely tripartite entangled. As can be seen, all states of the family with $\alpha, \beta \neq 0$ and $\alpha, \beta \neq \pi/2$ are entangled.

cumbersome, only the result is shown in Fig. 3.5 (for details please refer to the appendix of publication P3.3). As can be seen, the left-hand side of Eq. 3.30 is below 4 for all values of $\alpha, \beta \neq 0$ and $\alpha, \beta \neq \pi/2$ (which were excluded in the definition of $|\phi\rangle$ anyway). Hence, the entire family of states ϱ_ϕ^{nc} as defined in Eq. 3.29 is genuinely tripartite entangled although, by definition, it has vanishing tripartite correlations.

As an example state of the family ϱ_ϕ^{nc} , a mixture of the tree-qubit W state (obtained by setting $\alpha = \pi/3$ and $\beta = \cos^{-1}(1/\sqrt{3})$) with its antistate was considered, i.e.,

$$\varrho_W^{nc} = 0.5 |W\rangle\langle W| + 0.5 |\bar{W}\rangle\langle \bar{W}|. \quad (3.31)$$

Experimentally, such a state can be prepared by means of a four-photon symmetric Dicke state $|D_4^{(2)}\rangle$. As described by Wieczorek et al. [127], symmetric Dicke states can be used as a resource to prepare other entangled states by appropriate projection measurements. The W state for example can be obtained from a $|D_4^{(2)}\rangle$ state by projection of one qubit on $|1\rangle$. Correspondingly, the state \bar{W} can be obtained from projection on $|0\rangle$. The target state ϱ_W^{nc} can then be generated by evenly mixing the states W and \bar{W} which is in practice realized by tracing out the respective qubit. For details on the experimental implementation including a description of the utilized optical setup, see publication P3.3.

At this point one might ask if a vanishing correlation does also mean that one cannot infer another party's measurement result from one's own result. In

3. Entanglement detection by correlation measurements

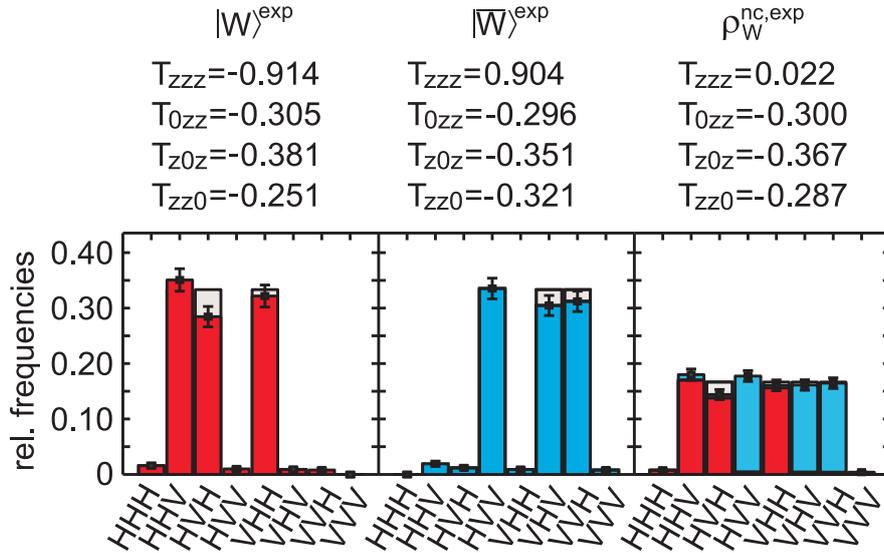


Figure 3.6.: Observed detection frequencies of the states $|W\rangle^{\text{exp}}$ (red) and $|\overline{W}\rangle^{\text{exp}}$ (blue) and $\rho_W^{\text{nc,exp}}$ (red and blue) when measuring in the $\sigma_z^{\otimes 3}$ basis. The gray bars represent the theoretically expected values. From the detection frequencies, one can infer the correlation value $T_{z,z,z}$ for the respective states. Although $T_{z,z,z}$ vanishes for the state $\rho_W^{\text{nc,exp}}$, from one's own, measurement result, one can still make predictions about the other parties outcomes with better success probability than for pure guessing due to nonvanishing two-party correlations, like, e.g., $T_{z,z,0}$.

order to answer this question, let us consider Fig. 3.6 which displays the relative frequencies that were measured in order to determine the correlation $T_{z,z,z} = \langle \sigma_z \otimes \sigma_z \otimes \sigma_z \rangle$ of the experimentally prepared states $|W\rangle$, $|\overline{W}\rangle$, and ρ_W^{nc} . For details see publication P3.3. Please note that the computational basis $|0\rangle$ and $|1\rangle$ used so far was encoded in the polarization degree of freedom of single photons as $|0\rangle \rightarrow |H\rangle$ and $|1\rangle \rightarrow |V\rangle$ where H stands for horizontal and V for vertical polarization. Commonly, horizontal polarization is associated with a parity of $+1$ and for vertical polarization one has a parity of -1 .

At first, one observes that the experimentally determined values of $T_{z,z,z}$ for $|W\rangle$ and $|\overline{W}\rangle$ are close to the theoretically expected values of -1 and 1 and that by mixing the two states, $T_{z,z,z}$ becomes zero. Also the experimentally determined bipartite correlations $T_{0,z,z}$, $T_{z,0,z}$, and $T_{z,z,0}$ are in good agreement with their theoretical values of $-1/3$. With these correlations, one can make predictions based on the parity of the measurement results of one or two parties. Let us therefore consider the W state that is shared between three parties. Since the value of $T_{z,z,z}$ is -1 , it follows that if the first two parties measure the same polarization, the third party has to measure vertical polarization. In case the first two parties measure orthogonal polarization, the third party has to measure horizontal polarization. Similarly, from the bipartite correlations one can conclude that two

parties will measure the same polarization with probability $1/3$ and orthogonal polarization with probability $2/3$. Hence, for the state $|W\rangle$ and correspondingly for $|\overline{W}\rangle$, from the parity of the measurement results of one or two parties, one can make predictions about the third party's outcome with better success probability than pure guessing.

Interestingly, similar statements can also be made for the state ϱ_W^{nc} . The experimental value of $T_{z,z,z}$ is close to the theoretical value of 0. The bipartite correlations $T_{0,z,z}$, $T_{z,0,z}$, and $T_{z,z,0}$ are also in good agreement with their expected values of $-1/3$. In contrast to the states $|W\rangle$ and $|\overline{W}\rangle$, here, due to the vanishing correlation $T_{z,z,z}$, it is not possible to infer the third party's measurement result from the product of the parities of the first two parties. However, from the bipartite correlations one can still infer that two parties will measure the same polarization with probability $1/3$ and orthogonal polarization with probability $2/3$. In summary, although the correlation $T_{z,z,z}$ vanishes for the state ϱ_W^{nc} , one can, based on the bipartite correlations $T_{0,z,z}$, $T_{z,0,z}$, and $T_{z,z,0}$, still make predictions with a success probability higher than for pure guessing.

As a closing remark to this section, it seems appropriate to shortly comment on the term correlation. In response to the paper by Kaszlikowski et al. [74], a vivid discussion on a proper definition of classical and quantum correlations set in. Bennet et al. [75], for example, proposed a set of postulates that every measure of genuine multipartite correlations should fulfill:

Postulate 1: If an N -partite state does not have genuine N -partite correlations and one adds a party in a product state, then the resulting $N+1$ -partite state does not have genuine N -partite correlations.

Postulate 2: If an N -partite state does not have genuine N -partite correlations, then local operations and unanimous postselection (which mathematically corresponds to the operation $\Lambda_1 \otimes \Lambda_2 \otimes \dots \otimes \Lambda_N$, where N is the number of parties and each Λ_i is a trace nonincreasing operation acting on the i th party's subsystem) cannot generate genuine N -partite correlations.

Postulate 3: If an N -partite state does not have genuine N -partite correlations, then if one party splits his subsystem into two parts, keeping one part for himself and using the other to create a new $N+1$ -st subsystem, then the resulting $N+1$ -partite state does not have genuine $N+1$ -partite correlations.

According to these postulates, the correlation function as defined in Eq. 2.41 cannot be called a measure of genuine N -partite correlations as can be directly shown. Let us therefore consider the N -qubit state $|00\dots 0\rangle$ which has a vanishing $T_{x,x,\dots,x}$ correlation, i.e., $T_{x,x,\dots,x} = \text{Tr}[|00\dots 0\rangle\langle 00\dots 0| \sigma_x \otimes \sigma_x \otimes \dots \otimes \sigma_x] = 0$. However, by applying the local unitary transformation $U = \text{HWP}(22.5^\circ)^{\otimes N}$ on the state one obtains $T_{x,x,\dots,x} = \text{Tr}[U |00\dots 0\rangle\langle 00\dots 0| U^\dagger \sigma_x \otimes \sigma_x \otimes \dots \otimes \sigma_x] = 1$

3. Entanglement detection by correlation measurements

which is a violation of postulate 2. Hence, $T_{x,x,\dots,x}$ and correspondingly all other correlations as defined in Eq. 2.41 cannot be a measure of genuine N -partite correlation according to the above postulates². Interestingly, Bennet et al. could not find a measure satisfying all their postulates.

In another approach, Giorgi et al. [76] used an information-theoretic definition of multipartite correlations. According to their definition, genuine classical tripartite correlations of 0.813 bit and tripartite genuine quantum tripartite correlations of 0.439 bit are assigned to the state ρ_W^{nc} . However, also the measure by Giorgi et al. does not fulfill all postulates by Bennet et al. At the time of writing, it is still an open question whether a rigorous and easy to calculate measure of genuine correlations which is based on plausible postulates can be given.

3.3. Publications

²Please note that correlations according to Eq. 2.41 cannot be a measure at all as they can increase under local unitary operations.

P3.1

Experimental Schmidt Decomposition and State Independent Entanglement Detection

Wiesław Laskowski,^{1,2,3} Daniel Richart,^{2,3} Christian Schwemmer,^{2,3} Tomasz Paterek,⁴ and Harald Weinfurter^{2,3}¹*Institute of Theoretical Physics and Astrophysics, University of Gdańsk, PL-80-952 Gdańsk, Poland*²*Max-Planck-Institut für Quantenoptik, Hans-Kopfermann-Strasse 1, D-85748 Garching, Germany*³*Department für Physik, Ludwig-Maximilians-Universität, D-80797 München, Germany*⁴*Centre for Quantum Technologies, National University of Singapore, 3 Science Drive 2, 117543 Singapore, Singapore*
(Received 17 November 2011; revised manuscript received 6 April 2012; published 11 June 2012)

We introduce an experimental procedure for the detection of quantum entanglement of an unknown quantum state with a small number of measurements. The method requires neither *a priori* knowledge of the state nor a shared reference frame between the observers and can thus be regarded as a perfectly state-independent entanglement witness. The scheme starts with local measurements, possibly supplemented with suitable filtering, which essentially establishes the Schmidt decomposition for pure states. Alternatively we develop a decision tree that reveals entanglement within few steps. These methods are illustrated and verified experimentally for various entangled states of two and three qubits.

DOI: [10.1103/PhysRevLett.108.240501](https://doi.org/10.1103/PhysRevLett.108.240501)

PACS numbers: 03.67.Mn, 03.65.Ta, 03.67.Bg, 42.50.Ex

Introduction.—Entanglement is the distinguishing feature of quantum mechanics and it is the most important resource for quantum information processing [1,2]. For any experiment it is thus of utmost importance to easily reveal entanglement, ideally with as little effort as possible. Common methods suffer from disadvantages. On the one hand, employing the Peres–Horodecki criterion [3,4] or evaluating entanglement measures, one can identify entanglement in arbitrary states; however, it requires full state tomography. On the other hand, various entanglement witnesses [4–10] can be determined with much fewer measurements, but they give conclusive answers only if the state under investigation is close to the witness state; i.e., they require *a priori* knowledge.

Recently, it has been shown that the existence of entanglement can be inferred from analyzing correlations among the measurement results on the subsystems of a quantum state. The properly weighted sum of correlations will overcome characteristic thresholds only if the state is entangled [11]. Here we further develop this approach to obtain a simple and practical method to detect entanglement of all pure states and some mixed states by measuring only a small number of correlations. Since the method is adaptive, it does not require *a priori* knowledge of the state nor a shared reference frame between the possibly remote observers, and thus it greatly simplifies the practical application. We describe two schemes. The first one essentially can be seen as a direct implementation of Schmidt decomposition, which identifies the maximal correlation directly. For bipartite pure systems, this approach can be divided conceptually into two stages: (i) calibration that establishes the experimental Schmidt decomposition [12,13] of a pure state by local measurements and suitable filtering and (ii) two correlation measurements to verify the entanglement criterion. The second scheme shows how to use a

decision tree to obtain a rapid violation of the threshold, thereby identifying entanglement.

Entanglement criterion.—For a two-qubit quantum state ρ , Alice and Bob observe correlations between their local Pauli measurements σ_k and σ_l , respectively. They are defined as the expectation values of the product of the two measurements, $T_{kl} = \text{Tr}[\rho(\sigma_k \otimes \sigma_l)]$, with the so-called correlation tensor elements $T_{kl} \in [1, -1]$. The local values T_{k0} (T_{0l}), with σ_0 being the identity operator, form the local Bloch vector of Alice (Bob). Using these measurements, a sufficient condition for entanglement can be formulated as [11,14]:

$$\sum_{k,l=x,y,z} T_{kl}^2 > 1 \Rightarrow \rho \text{ is entangled.} \quad (1)$$

For pure states this is also a necessary condition, while for mixed states care has to be taken. For mixed states, the likelihood of detecting the entanglement decreases with purity [15]. An extension of (1) can generally identify entanglement of an arbitrary mixed state, however, then losing the state independence [11,16]. Note two important facts. First, Eq. (1) can be seen as a state-independent entanglement witness, derived without any specific family of entangled states in mind. Second, to test whether the state is entangled, it is sufficient to break the threshold; i.e., it is neither required to measure all correlations nor to compute the density matrix [17]. Rather, it is now the goal to find strategies that minimize the number of correlation measurements. We show how this can be done by a particularly designed decision tree, or by identifying a Schmidt decomposition from local results and filtering when necessary.

Schmidt decomposition.—Consider pure two-qubit states. Any such state has a Schmidt decomposition

3. Entanglement detection by correlation measurements

$$|\psi_S\rangle = \cos\theta|a\rangle|b\rangle + \sin\theta|a_\perp\rangle|b_\perp\rangle, \quad \theta \in \left[0, \frac{\pi}{4}\right], \quad (2)$$

where the coefficients are real and the local bases $\{|a\rangle, |a_\perp\rangle\}$ and $\{|b\rangle, |b_\perp\rangle\}$ are called the Schmidt bases. Once the bases are known, Alice constructs her local measurements $\sigma_{z'} = |a\rangle\langle a| - |a_\perp\rangle\langle a_\perp|$ and $\sigma_{y'} = i|a_\perp\rangle\langle a| - i|a\rangle\langle a_\perp|$, and so does Bob in analogy. They can now detect entanglement with only two correlation measurements because $T_{z'z'}^2 + T_{y'y'}^2 = 1 + \sin^2 2\theta > 1$ for all pure entangled states. Note, the laboratories are not required to share a common reference frame.

In order to extract the Schmidt bases from experimental data, one starts with local measurements, determining the local Bloch vectors $\vec{\alpha}(\vec{\beta})$ of Alice (Bob). (Those vectors are related to the correlation tensor coefficients via $\alpha_i = T_{i0}/\sqrt{T_{x0}^2 + T_{y0}^2 + T_{z0}^2}$.) We consider two cases. First, suppose that a pure state has nonvanishing local Bloch vectors. Their directions define the Schmidt bases of Alice and Bob up to a global phase ϕ . Writing these bases in the computational basis

$$\begin{aligned} |a\rangle &= \cos\xi_A|0\rangle + e^{i\varphi_A}\sin\xi_A|1\rangle, \\ |a_\perp\rangle &= \sin\xi_A|0\rangle - e^{i\varphi_A}\cos\xi_A|1\rangle, \\ |b\rangle &= \cos\xi_B|0\rangle + e^{i\varphi_B}\sin\xi_B|1\rangle, \\ |b_\perp\rangle &= e^{i\phi}(\sin\xi_B|0\rangle - e^{i\varphi_B}\cos\xi_B|1\rangle), \end{aligned} \quad (3)$$

we see that the required coefficients can be inferred directly from the local Bloch vectors, $\vec{\alpha} = (\sin 2\xi_A \cos\varphi_A, \sin 2\xi_A \sin\varphi_A, \cos 2\xi_A)$ on Alice's side, and similarly for Bob. The global phase of $|b_\perp\rangle$ shows up as the relative phase in the decomposition (2); i.e., $|\psi_S\rangle = \cos\theta|a\rangle|b\rangle + \sin\theta e^{i\phi}|a_\perp\rangle|b_\perp\rangle$ (with $|b_\perp\rangle = e^{i\phi}|\tilde{b}_\perp\rangle$). It can be determined, for example, from the T_{yy} correlation as $\cos\phi = T_{yy}/\sqrt{1 - T_{x0}^2 - T_{y0}^2 - T_{z0}^2}$. If Bob would use the basis $\{|b\rangle, |\tilde{b}_\perp\rangle\}$ to build his observables $\sigma_{z''}$ and $\sigma_{y''}$, the corresponding correlations $T_{y'y''} = \sin 2\theta \cos\phi$ would vanish for $\cos\phi = 0$ and the two measurements $T_{z'z''}$ and $T_{y'y''}$ would not suffice to detect entanglement. In such a case, however, the other two correlations, $T_{x'x''}$ and $T_{y'x''}$, are nonzero, and can be used to reveal entanglement. Therefore, the determination of ϕ in the calibration is not essential if one accepts possibly one more correlation measurement.

Second, in the case of vanishing local Bloch vectors, the pure state under consideration $|\psi_m\rangle$ is maximally entangled and admits infinitely many Schmidt decompositions. In order to truly prove entanglement, Bob can thus freely choose some basis, say computational basis, for which the state will now be of the form $|\psi_m\rangle = \frac{1}{\sqrt{2}} \times (|a\rangle|0\rangle + |a_\perp\rangle|1\rangle)$. The basis of Alice can be found after filtering by Bob in his Schmidt basis: $F = |0\rangle\langle 0| + \varepsilon|1\rangle\langle 1|$. (For an actual implementation, see the experimental

section.) When Bob informs Alice that his detector behind the filter clicked, the initial state becomes

$$(1 \otimes F)|\psi_m\rangle \rightarrow \frac{1}{\sqrt{1 + \varepsilon^2}}(|a\rangle|0\rangle + \varepsilon|a_\perp\rangle|1\rangle). \quad (4)$$

Note that, due to filtering, a nonvanishing local Bloch vector emerges for Alice. Thus, the respective Schmidt basis can be found with the method described above and used for the evaluation of $T_{z'z'}^2 + T_{y'y'}^2$.

Decision tree.—Our second algorithm for entanglement detection does not even require calibration and also applies directly to mixed states. Alice and Bob choose three orthogonal local directions x, y , and z independently from each other and agree to only measure correlations along these directions. In Fig. 1 we show exemplarily which correlations should be measured in order to detect entanglement in a small number of steps. Starting with a measurement of T_{zz} , one continues along the solid (or dotted) arrow if the correlation is higher (or lower) than some threshold value (e.g., $1/2$ in Fig. 1). The tree is based on the principle of correlation complementarity [19–22]: in quantum mechanics there exist trade-offs for the knowledge of dichotomic observables with corresponding anti-commuting operators. For this reason, if the correlation $|T_{zz}|$ is big, correlations $|T_{zx}|, |T_{zy}|, |T_{xz}|$, and $|T_{yz}|$ have to be small, because their corresponding operators anticommute with the operator $\sigma_z \otimes \sigma_z$. Therefore, the next significant correlations have to lie in the xy plane of the correlation tensor, and thus the tree continues with a measurement of the T_{yy} correlation. This concept can be

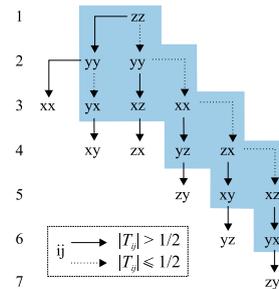


FIG. 1 (color online). The decision tree for efficient two-qubit entanglement detection. No shared reference frame is required between Alice and Bob; i.e., they choose their local x, y, z directions randomly and independently, which effectively gives rise to a basis $\{x_A, y_A, z_A\}$ for Alice and $\{x_B, y_B, z_B\}$ for Bob (not detailed in the figure or the main text). The scheme starts with measuring T_{zz} and follows at each step along the dashed arrow if the modulus of correlation is less than $\frac{1}{2}$ and otherwise along the continuous arrow. The algorithm succeeds as soon as $\sum T_{ij}^2 > 1$. The measurements in the blue shaded area suffice to detect all maximally entangled pure states with Schmidt-base vectors along x, y , or z .

generalized to multiqubit states. A decision tree for three qubits is given in the Supplemental Material [15]. The number of detected states grows with the number of steps through the decision tree. Since condition (1) is similar to the purity of a state, the scheme succeeds faster the more pure a state is (see Supplemental Material [15] for detailed analysis). Varying the threshold value does not lead to any significant changes in the statistic of detected states.

Finally, we connect both methods discussed here for the analysis of multiqubit states. A numerical simulation for pure states reveals that the correlation measurement along local Bloch vectors gives correlations close to the maximal correlations in more than 80% of cases. Therefore, these local directions give an excellent starting point for the decision tree.

Experiment.—For the demonstration of these new simple analysis methods we first use two photon-polarization entangled states. In the following, we will thus replace the computational basis states by horizontal ($|0\rangle \rightarrow |H\rangle$) and vertical ($|1\rangle \rightarrow |V\rangle$) linear polarization, respectively. The photon source (Fig. 2) is based on the process of spontaneous parametric down-conversion (SPDC), using a pair of crossed type I cut β -barium-borate (BBO) crystals pumped by a cw laser diode at a wavelength of $\lambda_{\text{pump}} = 402$ nm, with linear polarization of 45° . It emits pairs of horizontally and vertically polarized photons that superpose to the state $|\Psi\rangle = \frac{1}{\sqrt{2}}(|H\rangle|H\rangle + e^{i\delta}|V\rangle|V\rangle)$ [23]. The spectral bandwidth of the photons is reduced to 5 nm using interference filters, and two spatial emission modes are selected by coupling the photon pairs into two separate single-mode fibers.

For the purpose of preparing any pure two-qubit state, the polarization of each photon can be rotated individually by a set of quarter- (QWP) and half-wave plates (HWP) in each mode. By tilting an yttrium vanadate crystal (YVO_4) in front of the BBOs, the relative phase δ among the photon pairs can be set. Additionally, the state can be made asymmetric by removing a portion of vertically polarized light in one spatial mode with a Brewster plate

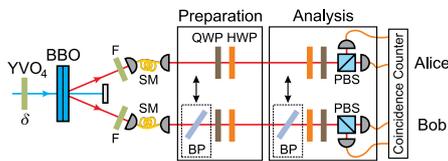


FIG. 2 (color online). Scheme of the experimental setup. The state $|\Psi\rangle = \frac{1}{\sqrt{2}}(|H\rangle|H\rangle + e^{i\delta}|V\rangle|V\rangle)$ is created by type I SPDC process. An yttrium vanadate crystal (YVO_4) is used to manipulate the phase δ of the prepared state. For preparation and analysis of the state, half- (HWP) and quarter-wave plates (QWP) are employed. Brewster plates (BP) can be introduced to make the state asymmetric and to perform the filter operation, respectively.

(BP). In the last step of the experiment, the polarization of each photon is analyzed with additional quarter- and half-wave plates and projection on $|H\rangle$ and $|V\rangle$ using a polarizing beam splitter (PBS). The local filtering of a maximally entangled state can be accomplished by placing a Brewster plate in front of the analysis wave plates. This Brewster plate reflects with a certain probability vertically polarized photons and, together with detection of a photon behind the Brewster plate, implements the filtering operation (4). Finally, the photons are detected by fiber-coupled single-photon detectors connected to a coincidence logic.

Experimental Schmidt decomposition.—Let us consider the state shown in Fig. 3(a). The protocol starts with Alice and Bob locally measuring the polarization of the photons, enabling them to individually determine the local Bloch vectors. For high efficiencies, which are possible in experiments with atoms or ions, the local measurements can indeed be done independently [24]. If nonvanishing local Bloch vectors can be identified, one can proceed to the next

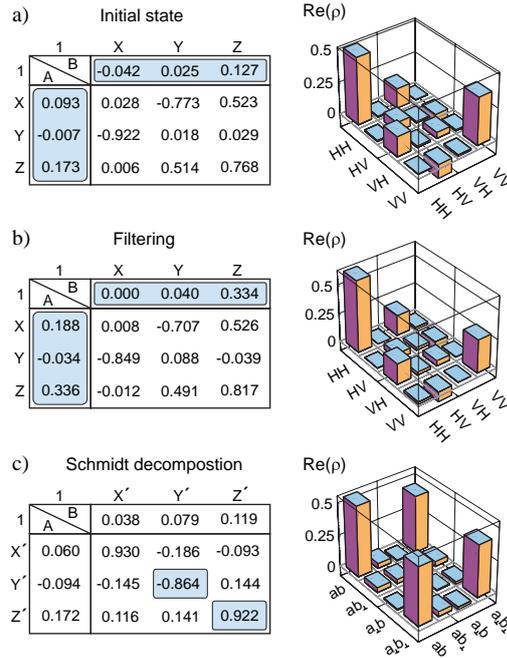


FIG. 3 (color online). Demonstration of Schmidt decomposition of a maximally entangled state prepared in unknown bases. The correlation tensor and corresponding density matrix are depicted for (a) the unknown state, (b) the state after applying local filtering, and (c) the state analyzed in the Schmidt bases. It is important to note that only the blue shaded elements of the correlation tensors will be measured, as this suffices to prove entanglement. The full correlation tensors and the corresponding states are only shown for completeness and didactic reasons.

3. Entanglement detection by correlation measurements

step. For the example here, the local expectation values are close to zero and filtering has to be applied. By using a Brewster plate in front of Bob's analysis wave plate, local Bloch vectors emerge as long as the filtering operation is successful [Fig. 3(b)] [25]. In this case, we obtain $T_{0l} = (0.000, 0.040, 0.334)$ and $T_{k0} = (0.188, -0.034, 0.336)$.

In the next step, Alice and Bob use their local Bloch vectors to realign their analyzers to the new local Schmidt bases $\{|a\rangle, |a_{\perp}\rangle\}$ and $\{|b\rangle, |b_{\perp}\rangle\}$, respectively. This process diagonalizes the correlation tensor, as depicted in Fig. 3(c). Therefore, it is only necessary to measure $T_{z'z''} = 0.922 \pm 0.015$ and $T_{y'y''} = -0.864 \pm 0.015$ to prove entanglement, since $T_{z'z''}^2 + T_{y'y''}^2 = 1.597 \pm 0.038 > 1$. Hence, 2×3 local measurements are needed in the first step of the algorithm, three combined measurements are needed for filtering if necessary, and finally only two correlation measurements have to be performed for entanglement detection.

Application of the decision tree.—In order to demonstrate the application of the decision tree, we will apply it to three states. For the first state $\frac{1}{\sqrt{2}}(|H\rangle|H\rangle + |V\rangle|V\rangle)$, whose correlation tensor is depicted in Fig. 4(a), the decision tree (Fig. 1) starts with the measurement of the correlation $T_{zz} = 0.980 \pm 0.015$ and continues with $T_{yy} = -0.949 \pm 0.015$. These two measurements already prove entanglement since $T_{zz}^2 + T_{yy}^2 = 1.869 \pm 0.041 > 1$. For a second state, $\frac{1}{\sqrt{2}}(|R\rangle|R\rangle + i|L\rangle|L\rangle)$, we obtain a correlation of $T_{zz} = -0.056 \pm 0.015$, close to zero [Fig. 4(b)]. Consequently, the next steps according to our algorithm (Fig. 1) are to determine the correlation $T_{yy} = 0.978 \pm 0.015$, followed by $T_{xz} = -0.959 \pm 0.015$, with their squares adding up to

a value of $1.879 \pm 0.041 > 1$ and hence proving entanglement. As a last example, we consider the initial state of Fig. 3. According to our decision tree, we need to measure $T_{zz} = 0.768 \pm 0.015$, $T_{yy} = 0.018 \pm 0.015$, and $T_{yx} = -0.922 \pm 0.015$, thus giving a value of $1.440 \pm 0.036 > 1$ and proving entanglement with only three steps.

Many qubits.—For the demonstration of multiqubit entanglement detection, we use two three-photon, polarization-entangled states: the W state [26] and the G state [27] (Fig. 5). In order to observe these states, a collinear type II SPDC source is used together with a linear setup to prepare the four-photon Dicke state $D_4^{(2)}$ [28,29]. Once the first photon is measured to be vertically polarized, the other three photons are projected into the W state. Similarly, the three-photon G state is obtained if the first photon is measured to be $+45^\circ$ polarized.

The protocol for entanglement detection starts with observers locally measuring the polarization of the photons, enabling them to individually determine the local Bloch vectors. For the G state we obtain $T_{i00} = (0.636, -0.008, -0.015)$, $T_{0j0} = (0.623, -0.092, 0.010)$, and $T_{00k} = (0.636, 0.070, 0.022)$. The local Bloch vectors suggest that the correlation T_{xxx} is big. Therefore, the decision tree starts with the measurement of $T_{xxx} = 0.904 \pm 0.025$ and continues with $T_{zzz} = -0.578 \pm 0.025$ (see Fig. 2 in the Supplemental Material [15]). These two measurements already prove entanglement because $T_{xxx}^2 + T_{zzz}^2 = 1.152 \pm 0.038 > 1$. For the W state, the local Bloch vectors $T_{i00} = (0.016, -0.070, 0.318)$, $T_{0j0} = (-0.010, -0.073, 0.308)$, and $T_{00k} = (-0.011, -0.0547, 0.319)$ suggest that now the correlation T_{zzz} is big. Indeed, we observe $T_{zzz} = -0.882 \pm 0.025$. The decision tree is the same as above but with local axes renamed as follows: $x \rightarrow z \rightarrow y \rightarrow x$. Therefore, the second measurement has to be T_{zyy} . With $T_{zyy} = 0.571 \pm 0.025$, we again prove entanglement as $T_{xxx}^2 + T_{zyy}^2 = 1.104 \pm 0.037 > 1$.

Conclusions.—We discussed and experimentally implemented two methods for fast entanglement detection for states about which we have no *a priori* knowledge. They are well suited for quantum communication schemes as the parties do not have to share a common reference frame, making the scheme insensitive to a rotation of the qubits during their transmission to the distant laboratories.

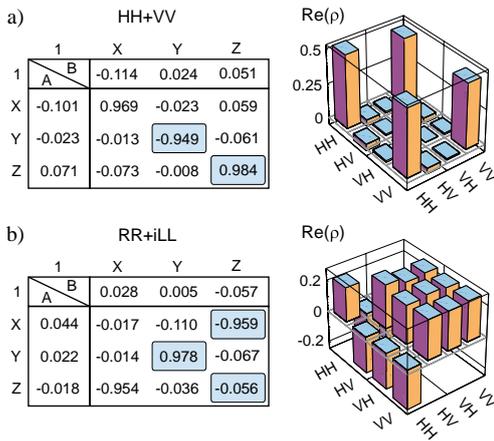


FIG. 4 (color online). Correlation tensors and density matrices of the experimental realization of two different states. The imaginary parts of the density matrices are negligible and therefore skipped. Using the decision tree, only the blue shaded correlations have to be measured for detecting entanglement. The errors of the correlations are <0.015 for (a) and <0.023 for (b).

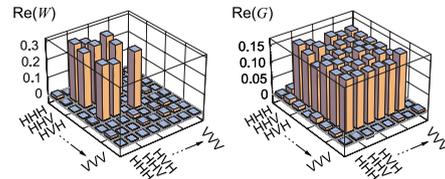


FIG. 5 (color online). Density matrices of the experimental realization of the G and W state. The corresponding fidelities are equal to 92.23% and 89.84%.

The two methods use a particularly simple and practical entanglement identifier [11]. One of them can be seen as experimental Schmidt decomposition and the other establishes a sequence of correlation measurements, leading to entanglement detection in a small number of steps.

We thank M. Żukowski for stimulating discussions. This work is supported by the EU project QESSENCE, the DAAD/MNiSW, the DFG-Cluster of Excellence MAP, and by the National Research Foundation and Ministry of Education in Singapore. W.L. is supported by the MNiSW Grant No. N202 208538 and by the Foundation for Polish Science. CS thanks QCCC of the Elite Network of Bavaria for support.

-
- [1] M. A. Nielsen and I. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, 2000).
- [2] R. Horodecki, P. Horodecki, M. Horodecki, and K. Horodecki, *Rev. Mod. Phys.* **81**, 865 (2009).
- [3] A. Peres, *Phys. Rev. Lett.* **77**, 1413 (1996).
- [4] M. Horodecki, P. Horodecki, and R. Horodecki, *Phys. Lett. A* **223**, 1 (1996).
- [5] M. Bourennane, M. Eibl, C. Kurtsiefer, S. Gaertner, H. Weinfurter, O. Gühne, P. Hyllus, D. Bruß, M. Lewenstein, and A. Sanpera, *Phys. Rev. Lett.* **92**, 087902 (2004).
- [6] B. M. Terhal, *Phys. Lett. A* **271**, 319 (2000).
- [7] M. Lewenstein, B. Kraus, J. I. Cirac, and P. Horodecki, *Phys. Rev. A* **62**, 052310 (2000).
- [8] D. Bruss, J. I. Cirac, P. Horodecki, F. Hulpke, B. Kraus, M. Lewenstein, and A. Sanpera, *J. Mod. Opt.* **49**, 1399 (2002).
- [9] O. Gühne and P. Hyllus, *Int. J. Theor. Phys.* **42**, 1001 (2003).
- [10] O. Gühne and G. Toth, *Phys. Rep.* **474**, 1 (2009).
- [11] P. Badziąg, Č. Brukner, W. Laskowski, T. Paterek, and M. Żukowski, *Phys. Rev. Lett.* **100**, 140403 (2008).
- [12] E. Schmidt, *Math. Ann.* **63**, 433 (1907).
- [13] A. Peres, *Quantum Theory: Concepts and Methods* (Kluwer Academic, Dordrecht, 1995).
- [14] R. Horodecki, P. Horodecki, and M. Horodecki, *Phys. Lett. A* **210**, 377 (1996).
- [15] See Supplemental Material at <http://link.aps.org/supplemental/10.1103/PhysRevLett.108.240501> for more details on the efficiency of the method based on the decision tree for pure and mixed states of two qubits and an algorithm that constructs the tree for many qubits.
- [16] W. Laskowski, M. Markiewicz, T. Paterek, and M. Żukowski, *Phys. Rev. A* **84**, 062305 (2011).
- [17] Although the effort to evaluate (1) generally scales as the one for full state tomography, we can directly evaluate the criterion from raw data without any further numerics required to reconstruct a physical density matrix [18].
- [18] D. F. V. James, P. G. Kwiat, W. J. Munro, and A. G. White, *Phys. Rev. A* **64**, 052312 (2001).
- [19] P. Kurzyński, T. Paterek, R. Ramanathan, W. Laskowski, and D. Kaszlikowski, *Phys. Rev. Lett.* **106**, 180402 (2011).
- [20] G. Toth and O. Gühne, *Phys. Rev. A* **72**, 022340 (2005).
- [21] S. Wehner and A. Winter, *J. Math. Phys. (N.Y.)* **49**, 062105 (2008).
- [22] S. Wehner and A. Winter, *New J. Phys.* **12**, 025009 (2010).
- [23] P. G. Kwiat, E. Waks, A. G. White, I. Appelbaum, and P. H. Eberhard, *Phys. Rev. A* **60**, R773 (1999).
- [24] Due to the low efficiency of the SPDC setup we have to use coincidence counts here. Using only the single counts, the local Bloch vector of Alice is slightly different. In this example we obtain $T_{0l} = (-0.073, -0.091, -0.059)$ and $T_{k0} = (0.031, 0.041, 0.037)$.
- [25] This in consequence means, that for the filtering Alice and Bob have to communicate with each other and hence this measurement is *not* local.
- [26] $|W\rangle = (|HHV\rangle + |HVV\rangle + |VHH\rangle)/\sqrt{3}$; W. Dür, G. Vidal, and J. I. Cirac, *Phys. Rev. A* **62**, 062314 (2000).
- [27] $|G\rangle = (|W\rangle + |\bar{W}\rangle)/\sqrt{2}$, where $|\bar{W}\rangle = (|VVH\rangle + |HVV\rangle + |VHV\rangle)/\sqrt{3}$; A. Sen(De), U. Sen, and M. Żukowski, *Phys. Rev. A* **68**, 032309 (2003).
- [28] $|D_4^{(2)}\rangle = (|HHVV\rangle + |HVHV\rangle + |VHHV\rangle + |HVHV\rangle + |VHVH\rangle + |VVHH\rangle)/\sqrt{6}$; R. Krischek, C. Schwemmer, W. Wieczorek, H. Weinfurter, P. Hyllus, L. Pezze, and A. Smerzi, *Phys. Rev. Lett.* **107**, 080504 (2011).
- [29] N. Kiesel, C. Schmid, G. Toth, E. Solano, and H. Weinfurter, *Phys. Rev. Lett.* **98**, 063604 (2007).

3. Entanglement detection by correlation measurements

Experimental Schmidt Decomposition and State Independent Entanglement Detection SUPPLEMENTARY INFORMATION

Wiesław Laskowski,^{1,2,3} Daniel Richart,^{2,3} Christian Schwemmer,^{2,3} Tomasz Paterek,⁴ and Harald Weinfurter^{2,3}

¹*Institute of Theoretical Physics and Astrophysics,
University of Gdańsk, PL-80-952 Gdańsk, Poland*

²*Max-Planck-Institut für Quantenoptik, Hans-Kopfermann-Strasse 1, D-85748 Garching, Germany*

³*Department für Physik, Ludwig-Maximilians-Universität, D-80797 München, Germany*

⁴*Centre for Quantum Technologies, National University of Singapore, 3 Science Drive 2, 117543 Singapore, Singapore*

In this Supplementary Information we discuss the efficiency of the method based on the decision tree for pure and mixed states of two qubits and present an algorithm that constructs the tree for many qubits.

DECISION TREE

This section is devoted to study the efficiency of the decision tree algorithm described in the main text. Some results in this section are analytical and some are numerical. In all our numerical investigations (unless explicitly stated otherwise) we used the decision tree of the main text (for two qubits) and in cases when going through the whole tree did not reveal entanglement we augmented it with additional measurements of those correlations which were not performed until that moment. The order of the additional measurements also results from the correlation complementarity (anti-commutation relations) [1]. With every remaining measurement we associate the “priority” parameter

$$P_{ij} = \sum_{k \neq i} P_{ij}(T_{kj}) + \sum_{l \neq j} P_{ij}(T_{il}), \quad (1)$$

that depends on the measured correlation tensor elements of the decision tree in the following way

$$P_{ij}(T_{mn}) = \begin{cases} T_{mn}^2 & \text{if } T_{mn} \text{ was performed before,} \\ 0 & \text{else.} \end{cases} \quad (2)$$

According to the correlation complementarity there is a bigger chance that this correlation is significant if the value of the corresponding parameter is small. Therefore, the correlations T_{ij} with lower values of P_{ij} are measured first.

Let us consider the following example. The measured correlations of the decision tree are as follows: $T_{zz} = 0.7$, $T_{xx} = 0.1$, and $T_{yy} = 0.4$. Therefore, $P_{xy} = P_{yx} = T_{xx}^2 + T_{yy}^2 = 0.17$, $P_{xz} = P_{zx} = T_{xx}^2 + T_{zz}^2 = 0.5$, and $P_{zy} = P_{yz} = T_{zz}^2 + T_{yy}^2 = 0.65$. Accordingly the order of the remaining measurements is as follows: first measure xy , then yx , next xz , zx , yz and zy .

Two qubits

Werner states

As an illustration of how the decision tree works for a well-known class of mixed states we first consider Werner states. It turns out that not all entangled states of the family can be detected.

Consider a family of states

$$\rho = p |\psi^-\rangle \langle \psi^-| + (1-p) \frac{1}{4} \mathbb{1}, \quad (3)$$

where $|\psi^-\rangle = \frac{1}{\sqrt{2}}(|01\rangle - |10\rangle)$ is the Bell singlet state, $\frac{1}{4} \mathbb{1}$ describes the completely mixed state of white noise, and p is a probability. Its correlation tensor, written in the same coordinate system for Alice and Bob, is diagonal with entries $T_{xx} = T_{yy} = T_{zz} = -p$, arising from the contribution of the entangled state. The states (3) are entangled if and only if $p > \frac{1}{3}$, whereas the decision tree reveals that these states are entangled for $p > \frac{1}{\sqrt{3}} \approx 0.577$. This is because only three elements contribute to the criterion. Note that the value of the decision parameter is irrelevant here. Furthermore, a random choice of local coordinate directions does not help. Although more steps would be involved in the decision tree the sum of squared correlations is invariant under local unitary operations and therefore only for $p > \frac{1}{\sqrt{3}}$ entanglement is detected for the Werner state.

Entangled state mixed with colored noise

An exemplary class of density operators for which the decision tree detects all entangled states is provided by:

$$\gamma = p |\psi^-\rangle \langle \psi^-| + (1-p) |01\rangle \langle 01|, \quad (4)$$

where entanglement is mixed with colored noise $|01\rangle$ bringing anti-correlations along local z axes. This state has the following non-vanishing elements of its correlation tensor $T_{xx} = T_{yy} = -p$ and $T_{zz} = -1$. Therefore, the decision tree allows detection of entanglement for this

class of states in two steps. Note that the state is entangled already for an infinitesimal admixture of the Bell singlet state as can be shown for a random choice of local coordinate systems.

Random mixed states

Fig. 1 shows how the efficiency of the algorithm grows with the purity of tested states. The efficiency is measured by the fraction of detected entangled states obtained from extensive Monte Carlo sampling in the two qubit state space. For nine steps the algorithm detects all the pure states. This is expected because Eq. (1) of the main text is a necessary and sufficient condition for the detection of entanglement of pure states.

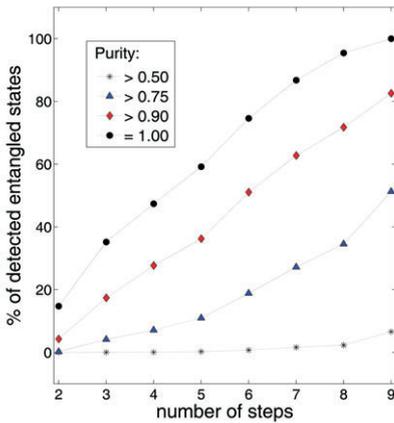


FIG. 1: Efficiency of the decision tree for two qubit random mixed states.

Many qubits

Algorithm for generation of the tree

The principle behind the decision tree is the correlation complementarity [1]. Correlation complementarity states that for a set of dichotomic anti-commuting operators $\{\alpha_1, \dots, \alpha_k\}$ the following trade-off relation is satisfied by all physical states:

$$T_{\alpha_1}^2 + \dots + T_{\alpha_k}^2 \leq 1, \quad (5)$$

where T_{α_1} is the average value of observable α_1 and so on. Therefore, if one of the average values is maximal, ± 1 , the other anti-commuting observables have vanishing averages. This motivates taking only sets of commuting operators as different branches of the decision tree.

Decision tree for three qubits

This tree is an example of the application of the algorithm presented in the previous section. Fig. 2 shows only one branch of the whole tree.

A numerical simulation reveals that the correlation measurement along local Bloch vectors gives correlations close to the maximal correlations of a pure multi-qubit state in more than 80% of the cases. Therefore, these local directions give an excellent starting point of the decision tree.

Exemplarily, the branch begins with T_{xxx} assuming that this correlation is big. If the local Bloch vectors indicate correlation along a different directions it is advantageous to correspondingly change the elements of the decision tree (see main text).

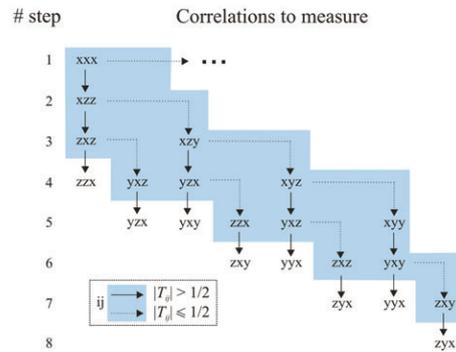


FIG. 2: One branch of the decision tree for three qubits.

[1] P. Kurzyński, T. Paterek, R. Ramanathan, W. Laskowski, and D. Kaszlikowski, Phys. Rev. Lett. **106**, 180402 (2011)

3. Entanglement detection by correlation measurements

P3.2

PHYSICAL REVIEW A **88**, 022327 (2013)

Optimized state-independent entanglement detection based on a geometrical threshold criterion

Wiesław Laskowski,¹ Christian Schwemmer,^{2,3} Daniel Richart,^{2,3} Lukas Knips,^{2,3} Tomasz Paterek,^{4,5} and Harald Weinfurter^{2,3}

¹*Institute of Theoretical Physics and Astrophysics, University of Gdańsk, PL-80-952 Gdańsk, Poland*

²*Max-Planck-Institut für Quantenoptik, Hans-Kopfermann-Strasse 1, D-85748 Garching, Germany*

³*Department für Physik, Ludwig-Maximilians-Universität, D-80797 München, Germany*

⁴*School of Physical and Mathematical Sciences, Nanyang Technological University, Singapore*

⁵*Centre for Quantum Technologies, National University of Singapore, Singapore*

(Received 11 June 2013; published 21 August 2013)

Experimental procedures are presented for the rapid detection of entanglement of unknown arbitrary quantum states. The methods are based on the entanglement criterion using accessible correlations and the principle of correlation complementarity. Our first scheme essentially establishes the Schmidt decomposition for pure states, with few measurements only and without the need for shared reference frames. The second scheme employs a decision tree to speed up entanglement detection. We analyze the performance of the methods using numerical simulations and verify them experimentally for various states of two, three, and four qubits.

DOI: 10.1103/PhysRevA.88.022327

PACS number(s): 03.67.Mn

I. INTRODUCTION

Entanglement is one of the most fundamental features of quantum physics and is considered to be the key resource for quantum information processing [1–3]. In order to detect entanglement, highly efficient witness operators are widely used nowadays [4–10]. However, these operators give conclusive answers only for states close to the target state. To detect entanglement of arbitrary states, positive, but not completely positive, maps [1,4] are the most universal entanglement identifiers. However, they are laborious to use as they require full state tomography. Therefore, more efficient schemes to detect entanglement are most desired.

It has been recently shown that the presence of entanglement in a quantum state is fully characterized by suitable combinations of experimentally accessible correlations and expectation values of local measurements [11]. This enables a simple and practical method to reveal entanglement of all pure states and some mixed states by measuring only few correlations [12]. Since the method is adaptive, it does not require *a priori* knowledge of the state nor a shared reference frame between the possibly remote observers, and thus greatly simplifies the practical application.

Here we extend these results and analyze in detail the possible performance of two schemes for entanglement detection. The first one can be seen as an experimental implementation of Schmidt decomposition, which identifies the maximal correlations through local measurements only. The second scheme shows how to deduce a strategy (decision tree) to find the maximal correlations of an unknown state and obtain a rapid violation of the threshold, identifying entanglement even for an arbitrary number of qubits. The physical principle behind both of our schemes is correlation complementarity [13]. It makes use of trade-offs between correlations present in quantum states. Once a measured correlation is big, other related correlations have to be small, and it is advantageous to move to measurements of the remaining correlations. This simplifies the entanglement detection scheme as a lower number of correlation measurements is required.

II. ENTANGLEMENT CRITERION

A quantum state is entangled if the sum of squared measured correlations exceeds a certain bound [11]. This identifier thus neither requires the measurement of all correlations in a quantum state nor the reconstruction of the density matrix. Rather, it is now the goal to find strategies that minimize the number of correlation measurements. We show how this can be done in different ways described in the subsequent sections. The first method is to identify a Schmidt decomposition from local results and filtering when necessary; the second is a particularly designed decision tree based on correlation complementarity.

Any N -qubit density matrix can be expressed as

$$\rho = \frac{1}{2^N} \sum_{\mu_1, \dots, \mu_N=0}^3 T_{\mu_1, \dots, \mu_N} \sigma_{\mu_1} \otimes \dots \otimes \sigma_{\mu_N}, \quad (1)$$

where $\sigma_{\mu_n} \in \{\sigma_0, \sigma_x, \sigma_y, \sigma_z\}$ is the respective local Pauli operator of the n th party (σ_0 being the identity matrix) and the real coefficients $T_{\mu_1, \dots, \mu_N} \in [-1, 1]$ are the components of the correlation tensor \hat{T} . They are given by the expectation values of the products of local Pauli observables, $T_{\mu_1, \dots, \mu_N} = \text{Tr}[\rho(\sigma_{\mu_1} \otimes \dots \otimes \sigma_{\mu_N})]$, and can be determined by local measurements performed on each qubit.

For N -qubit states, pure or mixed, the following sufficient condition for entanglement holds [11]:

$$\sum_{i_1, \dots, i_N=1}^3 T_{i_1, \dots, i_N}^2 > 1 \Rightarrow \rho \text{ is entangled.} \quad (2)$$

Note that to prove that a state is entangled, it is sufficient to break the threshold, i.e., in general it is not necessary to measure all correlations. Using fundamental properties of the correlation tensor, we design schemes to minimize the number of required correlation measurements.

III. SCHMIDT DECOMPOSITION

Any pure state of two qubits admits a Schmidt decomposition [14,15],

$$|\psi_S\rangle = \cos \theta |a\rangle |b\rangle + \sin \theta |a_\perp\rangle |b_\perp\rangle, \quad \theta \in [0, \frac{\pi}{4}], \quad (3)$$

where the local bases $\{|a\rangle, |a_\perp\rangle\}$ and $\{|b\rangle, |b_\perp\rangle\}$ are called the Schmidt bases of Alice and Bob, respectively.

This is an elementary description of bipartite pure quantum states, where the existence of a second term in the decomposition directly indicates entanglement. In addition, in the Schmidt bases, the correlation tensor of a two-qubit state takes a particularly simple form and shows maximal correlations in the state. Therefore, finding the Schmidt bases can be regarded as a redefinition of the measuring operators relative to the state and thus leads to rapid entanglement detection, in at most three subsequent measurements of correlations.

Once the bases are known, Alice constructs her local measurements, $\sigma_{z'} = |a\rangle\langle a| - |a_\perp\rangle\langle a_\perp|$ and $\sigma_{y'} = i|a_\perp\rangle\langle a| - i|a\rangle\langle a_\perp|$, and so does Bob. They can now detect entanglement by using the simple criterion (2) with only two correlation measurements because $T_{z'z'}^2 + T_{y'y'}^2 = 1 + \sin^2 2\theta > 1$ for all pure entangled states. Note that since the bases of Alice and Bob are determined on the fly, the laboratories are not required to share a common reference frame.

In the next sections, we present how to find the Schmidt bases (up to a global phase) from the experimental results gathered on individual qubits. We split the discussion into the two cases of nonvanishing and vanishing Bloch vectors, i.e., local averages (T_{x0}, T_{y0}, T_{z0}) , describing the states of the individual qubits.

This systematic procedure to verify entanglement in a pure two-qubit state is represented in Fig. 1, with the sections describing the particular steps.

A. From nonvanishing Bloch vectors to Schmidt bases

Consider first the case of nonzero Bloch vectors. The Schmidt bases of Alice and Bob are related to the standard bases as follows:

$$\begin{aligned} |a\rangle &= \cos \xi_A |0\rangle + e^{i\varphi_A} \sin \xi_A |1\rangle, \\ |a_\perp\rangle &= \sin \xi_A |0\rangle - e^{i\varphi_A} \cos \xi_A |1\rangle, \\ |b\rangle &= \cos \xi_B |0\rangle + e^{i\varphi_B} \sin \xi_B |1\rangle, \\ |b_\perp\rangle &= e^{i\delta} (\sin \xi_B |0\rangle - e^{i\varphi_B} \cos \xi_B |1\rangle). \end{aligned} \quad (4)$$

The global phase of $|b_\perp\rangle$ is relevant and required for the characterization of an arbitrary pure state, as can be seen from parameter counting. An arbitrary pure two-qubit state is parametrized by six real numbers (four complex amplitudes minus normalization condition and an irrelevant global phase). By plugging Eqs. (4) into the Schmidt decomposition (3), we indeed find the six relevant real parameters.

Any two-qubit state written in the standard bases of Alice and Bob can be brought into the Schmidt basis of Alice by the transformation

$$\begin{aligned} U(\xi_A, \varphi_A) &= |0\rangle\langle a| + |1\rangle\langle a_\perp| \\ &= \cos \xi_A |0\rangle\langle 0| + e^{-i\varphi_A} \sin \xi_A |0\rangle\langle 1| \\ &\quad + \sin \xi_A |1\rangle\langle 0| - e^{-i\varphi_A} \cos \xi_A |1\rangle\langle 1|. \end{aligned} \quad (5)$$

The coefficients ξ_A and φ_A of this transformation can be read from a nonvanishing normalized Bloch vector,

$$\vec{\alpha} \equiv \frac{\vec{T}^A}{|\vec{T}^A|} = (\sin 2\xi_A \cos \varphi_A, \sin 2\xi_A \sin \varphi_A, \cos 2\xi_A). \quad (6)$$

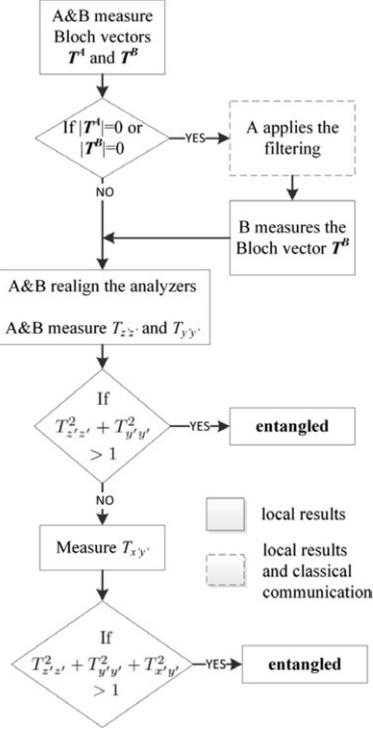


FIG. 1. The systematic way to experimentally verify entanglement of an arbitrary pure two-qubit state without any *a priori* knowledge and in the absence of a common reference frame. The steps of this diagram are described in detail in the main text.

Finally, the coefficients of the Schmidt basis in the standard basis are functions of components of vector $\vec{\alpha} = (T_{x0}, T_{y0}, T_{z0}) / \sqrt{T_{x0}^2 + T_{y0}^2 + T_{z0}^2}$ built out of experimentally accessible, local expectation values of Pauli measurements:

$$\begin{aligned} \cos \xi_A &= \sqrt{\frac{1 + \alpha_z}{2}}, & \sin \xi_A &= \sqrt{\frac{1 - \alpha_z}{2}}, \\ \cos \varphi_A &= \frac{\alpha_x}{\sqrt{1 - \alpha_z^2}}, & \sin \varphi_A &= \frac{\alpha_y}{\sqrt{1 - \alpha_z^2}}. \end{aligned} \quad (7)$$

If $T_{z0} = \pm 1$, the standard basis is the Schmidt basis. Note that instead of transforming the state, we can as well transform the measurement operators $\sigma_{n'} = U^\dagger \sigma_n U$. The new operators are given by

$$\begin{aligned} \sigma_{x'} &= \frac{-\alpha_x \alpha_z \sigma_x - \alpha_y \alpha_z \sigma_y + (1 - \alpha_z^2) \sigma_z}{\sqrt{1 - \alpha_z^2}}, \\ \sigma_{y'} &= \frac{\alpha_y \sigma_x + \alpha_x \sigma_y}{\sqrt{1 - \alpha_z^2}}, & \sigma_{z'} &= \alpha_x \sigma_x + \alpha_y \sigma_y + \alpha_z \sigma_z, \end{aligned} \quad (8)$$

and the Schmidt basis is the z' basis, i.e., $\sigma_{z'} = |a\rangle\langle a| - |a_\perp\rangle\langle a_\perp|$.

The equivalent analysis has to be done for the Schmidt basis of Bob. In summary, the Schmidt bases are defined by the direction of the Bloch vectors of the reduced states, up to a

3. Entanglement detection by correlation measurements

global phase. The global phase of $|b_{\perp}\rangle$ shows up as a relative phase in the Schmidt decomposition. As it is not obtainable by local measurements, it influences entanglement detection using operators (8).

B. Entanglement detection

Let us denote the basis established by local measurements of Bob by $\{|\tilde{b}\rangle, |\tilde{b}_{\perp}\rangle\}$, i.e., $|b\rangle = |\tilde{b}\rangle$ and $|b_{\perp}\rangle = e^{i\delta}|\tilde{b}_{\perp}\rangle$. Using the locally determined bases, the Schmidt decomposition takes the form

$$|\psi_S\rangle = \cos\theta|a\rangle|\tilde{b}\rangle + e^{i\delta}\sin\theta|a_{\perp}\rangle|\tilde{b}_{\perp}\rangle. \quad (9)$$

The correlations that Alice and Bob observe in the measurements related to locally determined bases are $T_{z'z'} = 1$ and $T_{x'x'} = \sin 2\theta \cos \delta$, $T_{y'y'} = -\sin 2\theta \cos \delta$, $T_{x'y'} = \sin 2\theta \sin \delta$, $T_{y'x'} = \sin 2\theta \sin \delta$. Note that the correlation $T_{y'y'}$ would vanish for $\cos \delta = 0$ and the two measurements $T_{z'z'}$ and $T_{y'y'}$ are no longer sufficient (they were sufficient if the full knowledge about the Schmidt bases had been available). In such a case, however, the other two correlations, $T_{x'y'}$ and $T_{y'x'}$, are nonzero and can be used to reveal entanglement. If the first two measurements are not sufficient to overcome the entanglement threshold of (2), then the third measurement of $T_{x'y'}$ correlations will definitely allow exceeding the threshold for every pure entangled state.

C. Vanishing Bloch vectors: Filtering

If the two-qubit state is maximally entangled, i.e., in the Schmidt decomposition $|\kappa\rangle = \frac{1}{\sqrt{2}}(|a\rangle|b\rangle + |a_{\perp}\rangle|b_{\perp}\rangle)$, then the Bloch vector is of zero length, $|T^A| = 0$, and the whole system admits infinitely many Schmidt decompositions. For every unitary operation of Alice, U , there exists an operation of Bob, U' , such that the state is unchanged,

$$U \otimes U'|\kappa\rangle = |\kappa\rangle. \quad (10)$$

Therefore, any basis of, say, Alice can serve as the Schmidt basis as soon as we accordingly update the basis of Bob. Our strategy to reveal the corresponding basis of Bob is to filter in the chosen Schmidt basis of Alice. It is best to explain it with an example. Assume Alice chooses the standard basis as her Schmidt basis. Due to the mentioned invariance of the maximally entangled state, there exists a Schmidt basis of Bob such that

$$|\kappa\rangle = \frac{1}{\sqrt{2}}(|0\rangle|b'\rangle + |1\rangle|b'_{\perp}\rangle). \quad (11)$$

The basis of Bob can be found by the filtering of Alice, $F = \varepsilon|0\rangle\langle 0| + |1\rangle\langle 1|$ with $\varepsilon \in [0, 1)$. We implemented this operation experimentally and provide the details later. In short, we use devices which are transparent to the $|1\rangle$ state, but probabilistically “reflect” the $|0\rangle$ state. If we imagine that a perfect detector is observing a part of this device where the reflected particle travels, and we see no detection, the filter operation is performed on the initial state. If Alice applies the filtering on her qubit and informs Bob that the filtering was successful, the initial state becomes

$$(F \otimes \mathbb{1})|\kappa\rangle \rightarrow \frac{1}{\sqrt{1+\varepsilon^2}}(\varepsilon|0\rangle|b'\rangle + |1\rangle|b'_{\perp}\rangle). \quad (12)$$

The result of filtering is that for Bob, a Bloch vector emerges and we can again use the method described above to find his Schmidt basis.

D. Performance

Summing up all required steps, we see that to experimentally verify entanglement of any pure two-qubit state without any further *a priori* knowledge requires at least 2×3 local measurements to determine the Schmidt bases and, sometimes, filtering requiring three local measurements more. Finally, two more (or three if $\delta = \pi/2$) correlation measurements allow one to verify the entanglement criterion.

IV. DECISION TREE

Our second algorithm for entanglement detection does not even require any initial measurements and directly applies also to mixed states. We will split the presentation into bipartite and multipartite cases. The decision tree provides an adaptive method to infer the next measurement setting from previous results.

A. Two qubits

Alice and Bob choose three orthogonal local directions x , y , and z independently from each other and agree to only measure correlations along these directions. In Fig. 2, we show exemplarily which correlations should be measured in order to detect entanglement in a small number of steps. Starting with a measurement of T_{zz} , one continues along the solid (dotted) arrow, if the correlation is higher (lower) than some threshold value t . We performed detailed numerical analysis on how the efficiency of entanglement detection depends on the threshold value. The efficiency is quantified by the percentage of entangled states detected at various steps of the decision

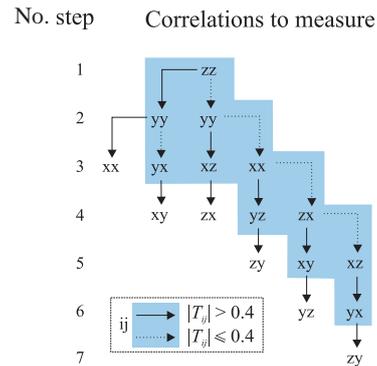


FIG. 2. (Color online) Decision strategies to detect entanglement. Start with a measurement of the correlation T_{zz} and proceed with the correlation along the solid (dotted) arrow if the measured correlation is higher (lower) than the chosen threshold value; here of $t = 0.4$. Due to correlation complementarity, there is a good chance of detecting entanglement in a small number of steps. The measurements in the blue shaded area suffice to detect all maximally entangled pure states with Schmidt-basis vectors x , y , or z .

tree. It turns out that the efficiency does not depend much on the threshold value and the best results are obtained for $t = 0.4$. We therefore set this threshold value in all of our simulations.

The construction of the tree is based on the principle of correlation complementarity [13,16–18]: in quantum mechanics, there exist trade-offs for the knowledge of dichotomic observables with corresponding anticommuting operators. For this reason, if the correlation $|T_{zz}|$ is big, then correlations $|T_{zx}|, |T_{zy}|, |T_{xz}|$, and $|T_{yz}|$ have to be small because their corresponding operators anticommute with the operator $\sigma_z \otimes \sigma_z$. Therefore, the next significant correlations have to lie in the xy plane of the correlation tensor and the next step in the tree is to measure the T_{yy} correlation.

In cases in which going through the whole tree did not reveal entanglement, we augmented it with additional measurements of correlations that were not established until that moment. The order of the additional measurements also results from the correlation complementarity. With every remaining measurement, we associate the “priority” parameter

$$P_{ij} = \sum_{k \neq i} P_{ij}(T_{kj}) + \sum_{l \neq j} P_{ij}(T_{il}), \quad (13)$$

which depends on the measurements of the decision tree in the following way:

$$P_{ij}(T_{mn}) = \begin{cases} T_{mn}^2 & \text{if } T_{mn} \text{ was performed before,} \\ 0 & \text{otherwise.} \end{cases} \quad (14)$$

According to the correlation complementarity, if the value of the corresponding parameter is small, there is a bigger chance that this correlation is significant. Therefore, the correlations T_{ij} with lower values of P_{ij} are measured first.

Let us illustrate this with the following example. The measured correlations of the decision tree are as follows: $T_{zz} = 0.7$, $T_{yy} = 0.6$, and $T_{xx} = 0.1$. Therefore, $P_{xy} = P_{yx} = T_{xx}^2 + T_{yy}^2 = 0.37$, $P_{xz} = P_{zx} = T_{xx}^2 + T_{zz}^2 = 0.5$, and $P_{yz} = P_{zy} = T_{zz}^2 + T_{yy}^2 = 0.85$. Accordingly, the order of the remaining measurements is as follows: first measure xy , then yx , and next xz, zx, yz , and zy .

B. Many qubits

Correlation complementarity, which holds also in the multipartite case, states that for a set $\{\alpha_1, \dots, \alpha_k\}$ of dichotomic mutually anticommuting multiparty operators, the following trade-off relation is satisfied by all physical states:

$$T_{\alpha_1}^2 + \dots + T_{\alpha_k}^2 \leq 1, \quad (15)$$

where T_{α_i} is the expectation value of observable α_i , and so on. Therefore, if one of the expectation values is maximal, say $T_{\alpha_1} = \pm 1$, then the other anticommuting observables have vanishing expectation values and do not have to be measured. In this way, we exclude exponentially many, in number of qubits, potential measurements because that many operators anticommute with α_1 , and we apply correlation complementarity pairwise to α_1 and one of the anticommuting operators. This motivates taking only sets of commuting operators along the branches of the decision tree that should be followed if the measured correlations are big.

We are thus led to propose the following algorithm generating one branch of the decision tree in which the first measurement, called $X \otimes X \otimes \dots \otimes X$, is assumed to have a big expectation value.

(i) Generate all N -partite Pauli operators that commute with $X \otimes X \otimes \dots \otimes X$.

Such operators have an even number of local Pauli operators different than X . Accordingly, their number is given by $\sum_{j=1}^{\lfloor \frac{N}{2} \rfloor} 2^{2j} \binom{N}{2j} = \frac{1}{2}(3^N - 1) - \text{Odd}(N)$, where $\text{Odd}(N) = 1$ if N is odd, and 0 otherwise. For example, in the three-qubit case, the set of operators commuting with XXX consists of 12 operators: $XZZ, ZZX, ZXZ, XYY, YYY, YXY, XYZ, XZY, YXZ, YZX, ZXY$, and ZYX .

(ii) Group them in strings of mutually commuting operators that contain as many elements as possible.

We verified for N up to eight qubits (and conjecture, in general) that the length of the string of mutually commuting operators is $L = 2^{N-1} + \text{Even}(N)$, where $\text{Even}(N) = 1$ if N is even, and 0 otherwise. In our three-qubit example, we have the following strings: $\{XXX, YXZ, XZZ, YZX\}$, $\{XXX, YYX, XYZ, YXZ\}$, $\{XXX, XZY, YZX, YXY\}$, $\{XXX, XZZ, ZXZ, ZZX\}$, $\{XXX, XYZ, ZXZ, ZYX\}$, $\{XXX, YXY, YYX, XYY\}$, $\{XXX, XYY, ZXY, ZYX\}$, and $\{XXX, ZZX, XZY, ZXY\}$. We denote the number of such strings by M .

(iii) Arrange the operators within the strings and sort the strings such that they are ordered with the same operator in the first position, then, if possible, second, third, etc.

In this way, we produce a set of strings $\{S_1, S_2, \dots, S_M\}$ such that in the first position of every string, we have $S_{j,1} = X \otimes X \otimes \dots \otimes X$; in the second position, the number of different operators is smaller or equal to the number of different operators in the third position, etc. After applying that operation in the three-qubit example, we obtain $\{XXX, XZZ, ZXZ, ZZX\}$, $\{XXX, XZZ, YXZ, YZX\}$, $\{XXX, XZY, YZX, YXY\}$, $\{XXX, XZY, ZZX, ZXY\}$, $\{XXX, XYZ, YXZ, YYX\}$, $\{XXX, XYZ, ZXZ, ZYX\}$, $\{XXX, XYY, YXY, YXX\}$, and $\{XXX, XYY, ZXY, ZYX\}$.

(iv) Connect the operators of the string S_1 with continuous arrows,

$$S_{1,1} \longrightarrow S_{1,2} \longrightarrow \dots \longrightarrow S_{1,L}. \quad (16)$$

(v) For all other strings S_j , with $j = 2, \dots, M$, check on which position string S_j differs from S_{j-1} . Let us denote this position by d . At these positions, strings are connected with another type of arrow, yielding the tree as

$$S_{j-1,d} \dashrightarrow S_{j,d} \longrightarrow S_{j,d+1} \longrightarrow \dots \longrightarrow S_{j,L}. \quad (17)$$

By using the strings of operators from step (iii) and choosing some threshold as to whether to follow one or the other string, we can now build up a decision tree, as shown in Fig. 3. Its essential feature is that an operator with big expectation value is followed only by the measurements of commuting operators, irrespectively of their expectation values.

3. Entanglement detection by correlation measurements

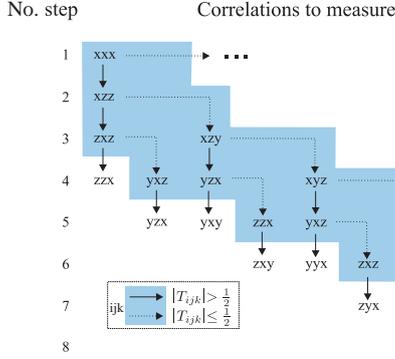


FIG. 3. (Color online) One branch of the decision tree for three qubits, starting with a measurement of the correlation T_{xxx} assumed to be big. The best efficiency is obtained for the threshold $t = 0.5$.

C. Bloch correlations

Finally, it would be useful to establish a measurement suitable as a starting point of the decision tree, i.e., such that the measured correlations have a good chance of being big. A natural candidate is to connect both methods discussed here and check whether the correlation measured along the Bloch vectors of every observer (we denote it as Bloch correlations) gives values close to the maximal correlation in a pure state. We verified this numerically and found that the Bloch correlations are larger than $\frac{3}{4}$ of maximal correlations in a pure state in 100% of two-qubit states and 69% of three-qubit states, but only in 27% of four-qubit states and 3% of five-qubit states. Therefore, the Bloch correlations give a very good starting point of the decision tree only for two and three qubits. We leave it as an open question whether a simple and reliable method exists that identifies the maximal correlations of a pure multiqubit state.

D. Performance

Let us analyze the results on the entanglement detection efficiency for different classes of two-qubit states. As explained in Sec. III B, at maximum, three correlation measurements are sufficient to detect entanglement once the local Schmidt bases of Alice and Bob are known. Here, in contrast, we study how many correlation measurements are needed when the decision tree is applied to an unknown entangled state.

The dependence of the efficiency of the algorithm on the number of steps involved can be seen in Fig. 4. The efficiency is defined by the fraction of detected entangled states with respect to all randomly generated entangled states. For nine steps, the algorithm detects all pure entangled states. This is expected because Eq. (2) is a necessary and sufficient condition for entanglement in the case of pure states. In the case of mixed states, Fig. 4 shows how the efficiency of the algorithm scales with the purity of the tested state. Since condition (2) is similar to the purity of a state, obviously, the scheme succeeds the faster a state becomes more pure.

Figure 4 also shows that the efficiency of the decision tree grows with the amount of entanglement in a state as

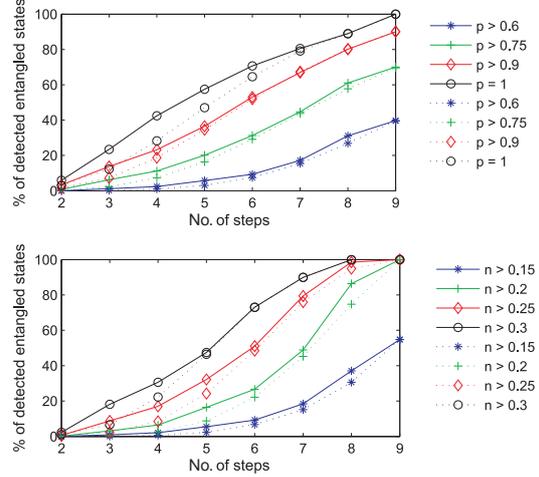


FIG. 4. (Color online) Efficiency of the decision tree for two-qubit random mixed states. The states were uniformly sampled according to the Haar measure. The efficiency increases with the purity of the state (top panel) as well as with the amount of entanglement in a tested state (bottom panel). Note that all pure entangled states are detected after nine steps, as well as all the states with negativity of more than 0.2 independently of their purity. Solid lines show the results when using the decision tree; dotted lines show the results when using random choices for the measurements.

characterized by the negativity [19]. It turns out that the tree detects all the states that have negativity of more than $\frac{1}{5}$.

We also compared the efficiency of the decision tree algorithm to entanglement detection based on a random order of measurements. In the first step of this protocol, Alice and Bob randomly choose one of nine measurements that also enter the decision tree. In the second step, they randomly measure one of the eight remaining measurements, and so on. At each step, condition (2) is checked for entanglement detection. Of course, the two methods converge for higher number of measurements. For a small number of measurements, the decision tree detects entanglement roughly one step faster than the random-measurement method. The advantage of the decision tree with respect to a random choice of the correlations is more pronounced for a higher number of qubits (see Sec. IV D3).

Condition (2) alone, i.e., without considering specific, state-dependent metrics (see [11]), cannot detect all mixed entangled states. As an illustration of how the decision tree works for mixed states, we first consider Werner states. It turns out that not all entangled states of this family can be detected, whereas the following example shows a family of mixed states for which all of the states are detected.

1. Werner states

Consider the family of states

$$\rho = p|\psi^-\rangle\langle\psi^-| + (1-p)\frac{1}{4}\mathbb{1}, \quad (18)$$

where $|\psi^-\rangle = \frac{1}{\sqrt{2}}(|01\rangle - |10\rangle)$ is the Bell singlet state, $\frac{1}{4}\mathbb{1}$ describes the completely mixed state (white noise), and p is

a probability [20]. Its correlation tensor, written in the same coordinate system for Alice and Bob, is diagonal with entries $T_{xx} = T_{yy} = T_{zz} = -p$, arising from the contribution of the entangled state. The states (18) are entangled if and only if $p > \frac{1}{3}$, whereas the decision tree reveals the entanglement only for $p > \frac{1}{\sqrt{3}} \approx 0.577$.

2. Entanglement mixed with colored noise

An exemplary class of density operators for which the decision tree detects all entangled states is provided by

$$\gamma = p|\psi^-\rangle\langle\psi^-| + (1-p)|01\rangle\langle 01|, \quad (19)$$

i.e., the maximally entangled state is mixed with colored noise $|01\rangle$ bringing anticorrelations along the local z axes. For this case, quite common for type-II parametric down-conversion sources, we obtain the following nonvanishing elements of its correlation tensor: $T_{xx} = T_{yy} = -p$ and $T_{zz} = -1$. Therefore, the decision tree allows detection of entanglement in this class of states in two steps. Note that the state is entangled already for an infinitesimal admixture of the Bell singlet state. We also verified numerically that for a hundred random choices of local coordinate systems, the decision tree detects entanglement even for $p > 10^{-3}$.

3. Three and more qubits

Similarly to the two-qubit case, we also studied the efficiency of the three-qubit decision tree of Fig. 3, as well as similar trees for higher number of qubits. The results for three qubits are presented in Fig. 5 and reveal that the decision tree is roughly two steps ahead of the protocol with a random order of measurements for small number of steps. In general, the number of steps the decision tree is ahead of the protocol with a random order of measurements grows exponentially with the number of qubits (see Fig. 6). The intuition behind this is that once big correlations are measured using the decision tree, a set of measurements exponential in size is excluded, whereas these measurements would still be randomly sampled in the other protocol.

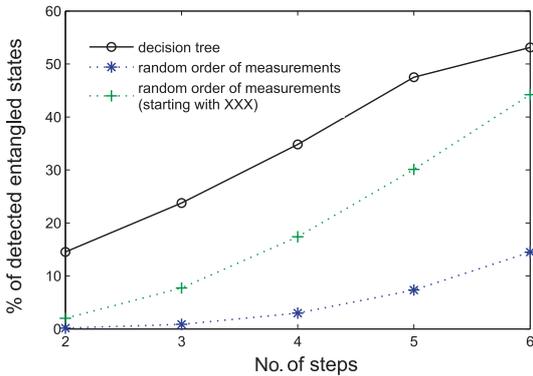


FIG. 5. (Color online) Efficiency of one branch of the decision tree for three-qubit random pure states. The states were uniformly sampled according to the Haar measure.

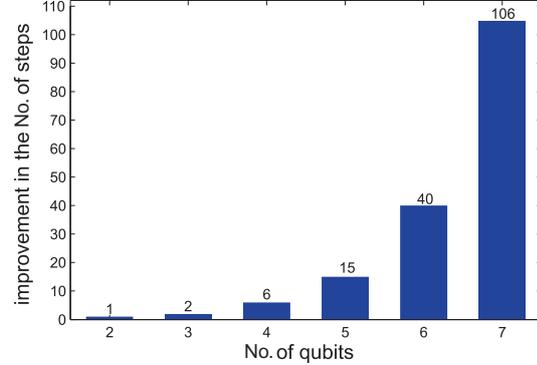


FIG. 6. (Color online) Efficiency of one branch of the decision tree for many qubits compared with a random choice of measurements. The plot shows the gain in the number of measurements provided by the decision tree. Pure states were uniformly sampled according to the Haar measure and the percentage of detected entangled states was calculated for different number of steps (measurements) in the tree as well as for the random order of measurements that start with $X \otimes \dots \otimes X$ for a fair comparison. We then compare the number of measurements for which the percentage of detected entangled states using the decision tree is the same as using the randomized measurements and plot here the maximal difference between them. The improvement provided by the tree grows exponentially with the number of qubits.

V. EXPERIMENTS

The entanglement detection schemes introduced above are experimentally evaluated by analyzing a variety of multiqubit entangled states. These states were created by spontaneous parametric down conversion (SPDC). Here, for the preparation of two-qubit entangled states, a type-I source with two crossed optically contacted β -barium-borate (BBO) crystals of 1 mm thickness is used; see Fig. 7 [21]. The computational basis $|0\rangle$ and $|1\rangle$ as introduced before is encoded in the polarization state $|H\rangle$ and $|V\rangle$, respectively. A continuous-wave laser diode (Nichia Corporation) at 402 nm is used to pump the BBO crystals with approximately 60 mW. The polarization of the

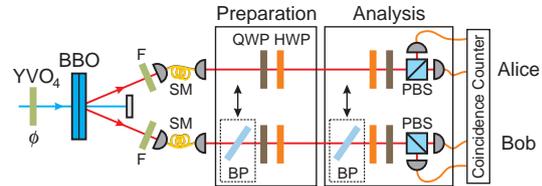


FIG. 7. (Color online) Scheme of the experimental type-I SPDC source used to prepare the state $\frac{1}{\sqrt{2}}(|HH\rangle + e^{i\phi}|VV\rangle)$. The phase ϕ can be set by an yttrium-vanadate crystal (YVO₄). Spectral filtering is performed by means of interference filters (F) and spatial filtering by single mode fibers (SM). Half- (HWP) and quarter-wave plates (QWP) are used for state preparation and analysis. Brewster plates (BP) enable the performance of the filter operation and the preparation of asymmetric states.

3. Entanglement detection by correlation measurements

pump light is oriented at 45° , allowing one to equally pump both crystals and to emit HH and VV polarized photon pairs with the same probability. However, a delay longer than the pump photon coherence length is acquired between the photon pairs generated in the first or second crystal over the length of the crystals, reducing their temporal indistinguishability. Therefore, an yttrium-vanadate (YVO_4) crystal of $200 \mu\text{m}$ thickness is introduced in front of the BBOs to precompensate for the delay and to set the phase ϕ between HH and VV . Using this configuration, entangled states of the form $|\Psi\rangle = \frac{1}{\sqrt{2}}(|H\rangle|H\rangle + e^{i\phi}|V\rangle|V\rangle)$ are generated [22].

In order to reduce the spectral bandwidth of the photon pairs, interference filters centered at 805 nm with a bandwidth of 7 nm are used. Spatial filtering is accomplished by coupling the photons at corresponding points of their emission cones into a pair of single-mode fibers. Polarization controllers allow for the compensation of the polarization rotation of the fibers. Then, the photons are transmitted through a set of quarter- (QWP) and half-wave plates (HWP), allowing an arbitrary transformation of the polarization state in each path. A set of Brewster plates with a loss rate up to $\approx 60\%$ for V and high transmission for H polarized light can be introduced in front of the wave plates to enable preparation of states. For the analysis, both Alice and Bob are provided with HWP and QWP as well as a filter (another Brewster plate) for Bob. Photons are then projected onto $|H\rangle$ and $|V\rangle$ implemented by a polarizing beam splitter (PBS) and respective detectors. Note that local filtering can also be accomplished by a polarizer. The output modes of the analyzing PBS are coupled into multimode fibers connected to avalanche photon detectors (SPCM-AQ4C, Perkin-Elmer module) with a photon detection efficiency of $\approx 50\%$. A coincidence logic is applied to extract the respective coincidence count rates within a time accuracy of $< 10 \text{ ns}$. The observed coincidence rate is approximately 200 s^{-1} and a measurement time of 10 s per basis setting allows one to register about 2000 events.

A. Schmidt decomposition

In order to perform the measurement in the Schmidt basis, we first have to determine basis vectors from the Bloch vectors observed by Alice and Bob. Let us consider the state depicted in Fig. 8(a). The table to the left shows the correlation tensor elements T_{ij} with the Bloch vectors of Alice (Bob) in the leftmost column (top row). For the application of the Schmidt decomposition method, Alice and Bob measure first their respective Bloch vectors (measurements actually to be performed are indicated by the blue shaded fields). Since they are close to 0, $\vec{T}_A = (0.002, 0.043, 0.017)$ and $\vec{T}_B = (0.109, -0.029, 0.029)$, the next step of the algorithm is to apply local filtering, as described by the scheme of Fig. 1. The filtered state shown in Fig. 8(b) has nonvanishing Bloch vectors, $\vec{T}_A = (0.338, -0.186, -0.136)$ and $\vec{T}_B = (-0.074, 0.147, 0.299)$, which can be used to find the corresponding Schmidt basis of the shared two-qubit state.¹

¹It should be noted that the Bloch vectors are determined from coincidence measurements. This is due to the low detection efficiency of correlated photons.

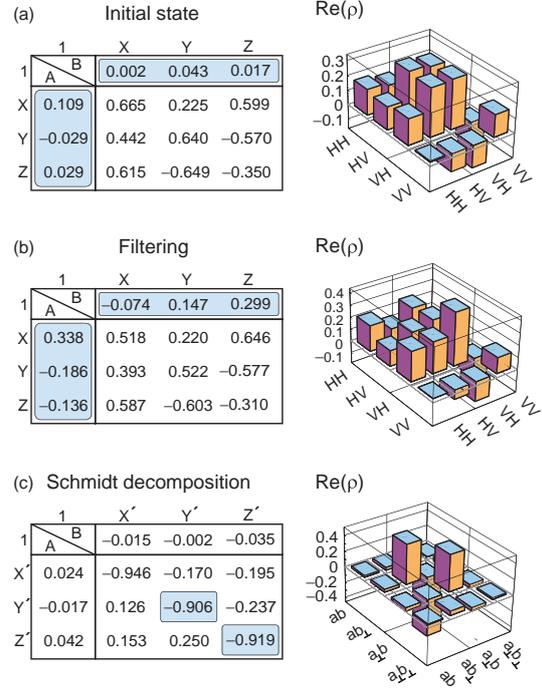


FIG. 8. (Color online) Schmidt decomposition of a maximally entangled unknown state. The correlation tensor and the density matrix are determined (a) before and (b) after applying local filtering. (c) After removing the filter, the state can be measured in its Schmidt basis.

If the phase ϕ is not determined by an additional correlation measurement, there are infinitely many such bases. As shown in Sec. III A, one possible choice is to redefine the local basis of Alice and Bob according to Eq. (8). Measuring along $\sigma_{i'}$ corresponds to a projection on its eigenstates $|\downarrow\rangle_{i'}$ and $|\uparrow\rangle_{i'}$. The task now is to find the angles for the wave plates of Alice and Bob, θ'_A/ϕ'_A and θ'_B/ϕ'_B , respectively. Since the PBS of the polarization analysis shown in Fig. 9 always projects on $|H\rangle$ and $|V\rangle$, the angles are calculated under the condition that $|\downarrow\rangle_{i'}$ ($|\uparrow\rangle_{i'}$) is rotated, up to a global phase τ , to $|H\rangle$ ($|V\rangle$), e.g.,

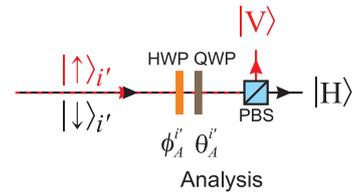


FIG. 9. (Color online) If Alice wants to measure in the basis $\sigma_{i'} = |\downarrow\rangle_{i'}\langle\downarrow|_{i'} - |\uparrow\rangle_{i'}\langle\uparrow|_{i'}$ ($i = x, y, z$), then the HWP and the QWP of the polarization analysis have to be aligned such that $|\downarrow\rangle_{i'}$ and $|\uparrow\rangle_{i'}$ are detected at different outputs of the PBS. The same holds for Bob.

TABLE I. Wave-plate settings for Alice and Bob to measure the maximally entangled state shown in Fig. 8(a) in the Schmidt basis and the complementary directions.

	Alice		Bob		
	$\frac{\lambda}{2}$	$\frac{\lambda}{4}$	$\frac{\lambda}{2}$	$\frac{\lambda}{4}$	
$\sigma_{x'}$	22.6°	25.8°	$\sigma_{x'}$	6.6°	4.6°
$\sigma_{y'}$	-15.8°	13.3°	$\sigma_{y'}$	7.2°	-30.6°
$\sigma_{z'}$	-9.9°	-12.8°	$\sigma_{z'}$	34.7°	13.5°

for Alice

$$U_{\text{QWP}}(\theta_A^i)U_{\text{HWP}}(\phi_A^i)|\downarrow\rangle_{i'} = e^{i\tau_1}|H\rangle, \quad (20)$$

$$U_{\text{QWP}}(\theta_A^i)U_{\text{HWP}}(\phi_A^i)|\uparrow\rangle_{i'} = e^{i\tau_2}|V\rangle, \quad (21)$$

where U labels the unitary operation of the corresponding wave plate. The angles θ_A^i/ϕ_A^i and θ_B^i/ϕ_B^i can be found by (numerically) solving the equation

$$|\langle H|U_{\text{QWP}}(\theta_A^i)U_{\text{HWP}}(\phi_A^i)|\downarrow\rangle_{i'}|^2 = 1, \quad (22)$$

and similarly for Bob. Using this scheme, we find the angles for Alice's and Bob's wave plates, such that their qubits are measured in the primed bases, presented in Table I.

After removing the filter, Alice and Bob can now measure in their new bases and reveal entanglement by measuring $T_{z'z'}$, followed by $T_{y'y'}$, and possibly $T_{y'x'}$. Here, the two measurements suffice to reveal entanglement as $T_{z'z'}^2 + T_{y'y'}^2 = 1.665 \pm 0.05 > 1$.

In full analogy to the previous example, it is also possible to apply the Schmidt decomposition scheme to a nonmaximally entangled state, e.g., as presented in Fig. 10(a). For using the Schmidt decomposition strategy, first both parties agree on measuring their respective Bloch vectors, $\vec{T}_A =$

TABLE II. Wave-plate settings for Alice and Bob to measure the asymmetric state shown in Fig. 10(a) in the Schmidt basis and the complementary directions.

	Alice		Bob		
	$\frac{\lambda}{2}$	$\frac{\lambda}{4}$	$\frac{\lambda}{2}$	$\frac{\lambda}{4}$	
$\sigma_{x'}$	0.7°	0.8°	$\sigma_{x'}$	21.9°	9.4°
$\sigma_{y'}$	-13.5°	17.9°	$\sigma_{y'}$	40.0°	-55.0°
$\sigma_{z'}$	-35.2°	-27.0°	$\sigma_{z'}$	42.0°	3.3°

$(0.072, -0.026, -0.213)$ and $\vec{T}_B = (-0.201, 0.279, 0.012)$. As they already can be distinguished from noise, Alice and Bob can find the Schmidt bases without applying the filter operation. The angle settings of the wave plates for analyzing in the Schmidt bases are again calculated using (22) and are shown in Table II. Again, the state is proved to be entangled after only two correlation measurements since $T_{z'z'}^2 + T_{y'y'}^2 = 1.624 \pm 0.047 > 1$; see Fig 10(b).

B. Decision tree

1. Two qubits

Let us first consider the two states analyzed above using Schmidt decomposition. For the first state (Fig. 8), we see that a direct application of the decision tree shown in Fig. 2 would require four correlation measurements to reveal entanglement, namely, $T_{zz}^2 + T_{yy}^2 + T_{xx}^2 + T_{zx}^2 = (-0.350)^2 + 0.640^2 + 0.599^2 + 0.615^2 = 1.33 \pm 0.03 > 1$. Similarly, the analysis of the second state (Fig. 10) would require four correlation measurements to determine entanglement, namely, $T_{zz}^2 + T_{yy}^2 + T_{xx}^2 + T_{zx}^2 = (-0.312)^2 + 0.582^2 + 0.579^2 + 0.622^2 = 1.158 \pm 0.030 > 1$. This shows that quite a few more correlation measurements are needed when using the decision tree. Yet, it saves measuring the Bloch vectors and filtering operations. To illustrate the entanglement detection scheme, we further apply it to a selection of maximally entangled states (Fig. 11) and to nonmaximally entangled states (Fig. 12).

For didactical reasons, the full correlation tensors are depicted, in both cases. In order to reveal entanglement, the decision tree requires the measurement of a number of correlations much smaller than needed to reconstruct the full density matrix. Following the lines as described in Sec. IV, only correlation measurements shaded red are required to detect entanglement. As an example, let us consider the state $\frac{1}{\sqrt{2}}(|RR\rangle + |LL\rangle)$ [Fig. 11(g)], for which a measurement of the two correlations $T_{zz} = 0.905$ and $T_{yy} = 0.977$ suffices to reveal entanglement since $T_{zz}^2 + T_{yy}^2 = 1.773 \pm 0.039 > 1$. In contrast, for the state $\frac{1}{\sqrt{2}}(|RP\rangle + i|LM\rangle)$ [Fig. 11(e)], the algorithm only stops after six steps, as the measurements of $T_{zz} = -0.089$, $T_{yy} = -0.091$, $T_{xx} = 0.099$, $T_{zx} = -0.194$, $T_{xz} = 0.941$, and $T_{yx} = 0.961$ are required to beat the threshold, i.e., $1.872 \pm 0.058 > 1$. A similar reasoning is applied to reveal entanglement of other two-qubit states.

The entanglement detection scheme is further applied to a selection of nonmaximally entangled states (Fig. 12). As an example, let us consider the state $0.83|LH\rangle + 0.56i|RV\rangle$

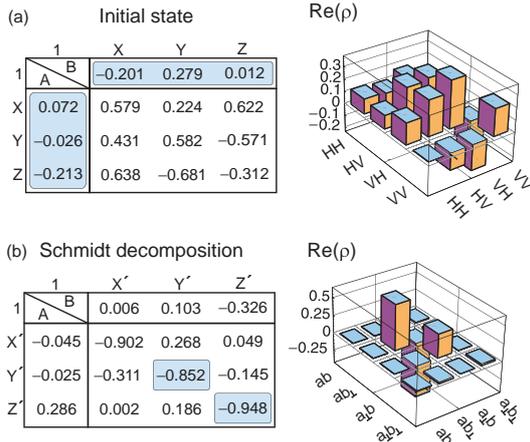


FIG. 10. (Color online) Schmidt decomposition of a nonmaximally entangled state. (a) The correlation tensor and density matrix are displayed for an unknown asymmetric state. (b) The state has nonzero Bloch vectors enabling one to determine the corresponding Schmidt basis for which the measured correlations are maximal.

3. Entanglement detection by correlation measurements

OPTIMIZED STATE-INDEPENDENT ENTANGLEMENT ...

PHYSICAL REVIEW A **88**, 022327 (2013)

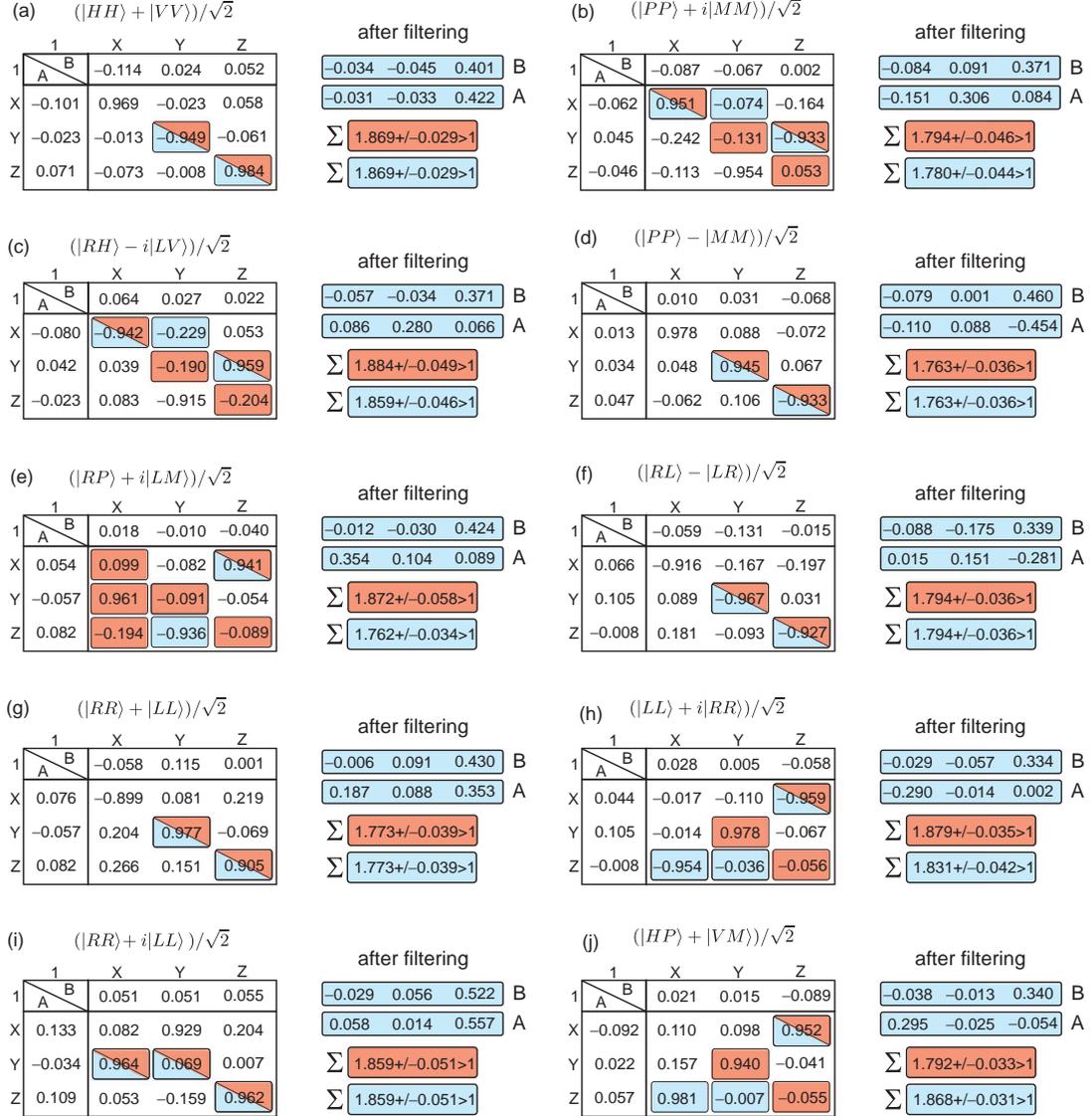


FIG. 11. (Color online) Application of the decision tree on a selection of maximally entangled states, allowing one to determine the entanglement of the state by measuring the correlations marked in red. As an alternative, local filtering is applied in order to extract the correlations with maximal value (blue correlations).

[Fig. 12(c)], for which our method reveals entanglement after four steps, as the measurements of $T_{zz} = 0.007$, $T_{yy} = 0.069$, $T_{xx} = -0.801$, and $T_{yz} = -0.968$ give a value of $1.583 \pm 0.067 > 1$. Similarly, as expected, for a separable state such as $|HH\rangle$ [Fig. 12(f)], our entanglement criterion delivers a value of $\sum_{k,l=1}^3 T_{kl}^2 = 0.964 \pm 0.062 < 1$ for measuring all correlations, not revealing entanglement clearly. These states, of course, can be analyzed also using Schmidt decomposition.

For maximally entangled states, the Bloch vectors after local filtering are also shown (blue color; see Fig. 11), while for nonmaximally entangled states (Fig. 12), no local filtering is required since the Bloch vectors are already nonvanishing. In all cases, only one entry of the respective Bloch vectors is large compared to the others. Therefore, no realignment of the analyzers is necessary. Due to Schmidt decomposition, the decision tree should start with a correlation measurement along

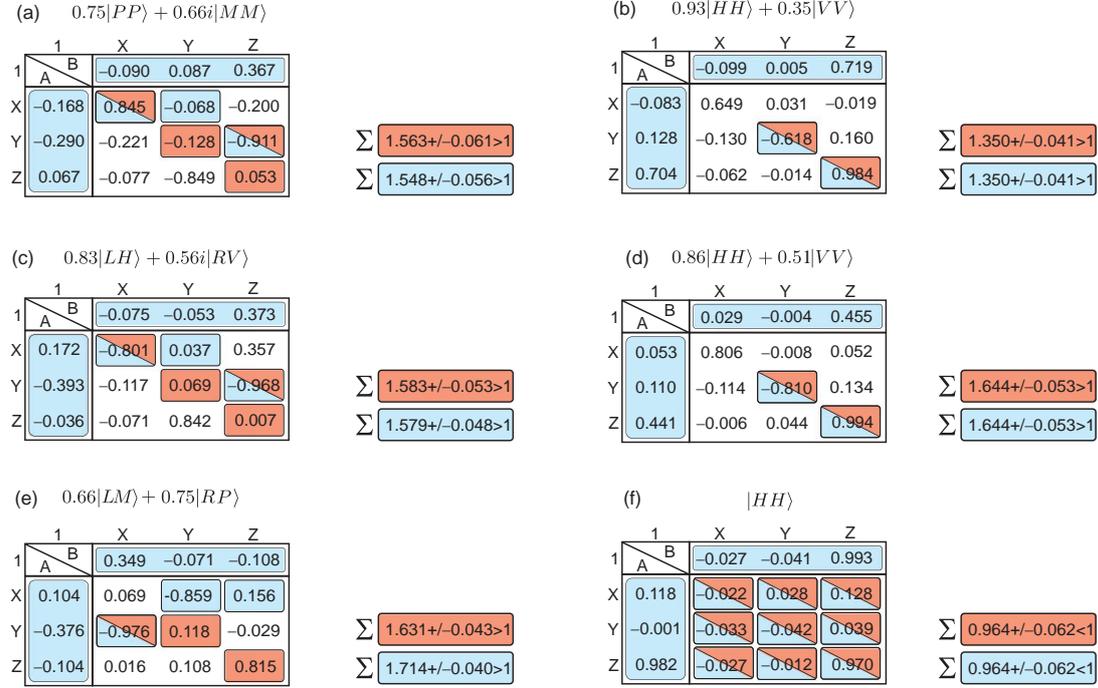


FIG. 12. (Color online) Application of the decision tree on a selection of nonmaximally entangled states, allowing one to determine the entanglement of the state by measuring the correlations marked in red. Due to the asymmetry of the states, local filtering is unnecessary, and the information on the Bloch vectors can be used to detect entanglement with a maximal number of three correlation measurements (blue correlations). (f) For a product state, the full set of correlations does not reveal entanglement, as it should be.

a direction in which we see a big local expectation value. In such a case, it is sufficient to cyclically relabel the required measurements as defined for the original decision tree. Following this method, it is possible to detect entanglement with a maximum number of three steps. The first correlation to be measured is determined by the Bloch vectors after applying local filtering.

2. Many qubits

For the demonstration of multiqubit entanglement detection, we use a family of three-photon polarization entangled Gdańsk (G) states [23] and the four-qubit Dicke state. The G states are defined by

$$|G(\alpha)\rangle = \cos(\alpha)|W\rangle + \sin(\alpha)|\bar{W}\rangle, \quad (23)$$

where $|W\rangle = \frac{1}{\sqrt{3}}(|HHV\rangle + |HVV\rangle + |VHH\rangle)$, and in order to obtain $|\bar{W}\rangle$, one exchanges H and V . The four-qubit Dicke state with two “excitations” reads

$$|D_4^{(2)}\rangle = \frac{1}{\sqrt{6}}(|HHVV\rangle + |HVVH\rangle + |VHHV\rangle + |HVVH\rangle + |VHVV\rangle + |VVHH\rangle). \quad (24)$$

Generalized three-qubit G state. In order to observe these states, a collinear type-II SPDC source together with a linear

setup to prepare the four-photon Dicke state $D_4^{(2)}$ is used [24,25]. The three-photon state is obtained if the first photon is measured to be $\cos(\alpha)|H\rangle + \sin(\alpha)|V\rangle$ polarized.

The protocol for entanglement detection starts with observers locally measuring the polarization of their respective photons, enabling them to individually determine the Bloch vectors.

(i) For the $G(\pi/4)$ state, we obtain $T_{i00} = (0.636, -0.008, -0.015)$, $T_{0j0} = (0.623, -0.092, 0.010)$, and $T_{00k} = (0.636, 0.070, 0.022)$. The Bloch vectors suggest that the correlation T_{xxx} is big. Therefore, the decision tree starts with the measurement of $T_{xxx} = 0.904 \pm 0.025$ and continues with $T_{xzz} = -0.578 \pm 0.025$ (see Fig. 3). These two measurements already prove entanglement because $T_{xxx}^2 + T_{xzz}^2 = 1.152 \pm 0.038 > 1$.

(ii) For the W state, $G(\pi/2)$, the Bloch vectors are $T_{i00} = (0.016, -0.070, 0.318)$, $T_{0j0} = (-0.010, -0.073, 0.308)$, and $T_{00k} = (-0.011, -0.0547, 0.319)$, which suggests that now the correlation T_{zzz} is big. Indeed, we observe $T_{zzz} = -0.882 \pm 0.025$. The decision tree is the same as above, but with local axes renamed as follows: $x \rightarrow z \rightarrow y \rightarrow x$. Therefore, the second measurement has to be T_{zyy} . With $T_{zyy} = 0.571 \pm 0.025$, we again prove entanglement as $T_{xxx}^2 + T_{zyy}^2 = 1.104 \pm 0.037 > 1$.

3. Entanglement detection by correlation measurements

OPTIMIZED STATE-INDEPENDENT ENTANGLEMENT ...

PHYSICAL REVIEW A **88**, 022327 (2013)

Four-qubit Dicke state. Here, we have vanishing Bloch vectors, $T_{i000} = (-0.020, -0.016, 0.007)$, $T_{0j00} = (-0.011, -0.029, 0.014)$, $T_{00k0} = (-0.018, -0.020, -0.004)$, and $T_{000l} = (-0.009, -0.022, 0.008)$. We construct a set of mutually commuting operators, which form the first branch of the four-qubit decision tree starting with $T_{zzzz}\{zzzz \rightarrow zzxx \rightarrow zxzx \rightarrow zxxz \rightarrow xzxx \rightarrow xxzz \rightarrow xzzx \rightarrow xxxx \rightarrow yyyy\}$. After measuring the correlations, $T_{zzzz} = 0.848 \pm 0.025$, $T_{zzxx} = -0.533 \pm 0.025$, $T_{zxzx} = -0.552 \pm 0.025$, our algorithm succeeds since $T_{zzzz}^2 + T_{zzxx}^2 + T_{zxzx}^2 = 1.3082 \pm 0.041 > 1$.

VI. CONCLUSIONS

The entanglement of arbitrary multiqubit states can be efficiently detected based on two methods described here. Both methods employ a criterion based on the sum of squared correlations. Combining this with an adaptive determination of the correlations to be measured allows one to succeed much faster than standard tomographic schemes. The first one, particularly designed for two-qubit states, determines the Schmidt decomposition from local measurements only,

where at most three correlation measurements are sufficient for entanglement detection. The second one employs a decision tree to speed up the analysis. Its design is based on correlation complementarity and prevents one from measuring less informative correlations. The performance of the scheme is numerically analyzed for arbitrary pure states and, in the two-qubit case, for mixed states. The schemes succeed, on average, at least one step earlier as compared with random sampling on two-qubit states, with an exponentially increasing speedup for a higher number of qubits. Our results encourage the application of these schemes in state-of-the-art experiments with quantum states of increasing complexity.

ACKNOWLEDGMENTS

We thank the EU-BMBF project QUASAR and the EU projects QWAD and QOLAPS for supporting this work. T.P. acknowledges support by the National Research Foundation, the Ministry of Education of Singapore, start-up grant of the Nanyang Technological University, and NCN Grant No. 2012/05/E/ST2/02352. C.S. thanks QCCC of the Elite Network of Bavaria for support.

-
- [1] R. Horodecki, P. Horodecki, M. Horodecki, and K. Horodecki, *Rev. Mod. Phys.* **81**, 865 (2009).
 - [2] J.-W. Pan, Z.-B. Chen, C.-Y. Lu, H. Weinfurter, A. Zeilinger, and M. Żukowski, *Rev. Mod. Phys.* **84**, 777 (2012).
 - [3] M. A. Nielsen and I. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, New York, 2000).
 - [4] M. Horodecki, P. Horodecki, and R. Horodecki, *Phys. Lett. A* **223**, 1 (1996).
 - [5] M. Bourennane, M. Eibl, C. Kurtsiefer, S. Gaertner, H. Weinfurter, O. Gühne, P. Hyllus, D. Bruß, M. Lewenstein, and A. Sanpera, *Phys. Rev. Lett.* **92**, 087902 (2004).
 - [6] B. M. Terhal, *Phys. Lett. A* **271**, 319 (2000).
 - [7] M. Lewenstein, B. Kraus, J. I. Cirac, and P. Horodecki, *Phys. Rev. A* **62**, 052310 (2000).
 - [8] D. Bruß, J. I. Cirac, P. Horodecki, F. Hulpke, B. Kraus, M. Lewenstein, and A. Sanpera, *J. Mod. Opt.* **49**, 1399 (2002).
 - [9] O. Gühne and P. Hyllus, *Int. J. Theor. Phys.* **42**, 1001 (2003).
 - [10] O. Gühne and G. Toth, *Phys. Rep.* **474**, 1 (2009).
 - [11] P. Badziąg, Č. Brukner, W. Laskowski, T. Paterek, and M. Żukowski, *Phys. Rev. Lett.* **100**, 140403 (2008).
 - [12] W. Laskowski, D. Richart, C. Schwemmer, T. Paterek, and H. Weinfurter, *Phys. Rev. Lett.* **108**, 240501 (2012).
 - [13] P. Kurzyński, T. Paterek, R. Ramanathan, W. Laskowski, and D. Kaszlikowski, *Phys. Rev. Lett.* **106**, 180402 (2011).
 - [14] E. Schmidt, *Math. Ann.* **63**, 433 (1906).
 - [15] A. Peres, *Quantum Theory: Concepts and Methods* (Kluwer Academic, Dordrecht, 1995).
 - [16] G. Toth and O. Gühne, *Phys. Rev. A* **72**, 022340 (2005).
 - [17] S. Wehner and A. Winter, *J. Math. Phys.* **49**, 062105 (2008).
 - [18] S. Wehner and A. Winter, *New J. Phys.* **12**, 025009 (2010).
 - [19] G. Vidal and R. F. Werner, *Phys. Rev. A* **65**, 032314 (2002).
 - [20] R. F. Werner, *Phys. Rev. A* **40**, 4277 (1989).
 - [21] P. G. Kwiat, E. Waks, A. G. White, I. Appelbaum, and P. H. Eberhard, *Phys. Rev. A* **60**, R773 (1999).
 - [22] P. Trojek and H. Weinfurter, *Appl. Phys. Lett.* **92**, 211103 (2008).
 - [23] A. Sen(De), U. Sen, and M. Żukowski, *Phys. Rev. A* **68**, 032309 (2003).
 - [24] R. Krischek, C. Schwemmer, W. Wieczorek, H. Weinfurter, P. Hyllus, L. Pezze, and A. Smerzi, *Phys. Rev. Lett.* **107**, 080504 (2011).
 - [25] N. Kiesel, C. Schmid, G. Tóth, E. Solano, and H. Weinfurter, *Phys. Rev. Lett.* **98**, 063604 (2007).

P3.3

PRL 114, 180501 (2015)

PHYSICAL REVIEW LETTERS

week ending
8 MAY 2015

Genuine Multipartite Entanglement without Multipartite Correlations

Christian Schwemmer,^{1,2} Lukas Knips,^{1,2} Minh Cong Tran,³ Anna de Rosier,⁴ Wiesław Laskowski,⁴
Tomasz Paterek,^{3,5,6,*} and Harald Weinfurter^{1,2}¹Max-Planck-Institut für Quantenoptik, Hans-Kopfermann-Straße 1, 85748 Garching, Germany²Department für Physik, Ludwig-Maximilians-Universität, 80797 München, Germany³School of Physical and Mathematical Sciences, Nanyang Technological University, Singapore, 637371 Singapore⁴Institute of Theoretical Physics and Astrophysics, University of Gdańsk, PL-80-952 Gdańsk, Poland⁵Center for Quantum Technologies, National University of Singapore, Singapore, 117543 Singapore⁶MajuLab, 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 multipartite 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 multipartite 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 σ_{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

3. Entanglement detection by correlation measurements

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

$$\rho_{\psi}^{nc} = \frac{1}{2}|\psi\rangle\langle\psi| + \frac{1}{2}|\bar{\psi}\rangle\langle\bar{\psi}|, \quad (2)$$

we obtain a state ρ_{ψ}^{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 α_{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 ρ and $\bar{\rho}$ therefore produces a state with no l -partite 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 ± 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 ρ_{ϕ}^{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 $L = \max_{T^{\text{bi-prod}}} (T, T^{\text{biproduct}})$ is the left-hand side of Eq. (6). In the ideal case of preparing ρ^{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 ρ_{ϕ}^{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 ρ_{ϕ}^{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 multiparty nonclassicality in Ref. [26]. Kaszlikowski *et al.* [9], however, showed that such a quantification is not sufficient as the state ρ_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 ρ_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 ρ_ϕ^{nc}). While this approach does assign classical correlations in the context of Giorgi *et al.* [15] to ρ_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 ρ_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 ± 0.005 , which due to the high number of detected events [16] is compatible with the value 0.917 ± 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 ± 0.011 for $|W\rangle^{\text{exp}}$, 0.919 ± 0.010 for $|\bar{W}\rangle^{\text{exp}}$, and 0.961 ± 0.003 for $\rho_W^{nc,\text{exp}}$. Analogously, starting with a six-photon Dicke state $|D_6^{(3)}\rangle$

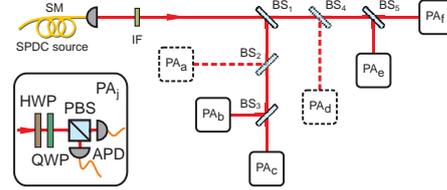


FIG. 1 (color online). Schematic of the linear optical setup used to observe symmetric Dicke states from which states with vanishing 3- and 5-partite correlations can be obtained. The photons are created by means of a cavity enhanced pulsed collinear type II spontaneous parametric down conversion source 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_2 and BS_4 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 ± 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 $T_{\theta_j j_3} = \text{Tr}(\rho\sigma_\theta \otimes \sigma_{j_2} \otimes \sigma_{j_3}) (T_{j_1 j_3})$. Indeed, as shown in

3. Entanglement detection by correlation measurements

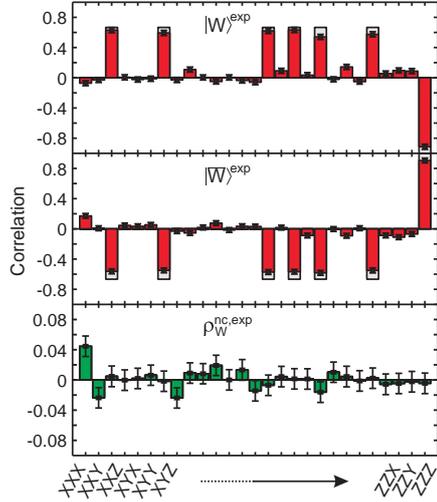


FIG. 2 (color online). Experimental tripartite correlations (red) for $|W\rangle^{\text{exp}}$, $|\bar{W}\rangle^{\text{exp}}$, and (green) $\rho_W^{\text{nc,exp}}$ in comparison to the theoretically expected values (gray). Note that the correlations of the state $\rho_W^{\text{nc,exp}}$ are magnified 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 θ (ϕ). 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 θ (ϕ). 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.

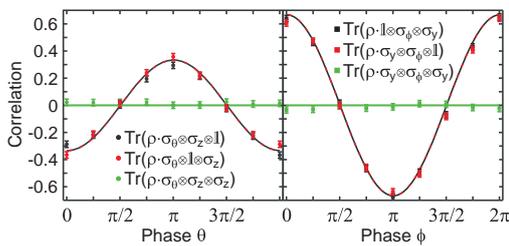


FIG. 3 (color online). Vanishing tripartite correlations for arbitrary measurements and non-vanishing bipartite correlations. Observable σ_θ (σ_ϕ) was measured on the first (second) qubit and σ_z (σ_y) measurements were performed on both other qubits (green curves) or one of them (red and black curves). 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 Research 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.

3. Entanglement detection by correlation measurements

Genuine n -partite entanglement without n -partite correlations SUPPLEMENTAL MATERIAL

Christian Schwemmer,^{1,2} Lukas Knips,^{1,2} Minh Cong Tran,³ Anna de Rosier,⁴
Wiesław Laskowski,⁴ Tomasz Paterek,^{3,5,6,*} and Harald Weinfurter^{1,2}

¹Max-Planck-Institut für Quantenoptik, Hans-Kopfermann-Straße 1, 85748 Garching, Germany

²Department für Physik, Ludwig-Maximilians-Universität, 80797 München, Germany

³School of Physical and Mathematical Sciences, Nanyang Technological University, 637371 Singapore

⁴Institute of Theoretical Physics and Astrophysics,
University of Gdańsk, PL-80-952 Gdańsk, Poland

⁵Centre for Quantum Technologies, National University of Singapore, 117543 Singapore

⁶MajuLab, 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 σ_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 σ_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 ρ_ϕ^{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 ρ_ϕ^{nc} contains only bipartite correlations:

$$\begin{aligned} T_{xx0} &= T_{yy0} = \sin(2\beta) \sin(\alpha), \\ T_{x0x} &= T_{y0y} = \sin(2\beta) \cos(\alpha), \\ T_{0xx} &= T_{0yy} = \sin^2(\beta) \sin(2\alpha), \\ T_{zx0} &= \cos(2\alpha) \sin^2(\beta) - \cos^2(\beta), \\ T_{z0z} &= -\cos(2\alpha) \sin^2(\beta) - \cos^2(\beta), \\ T_{0zz} &= \cos(2\beta), \end{aligned} \quad (9)$$

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:

$$a_x^2 + a_y^2 + a_z^2 = b_x^2 + b_y^2 + b_z^2 = \cos^2(2\theta), \quad (14)$$

3. Entanglement detection by correlation measurements

3

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 $T_{x0x}(P_{xx} + P_{yy}) + T_{z0z}P_{zz}$:

- (i) $|P_{xx}| = 1$ and $|P_{yy}| = |P_{zz}| = \sin(2\theta)$ with their signs matching those of T_{x0x} and T_{z0z} 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

$$\begin{aligned} L_{(i)} &= 1 + |T_{x0x}| + \sin(2\theta)(|T_{x0x}| + |T_{z0z}|) + \cos(2\theta)(\max(|T_{x0x}|, |T_{z0z}|) + \max(|T_{0xx}|, |T_{0zz}|)) \\ &\leq 1 + |T_{x0x}| + \sqrt{(|T_{x0x}| + |T_{z0z}|)^2 + (\max(|T_{x0x}|, |T_{z0z}|) + \max(|T_{0xx}|, |T_{0zz}|))^2}, \end{aligned} \quad (17)$$

where in the last step we optimized over θ . The same procedure applied to the other two cases gives:

$$L_{(ii)} \leq 1 + |T_{z0z}| + \sqrt{4T_{x0x}^2 + (\max(|T_{x0x}|, |T_{z0z}|) + \max(|T_{0xx}|, |T_{0zz}|))^2}, \quad (18)$$

$$L_{(iii)} \leq 1 + |T_{z0z}| + \max(|T_{x0x}|, |T_{z0z}|) + \max(|T_{0xx}|, |T_{0zz}|). \quad (19)$$

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 α or β close to 0 and $\pi/2$ that already achieve the bound of 4. Here, we give a simple analytical argument showing that ρ_ϕ^{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 ρ_ϕ^{nc} can only be bi-separable if there are bi-product pure states in its support. The support of ρ_ϕ^{nc} is spanned by $|\phi\rangle$ and $|\bar{\phi}\rangle$, i.e., ρ_ϕ^{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 ρ_ϕ^{nc} into pure states can only use pure states of the form

$$|\Phi\rangle = a|\phi\rangle + b|\bar{\phi}\rangle. \quad (20)$$

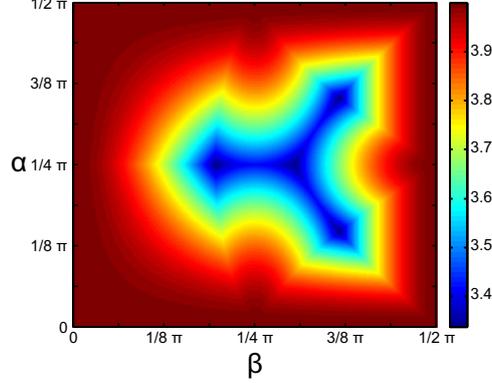


FIG. 1. Contour plot showing the maximal value of the left-hand side of our entanglement criterion for the states ρ_ϕ^{nc} defined above (7). Whenever the value is below 4, i.e., the right-hand side value as given in (10), the criterion detects genuine tripartite entanglement. This shows that all the states ρ_ϕ^{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 ρ_ϕ^{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 ρ_ϕ^{nc} . However, entanglement is detected only for a range of roughly $\alpha \in [0.590, 1.31]$ and $\beta \in [0.333, 1.24]$.

3. Entanglement detection by correlation measurements

5

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 ρ_ϕ^{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.

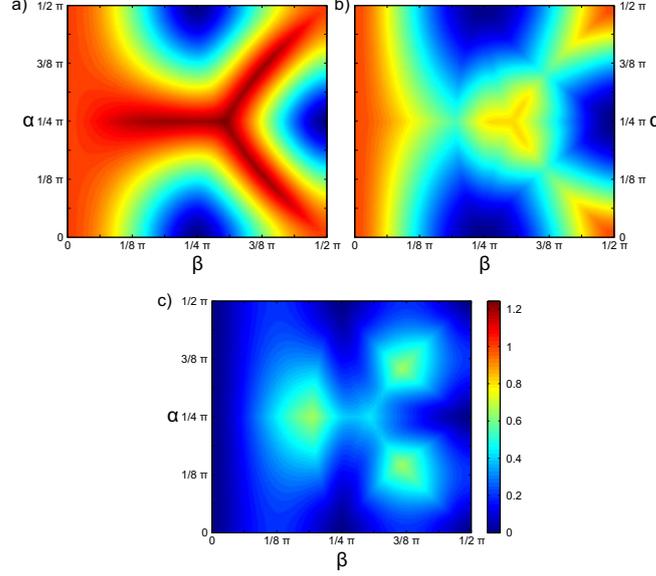


FIG. 2. Correlation content [15] of the states $\rho_p^{nc} = \frac{1}{2} |\phi\rangle\langle\phi| + \frac{1}{2} |\bar{\phi}\rangle\langle\bar{\phi}|$ with the pure states given in Eq. (7). (a) *Total genuine tripartite correlations*. The genuine tripartite correlations vanish only for mixtures of bi-product 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 ± 0.011 for $|W\rangle^{exp}$, 0.919 ± 0.010 for $|\bar{W}\rangle^{exp}$, and 0.961 ± 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 ± 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 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 $F_{|D_5^{(3)}\rangle}(|\Theta^{(3)}\rangle) = 0.979 \pm 0.012$). Equally mixing the states

3. Entanglement detection by correlation measurements

7

$|\Theta^{(2)}\rangle^{exp}$ 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.

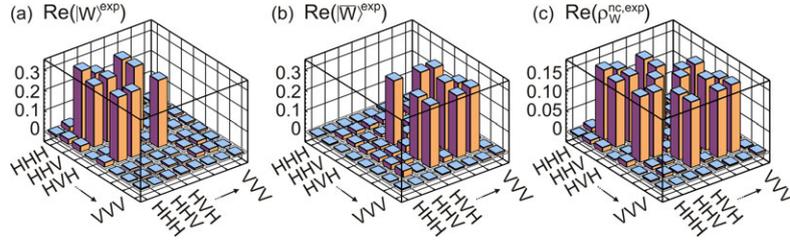


FIG. 3. Experimental three qubit states as obtained from the state $|D_4^{(2)}\rangle^{exp}$. (a) The state $|W\rangle^{exp}$ is obtained by projection of the fourth qubit of $|D_4^{(2)}\rangle^{exp}$ on V . (b) The state $|\bar{W}\rangle^{exp}$ is prepared by projecting the fourth qubit of $|D_4^{(2)}\rangle^{exp}$ 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 ± 0.011 for $|W\rangle^{exp}$, 0.919 ± 0.010 for $|\bar{W}\rangle^{exp}$, and 0.961 ± 0.003 for $\rho_W^{nc,exp}$.

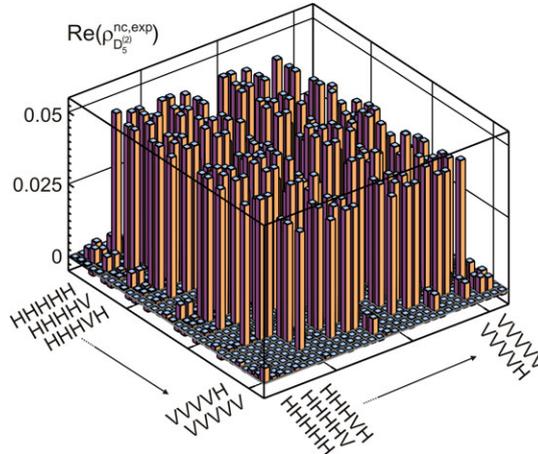


FIG. 4. Experimental state $\rho_{D_5^{(2)}}^{nc,exp}$ determined from five-fold coincidences together with permutational invariant tomography [18]. The fidelity with respect to the target state is 0.911 ± 0.004 .

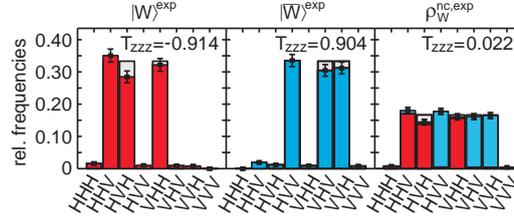


FIG. 5. (color online). Detection frequencies when observing the states $|W\rangle^{exp}$ (red) and $|\overline{W}\rangle^{exp}$ (blue) and $\rho_W^{nc,exp}$ (red and blue) in the $\sigma_z^{\otimes 3}$ basis. From these data T_{zzz} values can be calculated showing how the correlations of $|W\rangle^{exp}$ and $|\overline{W}\rangle^{exp}$ 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 $|\overline{W}\rangle^{exp}$. The state $|W\rangle^{exp}$ was observed with a slightly lower probability (0.485) than the state $|\overline{W}\rangle^{exp}$ (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 $|\overline{W}\rangle^{exp}$ were obtained from the state $|D_4^{(2)}\rangle^{exp}$ by projection of the fourth qubit onto horizontal/vertical polarization, i.e., from measuring σ_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 $\rho_W^{nc,exp}$.

3. Entanglement detection by correlation measurements

9

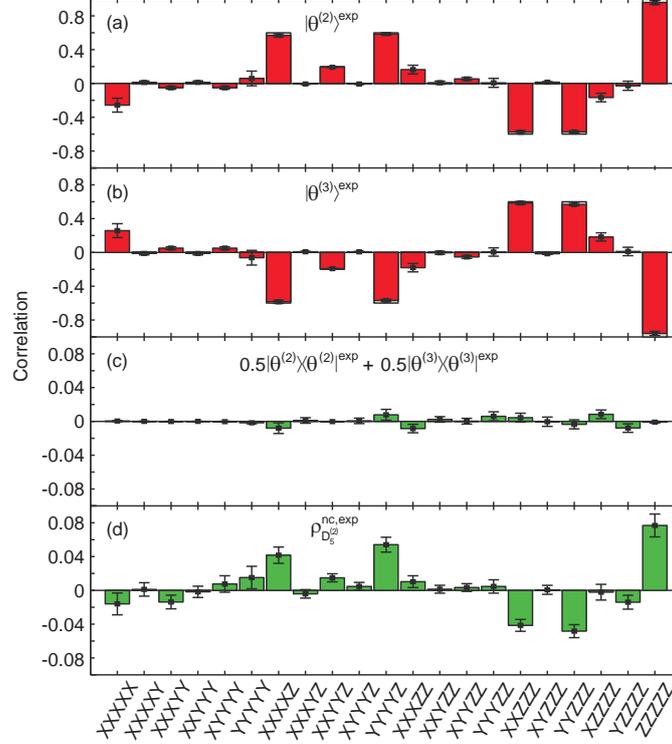


FIG. 6. Experimental five-partite symmetric correlations for the two most prominent states (a) $|\Theta^{(2)}\rangle^{exp}$ and (b) $|\Theta^{(3)}\rangle^{exp}$ in the eigen decomposition of the experimental density matrix $\rho_{D_5^{(2)}}^{nc,exp}$ shown in Fig. 4. The correlations of these states 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 $s_1 = x...xx$, $s_2 = x...xy$, ..., $s_{3^n} = z...zz$. 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., $r_1 = 0...00$, $r_2 = 0...01$, ..., $r_{2^n} = 1...11$. 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 $M_{r_j}^{s_i} = \frac{1}{2^n} \bigotimes_{k=1}^n \left(\frac{1}{3} + (-1)^{r_j(k)} \sigma_{s_i(k)} \right)$ [21, 22], where $\mathbf{1}$ denotes the 2×2 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 $\Delta^2 F_{|W\rangle}^{zzz} = 2.46e-05$.

r_j	000	001	001	011	100	101	110	111
$ \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

3. Entanglement detection by correlation measurements

11

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.

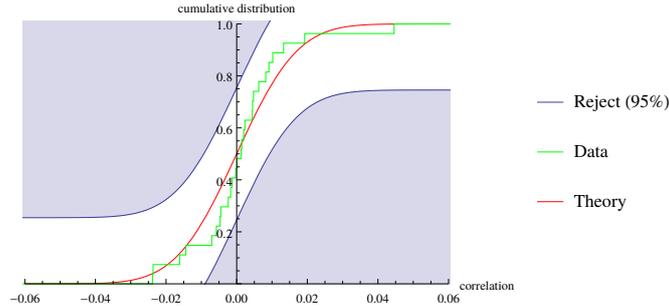


FIG. 7. The cumulative distribution of the experimentally determined correlations is compared to the cumulative distribution of the expected correlations ($\mu = 0$, $\sigma = 0.0135$). The shaded blue region contains points that would be sampled from the normal distribution with probability smaller 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 ρ_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

$$p = \Pr \left[\left(T, T_{D_5^{(2)}}^{nc,exp}\right) \geq 13.663 \mid H_0^{(5)} \right] < \frac{1}{\sqrt{2\pi}\sigma} \int_{13.663}^{\infty} dx \exp \left(-\frac{(x-\mu)^2}{2\sigma^2} \right) = 5.6 \times 10^{-3} \ll 0.05, \quad (33)$$

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]

$$\mathcal{B} = E_{\mathcal{P}(111110)} + E_{\mathcal{P}(22220)} + E_{\mathcal{P}(12220)} - E_{\mathcal{P}(21110)} - E_{\mathcal{P}(11000)} - E_{\mathcal{P}(22000)} \leq 6 \quad (34)$$

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

$$p = \Pr \left[\mathcal{B} \geq 6.358 \mid H_0^B \right] < \frac{1}{\sqrt{2\pi}\sigma} \int_{6.358}^{\infty} dx \exp \left(-\frac{(x-\mu)^2}{2\sigma^2} \right) = 0.0083 \ll 0.05. \quad (35)$$

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

$$\sigma_{(\theta_i, \phi_i)} = \sin(\theta_i) \cos(\phi_i) \sigma_x + \sin(\theta_i) \sin(\phi_i) \sigma_y + \cos(\theta_i) \sigma_z. \quad (37)$$

3. Entanglement detection by correlation measurements

13

Consequently,

$$\begin{aligned} T_{(\theta_1, \phi_1) (\theta_2, \phi_2) (\theta_3, \phi_3)} &= \sin(\theta_1) \cos(\phi_1) \sin(\theta_2) \cos(\phi_2) \sin(\theta_3) \cos(\phi_3) T_{xxx} \\ &\quad + \sin(\theta_1) \cos(\phi_1) \sin(\theta_2) \cos(\phi_2) \sin(\theta_3) \sin(\phi_3) T_{xxy} \\ &\quad + \dots \\ &\quad + \cos(\theta_1) \cos(\theta_2) \cos(\theta_3) T_{zzz}, \end{aligned} \tag{38}$$

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. 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, 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).

4. Quantum state tomography

The following chapter describes quantum state tomography, which has become one of the standard tools for the full analysis of quantum systems [32, 129, 211–215]. It considers both theoretical and experimental aspects of the topic. As throughout this thesis, here again, the focus lies entirely on qubits systems. For tomography of higher dimensional systems please refer to [216, 217] and for tomographic analysis in the continuous variable case, please see [218].

The chapter is structured as follows, in section 4.1 quantum state tomography in its most general form is discussed where no further assumptions about the structure of the state to be analyzed are made. The only constraint is that for each of the 4^N degrees of freedom of an N -qubit state, at least one measurement has to be performed. Then, a commonly applied tomography scheme is discussed, called Pauli tomography, where an overcomplete set of tomographic data is recorded by performing 6^N projection measurements. For both schemes, the experimentally determined relative frequencies are interpreted as approximations to the unknown underlying probabilities and are utilized to reconstruct the density matrix. However, this ansatz often fails to deliver a result that is physical. In such cases, numerical optimization algorithms are applied which yield a physical result best representing the observed data. The most widely used of these algorithms [211, 219] are discussed together with their numerical implementation.

In many situations, a complete analysis of a quantum state is not possible or just not attempted due to its huge experimental effort. However, certain classes of states possess an efficient representation with only few parameters and therefore allow for tomographic analysis at largely reduced cost. For example, states that can be represented as matrix product states of low bond dimension [117, 220] allow for the application of highly efficient tomography schemes as discussed in [77, 221–224]. The ansatz can even be generalized to *continuous* matrix product states as presented in [225–227]. Compressed sensing is another strategy for efficient tomography which benefits from the fact that in many experiments low rank states, i.e., states with only few non-zero eigenvalues, are prepared [55, 228–232]. For states which are permutationally invariant (PI), like GHZ states or symmetric Dicke states, PI tomography enables fully scalable tomographic analysis [56–58]. In general, any quantum state that can be expressed by means of a few parameter model allows for efficient tomography [233]. An extensive overview of the current state of the subject of quantum state reconstruction can be found in [234].

4. Quantum state tomography

In section 4.2, two partial tomography schemes, namely compressed sensing (CS) and PI tomography are presented in detail. Since many experimental states are at the same time low rank states and feature permutational invariance, also a combination of CS and PI tomography is considered. In publications P4.1 and P4.2, these highly efficient tomography schemes are compared against full tomography on the example of symmetric Dicke states with four and six photons. It could be shown that all tomography schemes deliver compatible results within their statistical errors. In this context, it is important that not only the measurement scheme itself is scalable but also data post processing. Therefore, an algorithm for the evaluation of permutationally invariant tomography is presented that on the one hand exploits the symmetry of the state and on the other hand resorts to methods from convex optimization, see publication P4.3.

Please note that all tomography schemes discussed in this work are static, i.e., a certain set of measurement operators is chosen *in advance* of the experiment. No information that is obtained in the course of the experiment is used to update the measurement operators for the remaining measurements. In contrast, in a so called adaptive scheme, the first n measurement results determine what the $n + 1$ st measurement will be. The main advantage of an adaptive tomography scheme is, that it converges faster towards the underlying unknown state. More precisely, this means that with the same amount of experimental data, a higher statistical significance can be achieved. However, this comes at the price of a higher numerical effort which can be considerable. Adaptive tomography schemes are treated both theoretically and experimentally in [235–242] for qubit systems. For adaptive homodyne quantum tomography please see [243, 244].

4.1. Standard tomography

The goal of full quantum state tomography is to experimentally determine the density matrix of an unknown quantum state ϱ from a multitude of identical copies of the state. In order to achieve this goal, the general ansatz is to find a decomposition of ϱ for a given set of measurement operators $\{M_i\}$ such that ϱ can be expressed as a linear function of the expectation values of the measurement operators $\{p_i\}$ with $p_i = \langle M_i \rangle$. More precisely, a decomposition of ϱ of the form

$$\varrho = \sum_i \mathcal{Y}_i p_i \tag{4.1}$$

is required. In the following, a general procedure is presented that allows to determine the generating set $\{\mathcal{Y}_i\}$ on the example of two widely used tomography schemes. The first scheme, projector-based complete tomography, requires 4^N projection measurements which corresponds to the number of free real parameters of an N -qubit state. It is best suited for experimental setups where

one detector per spatial mode is used [50, 211]. With a small modification, this scheme can be extended to *uniquely* determine ρ and infer the set $\{\mathcal{Y}_i\}$ even if more than 4^N projection measurements are performed. Please note that in the case of the minimum number of 4^N projection measurements, the $\{\mathcal{Y}_i\}$ form a *basis* of the N -qubit Hilbert space \mathbb{C}^{2^N} , i.e., a generating set with the smallest number of elements. The second scheme, Pauli tomography, utilizes the results of 6^N projection measurements where the projectors are given by the eigenvectors of the tensor products of all possible combinations of Pauli matrices σ_x , σ_y and σ_z . Pauli tomography is the method of choice for setups with two detectors per spatial mode [89, 129, 207]. In all other cases, projector-based complete tomography is to be preferred.

4.1.1. Projector-based complete tomography

Every N -qubit quantum state ρ can be represented by a Hermitian $2^N \times 2^N$ matrix with 4^N complex entries. Due to Hermiticity and normalization, the number of real parameters that are required to uniquely characterize ρ reduces to 4^N . Since for each parameter one measurement has to be performed, at least 4^N measurements are necessary to determine ρ . In the most general form, the measurement operators are a so-called positive operator valued measure (POVM) with all $\{M_i\}$ being positive-semidefinite Hermitian operators summing up to $\mathbb{1}^{\otimes N}$, i.e.,

$$\langle \phi | M_i | \phi \rangle \geq 0 \quad \forall \phi \quad \text{and} \quad \sum_i M_i = \mathbb{1}^{\otimes N}. \quad (4.2)$$

An important difference between a basis of the Hilbert space \mathcal{H} and a POVM is that the number of elements of the POVM can be larger than the dimension of \mathcal{H} . Experimentally, it is often advantageous to perform projection measurements because they are simpler to implement. One then speaks of a projector valued measure (PVM) instead of a POVM.

In the following, it will be explained how a decomposition of ρ of the form Eq. 4.1 can be inferred for PVMs with the minimal number of elements necessary for quantum state tomography [211]. Therefore, let us rewrite Eq. 4.1 as

$$\rho = \sum_{\substack{i_1, \dots, i_N \\ \in \{0,1,2,3\}}} \mathcal{Y}_{i_1, \dots, i_N} \langle \phi_{i_1} \otimes \dots \otimes \phi_{i_N} \rangle \quad (4.3)$$

where the set $\{\mathcal{Y}_{i_1, \dots, i_N}\}$ forms a basis of the Hilbert space and solely depends on the projection measurements $\phi_{i_1} \otimes \dots \otimes \phi_{i_N}$. In order to obtain the elements of $\{\mathcal{Y}_{i_1, \dots, i_N}\}$, it is best to first express ρ in the standard Pauli basis

$$\rho = \frac{1}{2^N} \sum_{\substack{i_1, \dots, i_N \\ \in \{0,x,y,z\}}} T_{i_1, \dots, i_N} \sigma_{i_1} \otimes \dots \otimes \sigma_{i_N} \quad (4.4)$$

4. Quantum state tomography

with correlations tensor T_{i_1, \dots, i_N} as introduced in section 2.2.1. For better readability, the shorthand notations $\Gamma_0 = \sigma_0 \otimes \sigma_0 \otimes \dots \otimes \sigma_0$, $\Gamma_1 = \sigma_0 \otimes \sigma_0 \otimes \dots \otimes \sigma_x$ etc., $T_{i_1, \dots, i_N} \rightarrow T_\mu$, $\mathcal{Y}_{i_1, \dots, i_N} \rightarrow \mathcal{Y}_\lambda$, and $\phi_{i_1} \otimes \dots \otimes \phi_{i_N} \rightarrow \phi_\lambda$ are used. In this notation, the expectation values $p_\lambda = \langle \phi_\lambda \rangle$ are then given by

$$p_\lambda = \text{Tr}(\varrho \phi_\lambda) = \text{Tr}\left(\frac{1}{2^N} \sum_\mu T_\mu \Gamma_\mu \phi_\lambda\right) = \frac{1}{2^N} \sum_{\mu=1}^{4^N} \text{Tr}(\phi_\lambda \Gamma_\mu) T_\mu. \quad (4.5)$$

Introducing the matrix $B_{\lambda, \mu} = \frac{1}{2^N} \text{Tr}(\phi_\lambda \Gamma_\mu)$ allows to interpret Eq. 4.5 as a system of linear equations

$$\vec{p} = B \vec{T}. \quad (4.6)$$

Now \vec{T} can be expressed as

$$\vec{T} = B^{-1} \vec{p} \quad (4.7)$$

or equivalently

$$T_\mu = B_{\mu, \lambda}^{-1} p_\lambda \quad (4.8)$$

where B^{-1} is the inverse of B . Please note that for a tomographically complete set of projectors, B^{-1} is mathematically always well-defined. Hence, the existence of B^{-1} can therefore also be seen as a check whether the set of projectors $\{\phi_\lambda\}$ can deliver a tomographically complete set of measurement data. Combining Eq. 4.8 and Eq. 4.4 delivers the desired result

$$\varrho = \frac{1}{2^N} \sum_\mu T_\mu \Gamma_\mu = \frac{1}{2^N} \sum_{\mu, \lambda} B_{\mu, \lambda}^{-1} \Gamma_\mu p_\lambda. \quad (4.9)$$

From a comparison of Eq. 4.9 with Eq. 4.3, it can be seen that

$$\mathcal{Y}_\lambda = \frac{1}{2^N} \sum_\mu B_{\mu, \lambda}^{-1} \Gamma_\mu. \quad (4.10)$$

Due to finite measurement statistics, the probabilities p_λ are experimentally not accessible and can only be approximated. Thus, for experimental data, the probabilities p_λ in Eq. 4.9 are replaced with the respective frequencies f_λ

$$\varrho = \sum_\lambda \mathcal{Y}_\lambda f_\lambda. \quad (4.11)$$

A common choice for the local projectors are the eigenvectors with eigenvalue +1 of the Pauli matrices σ_x , σ_y and σ_z plus one additional projector for normalization. One possible set of projectors is, e.g., $\phi_{0_k} = |x^+\rangle\langle x^+|$, $\phi_{1_k} = |y^+\rangle\langle y^+|$, $\phi_{2_k} = |z^+\rangle\langle z^+|$ and $\phi_{3_k} = |z^-\rangle\langle z^-|$ with $k = 1, \dots, N$. In order to correctly normalize the state, it is necessary that a subset of the measurement projectors adds up to identity. If one then measures all projectors of this subset for the same amount of

time t_{meas} and assumes that the number of copies generated by the experimental apparatus is constant over time, one can infer how many copies \mathcal{N} are generated per time interval. Then the relative frequencies can be calculated as $f_\lambda = \frac{c_\lambda}{\mathcal{N}}$ where c_λ is the number of detection events when measuring the projector ϕ_λ for time t_{meas} . Please note that the above procedure is not only applicable when the number of projectors to be measured is 4^N . It can easily be generalized to situations with more than the required 4^N projection measurements where a so called overcomplete set of tomographic data is recorded. Then, the matrix B is not a square matrix anymore and the inverse of B in Eq. 4.7 has to be replaced by the Moore-Penrose pseudo-inverse [245].

4.1.2. Pauli tomography scheme

As explained in the previous section, complete information about a quantum state can already be obtained from a minimal tomographic set. However, it is often preferable to measure an overcomplete tomographic set which is advantageous for a number of reasons. The first reason is that in many experiments [32, 129, 207], the results of two orthogonal projectors per qubit can be obtained at almost no additional experimental cost. Therefore, one often chooses to measure in the eigenbases of all 3^N possible combinations of tensor products of Pauli matrices each consisting of 2^N projection measurements. Hence, in total, 6^N measurement results are recorded. Another positive factor of this tomography scheme, also known as Pauli tomography, is that it is more robust with respect to fluctuations of the count rate since normalization is intrinsically guaranteed for each measurement basis. A further advantage is that, by using more detection devices (one for $|x^+\rangle\langle x^+|$ and one for $|x^-\rangle\langle x^-|$ etc.), in the same measurement time, more data is recorded compared to a minimal tomographic set because the detection probability for *any* state is always one, independent of the measurement basis. In our case, the projectors are chosen as $\phi_{0_k} = |x^+\rangle\langle x^+|$, $\phi_{1_k} = |x^-\rangle\langle x^-|$, $\phi_{2_k} = |y^+\rangle\langle y^+|$, $\phi_{3_k} = |y^-\rangle\langle y^-|$, $\phi_{4_k} = |z^+\rangle\langle z^+|$ and $\phi_{5_k} = |z^-\rangle\langle z^-|$ with $k = 1, \dots, N$. In order to determine the density matrix of the state, a decomposition of ϱ similar to Eq. 4.3 has to be found. In principle, it is possible to use the formalism presented in section 4.1.1 and [211, 246] to find such a decomposition. However, there is no need to resort to the Moore-Penrose pseudo-inverse since an elegant analytic derivation for a linear decomposition of ϱ exists [89] which delivers

$$\varrho = \frac{1}{2^N} \sum_{\substack{i_1, \dots, i_N \\ \in \{x, y, z\}}} \sum_{\substack{m_1, \dots, m_N \\ \in \{0, 1\}}} p_{m_1, \dots, m_N}^{i_1, \dots, i_N} \bigotimes_{k=1}^N \left(\frac{1}{3} \mathbb{1} + (-1)^{m_k} \sigma_{i_k} \right). \quad (4.12)$$

4. Quantum state tomography

The probabilities $p_{m_1, \dots, m_N}^{i_1, \dots, i_N}$ are given by the expectation values of the respective projection measurements

$$p_{m_1, \dots, m_N}^{i_1, \dots, i_N} = \frac{1}{2^N} \text{Tr} \left(\varrho \bigotimes_{k=1}^N (\mathbb{1} + (-1)^{m_k} \sigma_{i_k}) \right). \quad (4.13)$$

Again, for finite measurement statistics, the probabilities have to be replaced by the experimentally determined relative frequencies $f_{m_1, \dots, m_N}^{i_1, \dots, i_N} \approx p_{m_1, \dots, m_N}^{i_1, \dots, i_N}$. Please note that due to the overcompleteness of the tomographic set, the elements of the generating set $\{\bigotimes_{k=1}^N (\frac{1}{3}\mathbb{1} + (-1)^{m_k} \sigma_{i_k})\}$ are non-orthogonal and there are, in principle, infinitely many possible decompositions of ϱ . Nonetheless, Eq. 4.12 can be seen as a natural decomposition of ϱ where all the experimental data is considered and directly contributes to the final result.

4.1.3. Physicality constraint

Both tomography schemes presented in section 4.1.1 and 4.1.2 interpret the experimental data from a frequentistic point of view and do not impose any constraint on the result. Due to unavoidable statistical fluctuations of the count rates, the experimentally determined density matrix often fails to be positive-semidefinite, i.e., in many cases it contains at least one negative eigenvalue. Since the eigenvalues are interpreted as probabilities, this poses a severe conceptual problem. Especially when functions like, e.g., entanglement measures (see section 2.2.4) are to be calculated, the result might be meaningless for unphysical states or the function is mathematically just ill-defined outside the physical regime. For a more elaborate discussion about possible causes for negative eigenvalues, please see section 5.1. Here, the focus is entirely on the standard approach which guarantees that the experimentally determined state lies within the physical regime and, at the same time, optimally reflects the experimental data. More precisely, the task is to find a state such that a target function F_T is minimized for the experimentally determined set of frequencies $f = \{f_\lambda\}$

$$\hat{\varrho} = \arg \min_{\varrho \geq 0_T} F_T(\varrho|f). \quad (4.14)$$

The two most common choices for the target function F_T are maximum likelihood [219]

$$F_T^{MLE} = - \sum_{\lambda} f_{\lambda} \log(\text{Tr}(\varrho|\phi_{\lambda})\langle\phi_{\lambda}|) \quad (4.15)$$

and least squares [211]

$$F_T^{LS} = \sum_{\lambda} \frac{1}{w_{\lambda}} [f_{\lambda} - \text{Tr}(\varrho|\phi_{\lambda})\langle\phi_{\lambda}|]^2 \quad (4.16)$$

with weights w_λ . There are basically two options for the weights, to leave them fixed, $w = 1/f_\lambda$, or to leave them free, $w = 1/\text{Tr}(\varrho|\phi_\lambda\rangle\langle\phi_\lambda|)$. Please note that for the choice $w = 1/f_\lambda$ and especially for small sample sizes, the case $f_\lambda = 0$ may occur for some λ , leading to a singularity of F_T^{LS} . In the following, three widely used numerical procedures to find the optimum of Eq. 4.14 will be discussed.

Iterative fixed point algorithm

An iterative algorithm to find the optimum of Eq. 4.15 was proposed by Hradil and is discussed, e.g., in [244, 247, 248]. One starts with the completely mixed state $\varrho^{(0)} = \frac{1}{2^n}\mathbb{1}^{\otimes n}$ as initial guess. Then one repetitively applies the update rule

$$\varrho^{(k+1)} = \mathcal{N}[R(\varrho^{(k)})\varrho^{(k)}R(\varrho^{(k)})] \quad (4.17)$$

with

$$R(\varrho^{(k)}) = \sum_\lambda \frac{f_\lambda}{\text{Tr}(\varrho^{(k)}M_\lambda)} M_\lambda \quad (4.18)$$

and normalization constant \mathcal{N} such that $\text{Tr}(\varrho^{(k+1)}) = 1$ to determine $\varrho^{(k+1)}$. Although this algorithm is a generalization of the classical expectation-maximization algorithm as presented in [249], its convergence towards the optimum of Eq. 4.14 cannot always be guaranteed. Nonetheless, in practical applications, its convergence rate is often sufficiently high and it has to be noted that each step of the algorithm requires only matrix-matrix multiplication which can be implemented efficiently on a classical computer.

Cholesky decomposition

Another easy to implement numerical procedure to find the optimum of F_T is given by James et al. in [211]. It is based on the fact that every non-negative Hermitian matrix can be decomposed into a product of a lower triangular matrix Λ and its conjugate transposed Λ^\dagger . This decomposition, also known as Cholesky decomposition, is always given for non-negative Hermitian matrices. Thus, by parameterizing ϱ as

$$\varrho = \frac{\Lambda\Lambda^\dagger}{\text{Tr}(\Lambda\Lambda^\dagger)} \quad (4.19)$$

with

$$\Lambda = \begin{pmatrix} t_1 & 0 & \cdots & 0 & 0 \\ t_{n+1} + it_{n+2} & t_2 & \cdots & 0 & 0 \\ t_{n+3} + it_{n+4} & t_{n+5} + it_{n+6} & t_3 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ t_{4^{n-2^n-1}} + it_{4^{n-2^n}} & t_{4^{n-2^n+1}} + it_{4^{n-2^n+2}} & \cdots & \cdots & t_n \end{pmatrix}, \quad (4.20)$$

4. Quantum state tomography

it is guaranteed that ϱ is a normalized non-negative Hermitian matrix. Hence, for any choice of $\vec{t} \in \mathbb{R}^{4^N}$, $\varrho(\vec{t})$ can be interpreted as a valid density matrix of a quantum state. Then the constrained optimization problem of Eq. 4.14 can be transformed in an *unconstrained* one

$$\hat{\varrho} = \arg \min_{\vec{t} \in \mathbb{R}^{4^N}} F_T(\varrho(\vec{t})|f). \quad (4.21)$$

which can be solved with the aid of the build-in functions *FindMinimum* in Mathematica[®] or *fminunc* in MATLAB[®].

Convex optimization

The above algorithms to find the optimum of Eq. 4.15 and Eq. 4.16 are easy to implement on a digital computer, which is their great advantage, and most likely the main reason for their wide usage. However, both algorithms do not explicitly take advantage of the convexity of the target functions F_T^{MLE} and F_T^{LS} . For convex functions, the methods from the mathematically fully developed field of convex optimization can be applied which hugely benefit from the fact that convex functions have a unique optimum [57, 250, 251]. In consequence, this means that one can estimate by how much the true solution of Eq. 4.14 differs from the current iteration. Therefore, one can easily give a stopping condition for a fitting algorithm, i.e., a prescription when no further iterations are necessary.

In order to actually apply convex optimization to Eq. 4.22, one still has to restrict the solution to the physical regime. This can be achieved by adding a so called barrier term to the respective target function

$$F_T(\varrho) \longrightarrow \tilde{F}_T(\varrho) = F_T(\varrho) + \varepsilon \log[\det(\varrho)]. \quad (4.22)$$

The barrier term $\log[\det(\varrho)]$ diverges for rank-deficit states, i.e., for states with at least one vanishing eigenvalue. Thus, it penalized states that are “too close” to the border of the state space. In the limit $\varepsilon \rightarrow 0$, the function $\tilde{F}_T(\varrho)$ converges to the original function $F_T(\varrho)$. Please note that also the function of Eq. 4.22 is convex. In order to find the optimum of Eq. 4.14, one chooses a parametrization of the form

$$\varrho(x) = \frac{1}{2^N} \mathbb{1}^{\otimes N} + \sum_i x_i S_i \quad (4.23)$$

where $x \in \mathbb{R}^{4^N-1}$ and $\{S_i\}$ (together with $\frac{1}{2^N} \mathbb{1}^{\otimes N}$) is some appropriate operator basis. For each choice of ε , the optimum can be found by a Newton method [252]. Starting with $x_0 = 0$ and, e.g., $\varepsilon = 1$, one finds the next iteration by

$$x_{i+1} = x_i - \beta [H(\tilde{F}_T(x_i))]^{-1} \nabla \tilde{F}_T(x_i) \quad (4.24)$$

where $H(\tilde{F}_T(x_i))$ is the Hessian matrix at x_i and β is determined by a so-called backtracking line search [250, 252] to fulfill the Wolfe conditions [253]. For

more details on how to determine the Hessian matrix $H(\tilde{F}_T(x_i))$ and the gradient $\nabla\tilde{F}_T(x_i)$ for the respective target function, please see appendix A and [57]. When the 2-norm of the gradient $\|\nabla\tilde{F}_T(x_i)\|_2$ falls below some limit, e.g. $\|\nabla\tilde{F}_T(x_i)\|_2 < 10^{-7}$, ε is decreased further and the optimization procedure is repeated. Here, it is crucial to decrease ε in a way that during the next optimization, quadratic convergence is achieved after a few iterations [252]. In the work presented here, ε was reduced by a factor of 10 each time until $\varepsilon = 10^{-10}$ which turned out to deliver fast convergence and good numerical accuracy of the result. As already mentioned above, the great advantage of convex optimization is that the quality of the solution can directly be quantified. As shown in [57], the difference between $\tilde{F}_T(\varrho)$ and $F_T(\varrho)$ is bounded by

$$\tilde{F}_T(\varrho) - F_T(\varrho) \leq \varepsilon \dim(\mathcal{H}) \quad (4.25)$$

where $\dim(\mathcal{H})$ is the dimension of the Hilbert space. Thus, for small ε , the solution of Eq. 4.22 gives an excellent approximation to the solution of Eq. 4.14.

4.1.4. Comparison of algorithms

Next, the convergence of the fitting algorithms introduced in section 4.1.3 shall be investigated. Therefore, in a first step, tomographic data of random two- to six-qubit states ϱ_0 are simulated in order to eliminate any unwanted experimental imperfections and statistical noise. For these data, the result $\hat{\varrho}$ for the respective fit functions Eq. 4.15 and Eq. 4.16 has to be identical to the initial state. The figure of merit in this context is the number of iterations of the respective algorithm that is necessary until the distance between ϱ_0 and $\hat{\varrho}$ falls below some given threshold. In the following, the distance between ϱ_0 and $\hat{\varrho}$ is quantified by the trace norm. Please note that here, the different algorithms are compared by the number of iterations that are needed and not by the time they require because then, the actual implementation would have a too strong influence, i.e., in which programming language the respective algorithm is written, which linear algebra libraries are available etc. For a fair comparison, it has to be mentioned that the short run time of the convex optimization algorithm, as implemented in this work, is to a large part due to the usage of the Hessian matrix. However, for even larger qubit systems, the determination of the Hessian can become unfeasible. In order to also benefit from convex optimization techniques in such cases, one possibility is to resort to quasi Newton methods like, e.g., Broyden [254], DFP [255, 256] or BFGS [257–260], where the Hessian matrix is only approximated. Alternatively, conjugate gradient methods [261] allow for optimization without the need to determine second derivatives.

4. Quantum state tomography

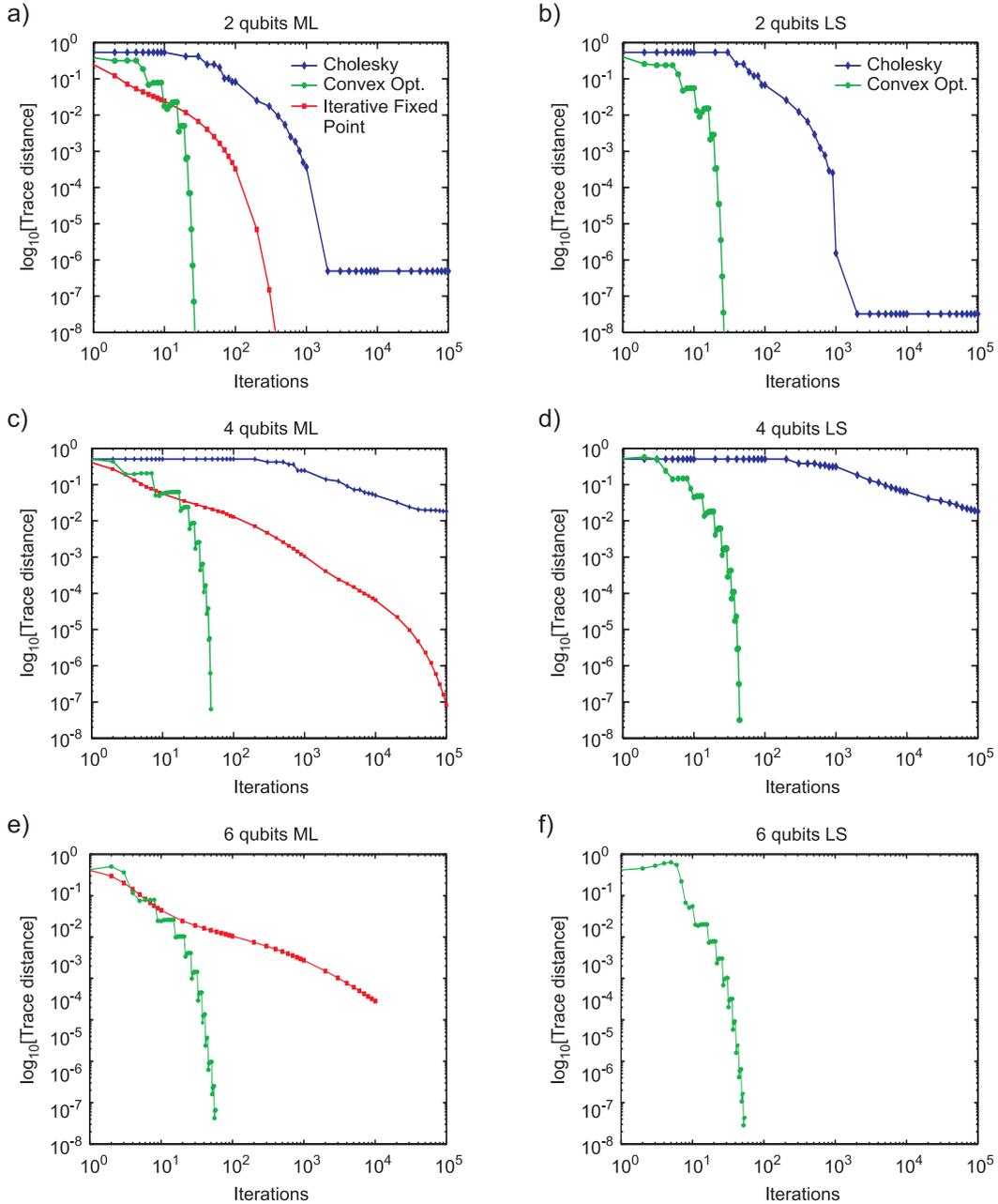


Figure 4.1.: Convergence of ML and LS fitting algorithms based on the Cholesky decomposition, the iterative fixed point algorithm, and convex optimization. The quality of the fit is quantified by the trace distance between the fitted and the initial state. For two qubits, as can be seen in subfigures a) and b), all algorithms achieve a good accuracy for the final result, although the convergence rates differ significantly. For higher qubit numbers however, as shown in subfigures c)-f), the supremacy of convex optimization becomes obvious. Please note that due to limitations of memory and computational power, the fitting algorithm based on the Cholesky decomposition could not be used for six qubits.

4.2. Partial tomography schemes

The major drawback of full quantum state tomography is the exponential scaling of the number of measurements with the system size. As shown above, at least 4^N measurements are necessary to gain complete knowledge of an N -qubit state. Moreover, this unfavorable scaling also poses, from today's point of view, an insurmountable obstacle for data post processing on digital computers. For example, in order to store the 4^{30} measurement results obtained from the complete tomographic analysis of a 30-qubit state in double precision, 10 Exabyte of disc space would be required! However, already for much smaller system sizes, the limits of full quantum state tomography are reached due to experimental restrictions, like the maximum measurement time or drifts of the setup, to name just a few. Hence, it is highly desirable to develop tomography schemes with better than exponential scaling by using additional knowledge like that the state is of low rank, features a certain symmetry or allows for some efficient representation [55–58, 77, 221–224, 228–233, 262–267].

In this thesis, the focus lies on two efficient partial tomography schemes: The first one is *compressed sensing*, which exploits the fact that most experimentally prepared states are close to pure states or are at least mixtures of only a few pure states. The second scheme is *permutationally invariant tomography* which is tailored to states that are invariant under particle swapping. The discussion starts with CS, as it is more closely related to the Pauli tomography scheme in the sense that it also uses Pauli measurements. After that, PI tomography is treated.

4.2.1. Compressed sensing

The first precursors of compressed or compressive sensing date back to the 1970's when seismologists observed that even from data that seemingly violated the Nyquist-Shannon theorem, it was possible to reconstruct images of reflective layers within the earth [268, 269]. However, at that time, a rigorous mathematical theory of the observed effect was still missing and it took two more decades until Candés et al. [270] and Donoho [271] were able to put the method on solid mathematical ground. Shortly after, CS was adopted to quantum state tomography [55, 228–231, 272]. Although the mathematical theory behind CS goes way beyond the scope of this work, an intuitive understanding of CS can already be obtained from a simple parameter counting argument. A pure quantum state is described by only $2^N - 1$ complex parameters, whereas a state with full rank requires $4^N - 1$ real parameters in order to be uniquely described. Hence, if one knows or if one can certify that a state is pure or of low rank, i.e., only few eigenvalues are non-zero, it is plausible that the measurement effort can be reduced. However, it is not a priori clear which settings should be measured in order to get the maximum information which leads to some oversampling c . More precisely, the measure-

4. Quantum state tomography

ment effort of CS is $O(\log(cr2^N))$ for a state with rank r [55]. In order to obtain the density matrix, the respective target function F_T from Eq. 4.15 or Eq. 4.16 has to be minimized for an tomographically *incomplete* set of measurement data. In practice, one performs the optimization for more and more measurement data until convergence sets in, i.e., further measurements do not alter the result any more and then one stops measuring, see Fig. 4.2. Please note, that in [55] CS

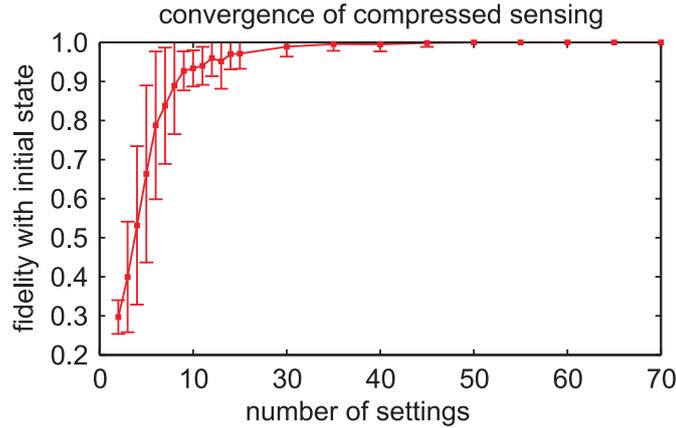


Figure 4.2.: Simulation of CS on the example of a four-qubit state of rank three. As can be seen, as soon as the number of (randomly) chosen measurement settings surpasses 30, the initial state is perfectly reconstructed.

for quantum state tomography is discussed for ℓ_1 minimization and in [228] optimization based on the matrix Dantzig selector and the matrix Lasso is considered. In this thesis, however, the target function F_T^{MLE} is chosen together with convex optimization for a better comparability with the standard approach.

4.2.2. Permutationally invariant tomography

Permutationally invariant states allow for tomographic analysis at largely reduced cost because they can be parametrized efficiently. In detail, PI tomography allows to determine the projection of a quantum state on the permutationally invariant subspace. Every quantum state ϱ can be decomposed into a PI part ϱ_{PI} and a part which is not PI $\varrho_{\overline{\text{PI}}}$,

$$\varrho = \varrho_{\text{PI}} + \varrho_{\overline{\text{PI}}}. \quad (4.26)$$

For any quantum state ϱ , the PI part can be determined as

$$\varrho_{\text{PI}} = \frac{1}{N!} \sum_i \Pi_i \varrho \Pi_i^\dagger \quad (4.27)$$

4.2. Partial tomography schemes

where the Π_i are permutation operators for qubits. For many relevant quantum states like GHZ states, W states or symmetric Dicke states, the non-permutationally invariant part vanishes and thus, even in many experimental realizations, only determining ϱ_{PI} is a fairly good approximation to ϱ . So far, determining ϱ_{PI} has been theoretically discussed for spin systems in [273, 274] and experimentally implemented for polarized photons in a single mode optical fiber in [275–277].

In the following, two different efficient representations of ϱ_{PI} shall be discussed and how they can be exploited for tomographic analysis. The first one is based on *PI* correlations, which allows for a simple derivation of the number of free parameters. The second representation utilizes a block diagonalization of ϱ_{PI} together with coarse graining, which proves to be advantageous especially for maximum likelihood and least squares fitting algorithms.

Formally, a PI correlation $T_{i_1, \dots, i_N}^{\text{PI}}$ is defined as

$$T_{i_1, \dots, i_N}^{\text{PI}} = \langle \sigma_{i_1} \otimes \dots \otimes \sigma_{i_N} \rangle_{\text{PI}} = \frac{1}{N!} \sum_i \text{Tr}[\varrho \Pi_i(\sigma_{i_1} \otimes \dots \otimes \sigma_{i_N}) \Pi_i^\dagger]. \quad (4.28)$$

This definition reflects the fact that for PI states, the result of a correlation measurement is independent under swapping of local measurements. A PI quantum state is uniquely described by its PI correlations, also called *generalized* Bloch vector. Thus, in order to determine the measurement effort, it is necessary to count the number of PI correlations. Therefore, let us consider Fig. 4.3a) where a simple scheme to count PI correlations on the example of four qubits is shown. There are three vertical red bars to separate the local settings σ_0 , σ_x , σ_y and σ_z . The horizontal black place holders left of the red bar 1 are filled up with σ_0 , the place holders between the red bars 1 and 2 are filled up with σ_x , those between the red bars 2 and 3 are filled up with σ_y , and, finally, those right of the red bar 3 are filled up with σ_z . The PI correlation corresponding to this arrangement of red the bars is shown beneath. A second example is shown in Fig. 4.3b). With this scheme, one can easily count all PI correlations of an N -qubit state. The total

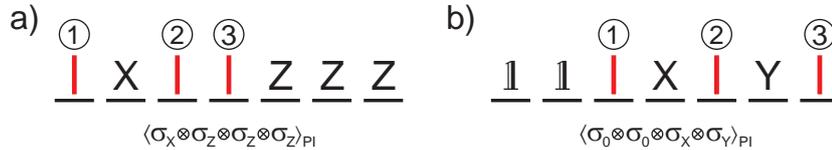


Figure 4.3.: With a simple combinatorial argument, the number of parameters that need to be determined for permutationally invariant tomography is $\binom{N+3}{N}$. For details, see the main text.

number of PI correlations is given then by the number of possible distributions of the three red bars which is, as can easily be seen, $\binom{N+3}{N} = \frac{1}{6}(N^3 + 6N^2 + 11N + 6)$.

4. Quantum state tomography

Thus the number of free parameters of a PI state is $O(N^3)$. Hence, for an optimal scheme, the measurement effort scales polynomially with the number of qubits.

In order to determine all PI correlations, an efficient approach is to measure a set of global operators $A_j^{\otimes N}$ ($j = 1, \dots, D_N$) where for each measurement setting, the same local operator is measured on all qubits, with $A_j = \vec{n}_j \cdot \vec{\sigma}$, $|\vec{n}_j| = 1$, $\vec{\sigma} = (\sigma_x, \sigma_y, \sigma_z)$ and with $D_N = \binom{N+2}{N} = \frac{1}{2}(N^2 + 3N + 2)$ [56]. In principle, the set of operators can be chosen quite arbitrarily, however, in order to achieve small errors, it is best to choose an optimized set of measurement operators. From these measurements, all expectation values of the form $\langle A_j^{\otimes(N-n)} \otimes \mathbb{1}^{\otimes n} \rangle_{\text{PI}}$ ($n = 0, \dots, N - 1$) can be inferred. All expectation values $\langle A_j^{\otimes(N-n)} \otimes \mathbb{1}^{\otimes n} \rangle_{\text{PI}}$ can be inferred from collective measurements alone, i.e., only the number of *up* and *down* clicks has to be recorded but not at which individual site they occurred. From each measurement setting, N measurement results are obtained. Please note that $n = N$ is excluded because $\langle \mathbb{1}^{\otimes N} \rangle \equiv 1$ does not deliver any additional information. The total number of parameters determined in this ansatz is hence $N \binom{N+2}{N} = \frac{1}{2}(N^3 + 3N^2 + 2N)$ which has, up to an oversampling of 3 in the limit $N \rightarrow \infty$, the optimal scaling. The PI correlations in the standard Pauli basis can be inferred from the measurement operators as [56]

$$\langle \sigma_{i_1} \otimes \dots \otimes \sigma_{i_N} \rangle_{\text{PI}} = \sum_{j=1}^{D_N} c_j^{k,l,m} \langle A_j^{\otimes(N-n)} \otimes \mathbb{1}^{\otimes n} \rangle_{\text{PI}} \quad (4.29)$$

with real coefficients $c_j^{k,l,m}$. For every fixed set of measurement operators A_j , the coefficients $c_j^{k,l,m}$ can be determined such that the sum of variances of all PI correlations

$$(\mathcal{E}_{\text{total}})^2 = \sum_{k+l+m+n=N} \mathcal{E}^2[(\sigma_{i_1} \otimes \dots \otimes \sigma_{i_N})_{\text{PI}}] \left(\frac{N!}{k!l!m!n!} \right) \quad (4.30)$$

is minimal as described in [56]. For Poissonian measurement statistics, as common in photonic experiments, Eq. 4.30 can be recast as

$$(\mathcal{E}_{\text{total}})^2 = \sum_{k+l+m+n=N} \sum_{j=1}^{D_N} |c_j^{k,l,m}|^2 \frac{[\Delta(A_j^{\otimes(N-n)} \otimes \mathbb{1}^{\otimes n})]}{\lambda_j - 1} \left(\frac{N!}{k!l!m!n!} \right) \quad (4.31)$$

where λ_j is the parameter of the Poissonian distribution.

In order to obtain possibly small variances, it is, additional to an optimal choice of coefficients $c_j^{k,l,m}$, necessary to optimize the measurement operators A_j . As already mentioned above, $A_j = \vec{n}_j \cdot \vec{\sigma}$ with $|\vec{n}_j| = 1$, and therefore the measurement operators can be visualized on the Bloch sphere. Here, the best choice turned out to be an even distribution of the A_j 's, as shown in Fig. 4.4. Please note that it is possible to include the settings $\sigma_x^{\otimes N}$, $\sigma_y^{\otimes N}$ and $\sigma_z^{\otimes N}$ without significantly changing the variances in Eq. 4.30. From these measurement settings, various fidelity

witnesses [278], entanglement witnesses [61, 158, 278, 279] and, more important in this context, a lower bound for the overlap with the symmetric subspace can be deduced [56].

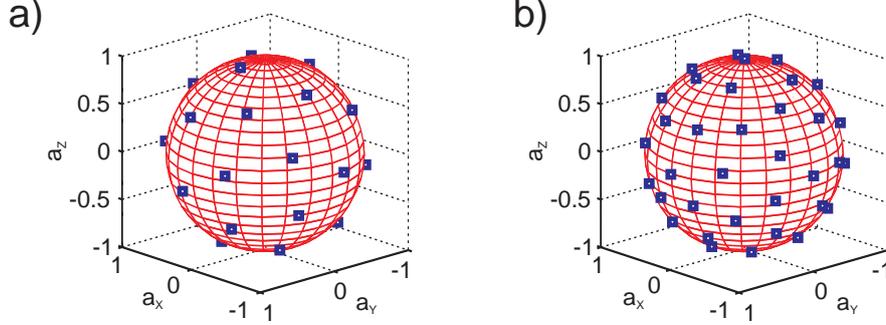


Figure 4.4.: Optimized set of measurement operators A_j visualized on the Bloch sphere for a) four and b) six-qubit PI tomography.

Alternatively to the approach presented above, which is closely linked to the concept of correlations, another ansatz based on a block diagonal decomposition of ϱ and coarse graining exists. It is advantageous both with respect to linear inversion and a maximum likelihood reconstruction. As already mentioned, the PI part of a state ϱ can be reconstructed from collective measurements alone. This property is reflected by the coarse grained measurement operators M_j^n [57]

$$M_j^n = \sum_i \Pi_i |0\rangle_j \langle 0|_j^{\otimes N-n} \otimes |1\rangle_j \langle 1|_j^{\otimes N-n} \Pi_i^\dagger \quad (4.32)$$

$$= \binom{N}{n} \left[|0\rangle_j \langle 0|_j^{\otimes N-n} \otimes |1\rangle_j \langle 1|_j^{\otimes N-n} \right]_{\text{PI}} \quad (4.33)$$

where Π_i are permutation operators for qubits and $|0\rangle_j$ and $|1\rangle_j$ are the eigenvectors of A_j with eigenvalue $+1$ and -1 , respectively. Using the expectation values of the coarse grained measurement operators, ϱ_{PI} can be expressed as a linear function of these expectation values in a similar way as in section 4.1.1. More precisely, a basis of the PI subspace $\{\mathcal{Y}_j^{n,\text{PI}}\}$, similar to the set $\{\mathcal{Y}_i\}$ in Eq. 4.1, has to be found, such that ϱ_{PI} can be expressed as

$$\varrho_{\text{PI}} = \sum_{j=1}^{D_N} \sum_{n=1}^N \mathcal{Y}_j^{n,\text{PI}} \langle M_j^n \rangle. \quad (4.34)$$

Similar to section 4.1.1, a shorthand notation is introduced, $\mathcal{Y}_j^{n,\text{PI}} \rightarrow \mathcal{Y}_\lambda^{\text{PI}}$ and $M_j^n \rightarrow M_\lambda$. As explained, ϱ_{PI} is uniquely described by its PI correlations, here labeled as T_μ^{PI} ,

$$\varrho_{\text{PI}} = \frac{1}{2^N} \sum_{\mu} T_\mu^{\text{PI}} \Gamma_\mu^{\text{PI}} \quad (4.35)$$

4. Quantum state tomography

with Γ_μ^{PI} the corresponding element of the symmetrized Pauli basis. Then the expectation value $p_\lambda = \langle M_\lambda \rangle$ can be determined as

$$p_\lambda = \text{Tr}(\varrho_{\text{PI}} M_\lambda) = \text{Tr}\left(\frac{1}{2^N} \sum_\mu T_\mu^{\text{PI}} \Gamma_\mu^{\text{PI}} M_\lambda\right) = \frac{1}{2^N} \sum_\mu \text{Tr}(M_\lambda \Gamma_\mu^{\text{PI}}) T_\mu^{\text{PI}} \quad (4.36)$$

Similar to Eq. 4.7 in section 4.1.1, introducing matrix $B_{\lambda,\mu} = \frac{1}{2^N} \text{Tr}(M_\lambda \Gamma_\mu^{\text{PI}})$, allows to interpret Eq. 4.36 as a system of linear equations

$$\vec{p} = B \vec{T}^{\text{PI}}. \quad (4.37)$$

The further derivation of the basis elements $\mathcal{Y}_\lambda^{\text{PI}}$ goes along the same line as in the case of standard tomography discussed in section 4.1. Therefore, only the final result shall be given

$$\mathcal{Y}_\lambda^{\text{PI}} = \frac{1}{2^N} \sum_\mu B_{\mu,\lambda}^{-1} \Gamma_\mu^{\text{PI}}. \quad (4.38)$$

As already mentioned before, due to finite measurement statistics, the probabilities p_λ can only be approximately determined in real experiments, i.e., $f_\lambda \approx p_\lambda$, and thus

$$\varrho_{\text{PI}} = \sum_\lambda \mathcal{Y}_\lambda^{\text{PI}} f_\lambda. \quad (4.39)$$

Note that the presented approach is simpler in the sense that it does not necessitate to determine the coefficients $c_j^{k,l,m}$ from Eq. 4.29. However, here, the drawback is that there is no direct way as in Eq. 4.31 to quantify the expected errors for a certain choice of measurement operators A_j . Nonetheless, as soon as a set of operators is proven to be appropriate in terms of the expected errors, Eq. 4.39 is a simple and straightforward way to infer the PI part of a quantum state.

Apart from their usefulness for a direct evaluation of the PI part of a quantum state, the coarse grained operators M_j^n are also beneficial for a *scalable* maximum likelihood reconstruction of ϱ_{PI} . Permutationally invariant tomography does not only offer an efficient measurement scheme but also offers the possibility to efficiently store and further process the obtained data on a digital computer. For this purpose, a basis change to the angular momentum basis is required. A possible choice of basis states for the N -qubit Hilbert space is $|j, j_z, \alpha\rangle$, with $\vec{J}^2 |j, j_z, \alpha\rangle = j(j+1) |j, j_z, \alpha\rangle$, $J_z |j, j_z, \alpha\rangle = j_z |j, j_z, \alpha\rangle$. The total spin numbers are restricted to $j = j_{\min}, j_{\min} + 1, \dots, \frac{N}{2}$ starting from $j_{\min} = 0$ for N even and $j_{\min} = \frac{1}{2}$ for N odd, while $j_z = -\frac{N}{2}, -\frac{N}{2} + 1, \dots, \frac{N}{2}$. Here, the label α was introduced to remove the degeneracy of the eigenstates of \vec{J}^2 and J_z^2 with degree d_j [280]. In this basis, any PI operator, like, e.g., ϱ_{PI} or M_j^n , can be written in a simple block diagonal form

$$\varrho_{\text{PI}} = \bigoplus_{k=k_{\min}}^{N/2} \frac{\mathbb{1}_{d_k}}{d_k} \otimes p_k \varrho_k \quad (4.40)$$

4.2. Partial tomography schemes

with ϱ_k being the density operators of the spin- k subspace and p_k a probability distribution, see Fig. 4.5. Hence, as an equally efficient alternative to storing

$$\rho_{\text{PI}} = \left(\begin{array}{c} \boxed{\tilde{\rho}_{N/2}} \\ \mathbb{1}_{d_k} \otimes \tilde{\rho}_k \\ \begin{array}{cc} \boxed{\tilde{\rho}_k} & 0 \\ 0 & \boxed{\tilde{\rho}_k} \end{array} \\ \boxed{\tilde{\rho}_0} \end{array} \right)$$

Figure 4.5.: Every PI state allows for a block diagonal decomposition.

the generalized Bloch vector introduced in Eq. 4.28, it is also possible to store only the $\frac{N}{2}$ blocks $\tilde{\varrho}_j = p_j \varrho_j / d_j$ (of which each has a multiplicity of d_j). The largest of these blocks, which corresponds to the symmetric subspace, is of size $(N+1) \times (N+1)$ and multiplicity $d_{\frac{N}{2}} = 1$. Thus, the block diagonal representation of PI states is an efficient scheme for data storage. Moreover, as also the coarse grained measurement operators M_k^n allow for a block diagonal representation,

$$M_k^n = \bigoplus_{j=j_{\min}}^{N/2} \mathbb{1}_{d_k} \otimes M_{k,j}^n, \quad (4.41)$$

expectation values of the kind $\text{Tr}(\varrho_{\text{PI}} M_j^n)$ can efficiently be calculated as follows

$$\text{Tr}(\varrho_{\text{PI}} M_j^n) = \sum_{k=k_{\min}}^{N/2} p_k \text{Tr}(\varrho_k M_{k,j}^n). \quad (4.42)$$

More precisely, in the standard representation the trace has to be taken over the product of two 2^N -dimensional matrices, now there are only about $\frac{N}{2}$ terms with traces of at most $(N+1)$ -dimensional matrices. Again, the effort reduces from exponential to polynomial. This speedup can be used beneficially to also make maximum likelihood estimation scalable, i.e., the computational effort increases polynomially with the system size. For details see publication P4.3.

Even if the state to be analyzed is not PI, as long as the observable to be measured is PI one can hugely benefit from the scheme thanks to the fact that a similarly scalable decomposition can be found for any PI operator O , i.e., $O = \bigoplus_j \mathbb{1}_{d_j} \otimes O_j$. Together with Eq. (4.40) this yields an efficient way to calculate the expectation values $\langle O \rangle = \text{Tr}(\varrho O) = \sum_j p_j \text{Tr}(\varrho_j O_j)$ also for non-PI states. Please note that if the operator whose expectation value is to be determined is PI, the result obtained by means of permutationally invariant tomography is correct, even in case the state under investigation is not PI.

4.2.3. Compressed sensing in the permutationally invariant subspace

In sections 4.2.1 and 4.2.2, two efficient tomography schemes were presented. The first one, CS, is suited for low rank states, i.e., states with only few non-zero eigenvalues. The second scheme, PI tomography, is ideal for the analysis of states that lie within the PI subspace. At this point, the question arises if it is also possible to combine these two schemes as there are states that are at the same time of low rank and PI. Prominent examples for such states are GHZ states and the symmetric Dicke states as discussed in this work. As expected, the answer turns out to be positive. More precisely, in order to carry out CS in the PI subspace, only a subset of the set of operators $A_j^{\otimes N}$ ($j = 1, \dots, D_N$) is measured. Then, the target function 4.14 is minimized under the constraint of ϱ being PI and positive-semidefinite. This scheme is experimentally implemented in publication P4.3 on the example of a symmetric six-photon Dicke state. There, it could be shown that CS in the PI subspace can further reduce the experimental without significantly changing the reconstructed state. For more details see publication P4.2 and for a description of the fitting algorithm, see publication P4.3.

4.3. Publications

P4.1

PRL 105, 250403 (2010)

PHYSICAL REVIEW LETTERS

week ending
17 DECEMBER 2010

Permutationally Invariant Quantum Tomography

G. Tóth,^{1,2,3} W. Wieczorek,^{4,5,*} D. Gross,⁶ R. Krischek,^{4,5} C. Schwemmer,^{4,5} and H. Weinfurter^{4,5}¹*Department of Theoretical Physics, The University of the Basque Country, P.O. Box 644, E-48080 Bilbao, Spain*²*IKERBASQUE, Basque Foundation for Science, E-48011 Bilbao, Spain*³*Research Institute for Solid State Physics and Optics, Hungarian Academy of Sciences, P.O. Box 49, H-1525 Budapest, Hungary*⁴*Max-Planck-Institut für Quantenoptik, Hans-Kopfermann-Strasse 1, D-85748 Garching, Germany*⁵*Fakultät für Physik, Ludwig-Maximilians-Universität, D-80797 München, Germany*⁶*Institute for Theoretical Physics, Leibniz University Hannover, D-30167 Hannover, Germany*

(Received 4 June 2010; revised manuscript received 30 August 2010; published 16 December 2010)

We present a scalable method for the tomography of large multiqubit quantum registers. It acquires information about the permutationally invariant part of the density operator, which is a good approximation to the true state in many relevant cases. Our method gives the best measurement strategy to minimize the experimental effort as well as the uncertainties of the reconstructed density matrix. We apply our method to the experimental tomography of a photonic four-qubit symmetric Dicke state.

DOI: 10.1103/PhysRevLett.105.250403

PACS numbers: 03.65.Wj, 03.65.Ud, 42.50.Dv

Because of the rapid development of quantum experiments, it is now possible to create highly entangled multiqubit states using photons [1–5], trapped ions [6], and cold atoms [7]. So far, the largest implementations that allow for an individual readout of the particles involve on the order of 10 qubits. This number will soon be overcome, for example, by using several degrees of freedom within each particle to store quantum information [8]. Thus, a new regime will be reached in which a complete state tomography is impossible even from the point of view of the storage space needed on a classical computer. At this point the question arises: Can we still extract useful information about the quantum state created?

In this Letter we propose permutationally invariant (PI) tomography in multiqubit quantum experiments [9]. Concretely, instead of the density matrix ρ , we propose to determine the PI part of the density matrix defined as

$$\rho_{\text{PI}} = \frac{1}{N!} \sum_k \Pi_k \rho \Pi_k, \quad (1)$$

where Π_k are all the permutations of the qubits. Reconstructing ρ_{PI} has been considered theoretically for spin systems (see, e.g., Ref. [10]). Recently it has been pointed out that photons in a single mode optical fiber will always be in a PI state and that there is only a small set of measurements needed for their characterization [11,12].

Here, we develop a provably optimal scheme, which is feasible for large multiqubit systems: For our method, the measurement effort increases only quadratically with the size of the system. Our approach is further motivated by the fact that almost all multipartite experiments are done with PI quantum states [2–4,6]. Thus, the density matrix obtained from PI tomography is expected to be close to the one of the experimentally achieved state. The expectation values of symmetric operators, such as some entanglement witnesses, and fidelities with respect to symmetric states are the same

for both density matrices and are thus obtained exactly from PI tomography [2–4]. Finally, if ρ_{PI} is entangled, so is the state ρ of the system, which makes PI tomography a useful and efficient tool for entanglement detection.

Below, we summarize the four main contributions of this Letter. We restrict our attention to the case of N qubits—higher-dimensional systems can be treated similarly.

(1) In most experiments, the qubits can be individually addressed whereas nonlocal quantities cannot be measured directly. The experimental effort is then characterized by the number of local measurement settings needed, where “setting” refers to the choice of one observable per qubit, and repeated von Neumann measurements in the observables’ eigenbases [13]. Here, we compute the minimal number of measurement settings required to recover ρ_{PI} .

(2) The requirement that the number of settings be minimal does not uniquely specify the tomographic protocol. On the one hand, there are infinitely many possible choices for the local settings that are both minimal and give sufficient information to find ρ_{PI} . On the other hand, for each given setting, there are many ways of estimating the unknown density operator from the collected data. We present a systematic method to find the optimal scheme through statistical error analysis.

(3) Next, we turn to the important problem of gauging the information loss incurred due to restricting attention to the PI part of the density matrix. We describe an easy test measurement that can be used to judge the applicability of PI tomography before it is implemented.

(4) Finally, we demonstrate that these techniques are viable in practice by applying them to a photonic experiment observing a four-qubit symmetric Dicke state.

Minimizing the number of settings.—We will now present our first main result.

Observation 1. For a system of N qubits, permutationally invariant tomography can be performed with

4. Quantum state tomography

$$\mathcal{D}_N = \binom{N+2}{N} = \frac{1}{2}(N^2 + 3N + 2) \quad (2)$$

local settings. It is not possible to perform such a tomography with fewer settings.

Proof.—First, we need to understand the information obtainable from a single measurement setting. We assume that for every given setting, the same basis is measured at every site [14]. Measuring a local basis $\{|\phi_1\rangle, |\phi_2\rangle\}$ is equivalent to estimating the expectation value of the traceless operator $A = |\phi_1\rangle\langle\phi_1| - |\phi_2\rangle\langle\phi_2|$. Merely by measuring $A^{\otimes N}$, it is possible to obtain all the N expectation values

$$\langle(A^{\otimes(N-n)} \otimes \mathbb{1}^{\otimes n})_{\text{PI}}\rangle, \quad (n = 0, \dots, N-1), \quad (3)$$

and, conversely, that is all the information obtainable about \mathcal{Q}_{PI} from a single setting.

Next, we will use the fact that any PI density operator can be written as a linear combination of the pairwise orthogonal operators $(X^{\otimes k} \otimes Y^{\otimes l} \otimes Z^{\otimes m} \otimes \mathbb{1}^{\otimes n})_{\text{PI}}$, where X , Y , and Z are the Pauli matrices. We consider the space spanned by these operators for one specific value of n . Simple counting shows that its dimension is $\mathcal{D}_{(N-n)}$. The same space is spanned by $\mathcal{D}_{(N-n)}$ generic operators of the type $(A_j^{\otimes(N-n)} \otimes \mathbb{1}^{\otimes n})_{\text{PI}}$. We draw two conclusions: First, any setting gives at most one expectation value for every such space. Hence the number of settings cannot be smaller than the largest dimension, which is \mathcal{D}_N . Second, a generic choice of \mathcal{D}_N settings is sufficient to recover the correlations in each of these spaces, and hence completely characterizes \mathcal{Q}_{PI} . This concludes the proof [15].

The proof implies that there are real coefficients $c_j^{(k,l,m)}$ such that

$$\begin{aligned} &\langle(X^{\otimes k} \otimes Y^{\otimes l} \otimes Z^{\otimes m} \otimes \mathbb{1}^{\otimes n})_{\text{PI}}\rangle \\ &= \sum_{j=1}^{\mathcal{D}_N} c_j^{(k,l,m)} \langle(A_j^{\otimes(N-n)} \otimes \mathbb{1}^{\otimes n})_{\text{PI}}\rangle. \end{aligned} \quad (4)$$

We will refer to the numbers on the left-hand side of Eq. (4) as the elements of the generalized Bloch vector. The expectation values on the right-hand side can be obtained by measuring the settings with A_j for $j = 1, 2, \dots, \mathcal{D}_N$.

Minimizing uncertainties.—We now have to determine the optimal scheme for PI tomography. To this end, we define our measure of statistical uncertainty as the sum of the variances of all the Bloch vector elements

$$\begin{aligned} (\mathcal{E}_{\text{total}})^2 &= \sum_{k+l+m+n=N} \mathcal{E}^2[(X^{\otimes k} \otimes Y^{\otimes l} \otimes Z^{\otimes m} \otimes \mathbb{1}^{\otimes n})_{\text{PI}}] \\ &\times \left(\frac{N!}{k!l!m!n!} \right), \end{aligned} \quad (5)$$

where the term with the factorials is the number of different permutations of $X^{\otimes k} \otimes Y^{\otimes l} \otimes Z^{\otimes m} \otimes \mathbb{1}^{\otimes n}$. Based on Eq. (4), the variance of a single Bloch vector element is

$$\begin{aligned} &\mathcal{E}^2[(X^{\otimes k} \otimes Y^{\otimes l} \otimes Z^{\otimes m} \otimes \mathbb{1}^{\otimes n})_{\text{PI}}] \\ &= \sum_{j=1}^{\mathcal{D}_N} |c_j^{(k,l,m)}|^2 \mathcal{E}^2[(A_j^{\otimes(N-n)} \otimes \mathbb{1}^{\otimes n})_{\text{PI}}]. \end{aligned} \quad (6)$$

Equation (5) can be minimized by changing the A_j matrices and the $c_j^{(k,l,m)}$ coefficients. We consider the coefficients first.

For any Bloch vector element, finding $c_j^{(k,l,m)}$'s that minimize the variance Eq. (6) subject to the constraint that equality holds in Eq. (4) is a least squares problem. It has an analytic solution obtained as follows: Write the operator on the left-hand side of Eq. (6) as a vector \vec{v} (with respect to some basis). Likewise, write the operators on the right-hand side as \vec{v}_j ; and define a matrix $V = [\vec{v}_1, \vec{v}_2, \dots, \vec{v}_{\mathcal{D}_N}]$. Then Eq. (4) can be cast into the form $\vec{v} = V\vec{c}$, where \vec{c} is a vector of the $c_j^{(k,l,m)}$ values for given (k, l, m) . If E is the diagonal matrix with entries $E_{j,j}^2 = \mathcal{E}^2[(A_j^{\otimes(N-n)} \otimes \mathbb{1}^{\otimes n})_{\text{PI}}]$, then the optimal solution is $\vec{c} = E^{-2}V^T(VE^{-2}V^T)^{-1}\vec{v}$, where the inverse is taken over the range [16].

Equipped with a method for obtaining the optimal $c_j^{(k,l,m)}$'s for every fixed set of observables A_j , it remains to find the best settings to measure. Every qubit observable can be defined by the measurement directions \vec{a}_j using $A_j = a_{j,x}X + a_{j,y}Y + a_{j,z}Z$. Thus, the task is to identify \mathcal{D}_N measurement directions on the Bloch sphere minimizing the variance. In general, finding the globally optimal solution of high-dimensional problems is difficult. In our case, however, $\mathcal{E}_{\text{total}}$ seems to penalize an inhomogeneous distribution of the \vec{a}_j vectors; thus, using evenly distributed vectors as an initial guess, usual minimization procedures can be used to decrease $\mathcal{E}_{\text{total}}$ and obtain satisfactory results [16].

The variance $\mathcal{E}^2[(A_j^{\otimes(N-n)} \otimes \mathbb{1}^{\otimes n})_{\text{PI}}]$ of the observed quantities depends on the physical implementation. In the photonic setup below, we assume Poissonian distributed counts. It follows that (see also Refs. [17,18])

$$\mathcal{E}^2[(A_j^{\otimes(N-n)} \otimes \mathbb{1}^{\otimes n})_{\text{PI}}] = \frac{[\Delta(A_j^{\otimes(N-n)} \otimes \mathbb{1}^{\otimes n})_{\text{PI}}]_{\mathcal{Q}_0}^2}{\lambda_j - 1}, \quad (7)$$

where $(\Delta A)_{\mathcal{Q}}^2 = \langle A^2 \rangle_{\mathcal{Q}} - \langle A \rangle_{\mathcal{Q}}^2$, \mathcal{Q}_0 is the state of the system, and λ_j is the parameter of the Poissonian distribution, which equals the expected value of the total number of counts for the setting j . The variance depends on the unknown state. If we have preliminary knowledge of the likely form of \mathcal{Q}_0 , we should use that information in the optimization. Otherwise, \mathcal{Q}_0 can be set to the completely mixed state. For the latter, straightforward calculation shows that $\mathcal{E}^2[(A_j^{\otimes(N-n)} \otimes \mathbb{1}^{\otimes n})_{\text{PI}}] = \binom{N}{n}^{-1}/(\lambda_j - 1)$. For another implementation, such as trapped ions, our scheme for PI tomography can be used after replacing Eq. (7) by a formula giving the variance for that implementation.

Estimating the information loss due to symmetrization.—It is important to know how close the PI quantum state is to the state of the system as PI tomography should

serve as an alternative of full state tomography for experiments aiming at the preparation of PI states.

Observation 2. The fidelity between the original state and the permutationally invariant state, $F(\varrho, \varrho_{\text{PI}})$, can be estimated from below as $F(\varrho, \varrho_{\text{PI}}) \geq \langle P_s \rangle_{\varrho}^2$, where $P_s = \sum_{n=0}^N |D_N^{(n)}\rangle\langle D_N^{(n)}|$ is the projector to the N -qubit symmetric subspace, and the symmetric Dicke state is defined as $|D_N^{(n)}\rangle = \binom{N}{n}^{-1/2} \sum_k \mathcal{P}_k(|0\rangle^{\otimes(N-n)} \otimes |1\rangle^{\otimes n})$, where the summation is over all the different permutations of the qubits. Observation 2 can be proved based on Ref. [19] and elementary matrix manipulations. Note that Observation 2 makes it possible to estimate $F(\varrho, \varrho_{\text{PI}})$ based on knowing only ϱ_{PI} .

Lower bounds on the fidelity to symmetric Dicke states, i.e., $\text{Tr}(|D_N^{(n)}\rangle\langle D_N^{(n)}| \varrho)$ can efficiently be obtained by measuring X , Y , and Z on all qubits, i.e., measuring only three local settings independent of N [20]. With the same measurements, one can also obtain a lower bound on the overlap between the state and the symmetric subspace. For four qubits, this can be done based on $P_s \geq [(J_x^4 + J_y^4 + J_z^4) - (J_x^2 + J_y^2 + J_z^2)]/18$, where $J_x = (1/2)\sum_k X_k$, $J_y = (1/2)\sum_k Y_k$, etc. Operators for estimating $\langle P_s \rangle$ for $N = 6, 8$ are given in Ref. [16]. This allows one to judge how suitable the quantum state is for PI tomography before such a tomography is carried out.

Experimental results.—We demonstrate the method and the benefits of our algorithm for PI tomography for a four-qubit symmetric Dicke state with two excitations $|D_4^{(2)}\rangle$. First, we optimize the \vec{a}_j 's and the $c_j^{(k,l,m)}$'s for $\varrho_0 = \mathbb{1}/16$ and only for the uncertainty of full four-qubit correlation terms, which means that when computing $\mathcal{E}_{\text{total}}$, we carry out the summation in Eq. (5) only for the terms with $n = 0$. With simple numerical optimization, we were looking for the set of A_j basis matrices that minimize the uncertainty of the full correlation terms. Then, we also looked for the basis matrices that minimize the sum of the squared error of all the Bloch vector elements and considered also density matrices different from white noise, such as a pure Dicke state mixed with noise. We find that the gain in terms of decreasing the uncertainties is negligible in our case and that it is sufficient to optimize for $\varrho_0 = \mathbb{1}/16$ and for the full correlation terms. To demonstrate the benefits of the optimization of the measurement directions, we also compare the results with those obtained with randomly distributed basis matrices.

The Dicke state was observed in a photonic system. Essentially, four photons emitted by the second-order collinear type-II spontaneous parametric down-conversion process were symmetrically distributed into four spatial modes. Upon detection of one photon from each of the outputs, the state $|D_4^{(2)}\rangle$ is observed. Polarization analysis in each mode is used to characterize the experimentally observed state. We collected data for each setting for 5 min, with an average count rate of 410 per minute. The experimental setup has been described in detail in Refs. [2,3].

First, to check the applicability of the PI tomography, we apply our tools described above requiring only the

measurement of the three settings, $X^{\otimes 4}$, $Y^{\otimes 4}$, and $Z^{\otimes 4}$. We determine the expectation value of the projector to the symmetric subspace, yielding $\langle P_s \rangle \geq 0.905 \pm 0.015$. Based on Observation 2, we obtain $F(\varrho, \varrho_{\text{PI}}) \geq 0.819 \pm 0.028$. These results show that the state is close to be PI and has a large overlap with the symmetric subspace. Thus, it makes sense to apply PI tomography.

For PI tomography of a four-qubit system, the measurement of 15 settings is needed. We used Eq. (4) to obtain the Bloch vector elements from the experimentally measured quantities. This way, we could obtain all the 34 symmetric correlations of the form $(X^{\otimes k} \otimes Y^{\otimes l} \otimes Z^{\otimes m} \otimes \mathbb{1}^{\otimes n})_{\text{PI}}$. In Fig. 1, we give the values of the correlations for optimized and for randomly chosen measurement directions, compared to the results obtained from full tomography, which needed 81 measurement settings. As can be seen in Fig. 1, the uncertainty for the optimized settings is considerably smaller than the one for the randomly chosen settings. Moreover, the results from the optimized settings fit very well the results of the full tomography. In Fig. 2, we compare the density matrices obtained from full tomography [Fig. 2(a)], from PI tomography for optimized [Fig. 2(b)], and for random measurement directions [Fig. 2(c)]. Because of noise, the fidelity of the result of the full tomography with respect to $|D_4^{(2)}\rangle$ is 0.873 ± 0.005 , which is similar to the fidelity of the results of the PI tomography with optimized settings, 0.852 ± 0.009 [21]. In contrast, for the method using random measurement directions, the fidelity is 0.814 ± 0.059 , for which the

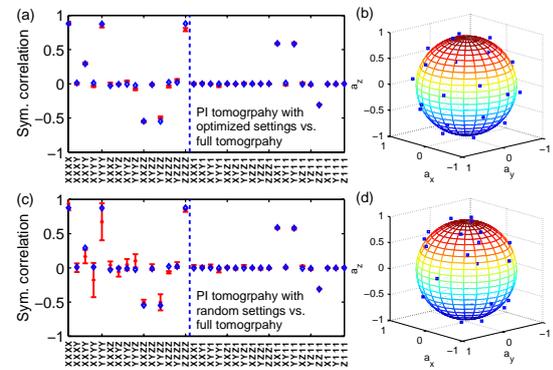


FIG. 1 (color online). (a) Comparison of the 34 symmetrized correlations coming from (crosses with error bars) 15 permutationally invariant measurement settings with optimized A_j matrices for $N = 4$ qubits and (diamonds) from full tomography requiring 81 local settings. The average uncertainty of all symmetrized correlations obtained from full tomography is ± 0.022 , and is not shown in the figure. The labels refer to symmetrized correlations of the form given in the left-hand side of Eq. (4). The results corresponding to the 15 full four-qubit correlations are left from the vertical dashed line. (b) Measurement directions. A point at (a_x, a_y, a_z) corresponds to measuring operator $a_x X + a_y Y + a_z Z$. (c) Results for randomly chosen A_j matrices and (d) corresponding measurement directions.

4. Quantum state tomography

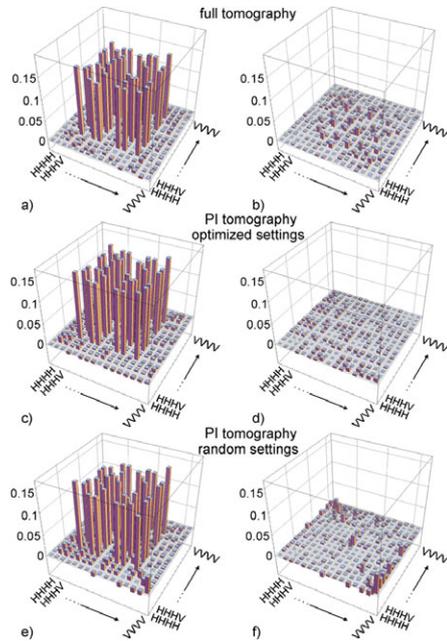


FIG. 2 (color online). (a) The real and (b) imaginary parts of the density matrix coming from full tomography. (c),(d) The same for permutationally invariant tomography with optimized and (e),(f) random measurement directions, respectively.

uncertainty is the largest compared to all previous fidelity values. Finally, we also computed the fidelity of the results with respect to the PI density matrix obtained from full tomography [22]. The results of the PI tomography with optimized settings show a good agreement with full tomography: the fidelity is 0.947, which is quite close to the fidelity between the results of full tomography and its PI part, 0.964. On the other hand, for the PI tomography with random settings the corresponding fidelity is much lower, 0.880. Overall, the PI tomography shows a good agreement with the full tomography for this particular experiment. However, a reasonable choice of measurement directions is needed to obtain uncertainties in the reconstructed Bloch vector elements similar to the ones from full tomography.

Finally, let us comment on how our method can be extended to larger systems. Permutationally invariant operators can be represented efficiently on a digital computer in the basis of $(X^{\otimes k} \otimes Y^{\otimes l} \otimes Z^{\otimes m} \otimes \mathbb{1}^{\otimes n})_{PI}$ operators. We determined the optimal A_j operators for PI tomography for systems with $N = 6, 8, \dots, 14$ qubits. To have the same maximum uncertainty of the Bloch vector elements as for the $N = 4$ case, one has to increase the counts per setting by less than 50% [16].

In summary, we presented a scalable method for permutationally invariant tomography, which can be used in place of full state tomography in experiments that aim at preparing permutationally invariant many-qubit states. For our approach, the same operator has to be measured on all qubits,

which is a clear advantage in some experiments. We showed how to choose the measurements such that the uncertainty in the reconstructed density matrix is the smallest possible. This paves the way of characterizing permutationally invariant states of many qubits in various physical systems. Moreover, this work also shows that, given some knowledge or justifiable assumptions, there is a way to obtain scalable state tomography for multiqubit entangled states.

We thank D. Hayes and N. Kiesel for discussions. We thank the Spanish MEC (Consolider-Ingenio 2010 project "QOIT," Project No. FIS2009-12773-C02-02), the Basque Government (Project No. IT4720-10), the ERC StG GEDENTQOPT, the DFG-Cluster of Excellence MAP, the EU projects QAP, Q-Essence, and CORNER, and the DAAD/MNISW for support. W.W. and C.S. thank the QCCC of the Elite Network of Bavaria for support.

*Present address: Faculty of Physics, University of Vienna, Boltzmanngasse 5, A-1090 Wien, Austria.

- [1] J.-W. Pan *et al.*, *Nature (London)* **403**, 515 (2000); M. Bourennane *et al.*, *Phys. Rev. Lett.* **92**, 087902 (2004); N. Kiesel *et al.*, *ibid.* **95**, 210502 (2005).
- [2] N. Kiesel *et al.*, *Phys. Rev. Lett.* **98**, 063604 (2007).
- [3] W. Wieczorek *et al.*, *Phys. Rev. Lett.* **103**, 020504 (2009); R. Krischek *et al.*, *Nat. Photon.* **4**, 170 (2010).
- [4] R. Prevedel *et al.*, *Phys. Rev. Lett.* **103**, 020503 (2009).
- [5] W. Wieczorek *et al.*, *Phys. Rev. Lett.* **101**, 010503 (2008).
- [6] C. A. Sackett *et al.*, *Nature (London)* **404**, 256 (2000); H. Häffner *et al.*, *Nature (London)* **438**, 643 (2005).
- [7] O. Mandel *et al.*, *Nature (London)* **425**, 937 (2003).
- [8] C. Cinelli *et al.*, *Phys. Rev. Lett.* **95**, 240405 (2005); G. Vallone *et al.*, *ibid.* **98**, 180502 (2007); W.-B. Gao *et al.*, *Nature Phys.* **6**, 331 (2010).
- [9] For other approaches see D. Gross *et al.*, *Phys. Rev. Lett.* **105**, 150401 (2010); M. Cramer and M.B. Plenio, *arXiv:1002.3780*; S.T. Flammia *et al.*, *arXiv:1002.3839*; O. Landon-Cardinal *et al.*, *arXiv:1002.4632*.
- [10] G. M. D'Ariano *et al.*, *J. Opt. B* **5**, 77 (2003).
- [11] R. B. A. Adamson *et al.*, *Phys. Rev. Lett.* **98**, 043601 (2007).
- [12] L. K. Shalm *et al.*, *Nature (London)* **457**, 67 (2009).
- [13] O. Gühne and G. Tóth, *Phys. Rep.* **474**, 1 (2009).
- [14] Otherwise more than \mathcal{D}_N settings are necessary [16].
- [15] This is connected to a general idea: It is expected that the determination of an operator within a subspace whose dimension depends polynomially on N needs a number of settings increasing also polynomially with N .
- [16] See supplementary material at <http://link.aps.org/supplemental/10.1103/PhysRevLett.105.250403> for additional derivations and experimental results.
- [17] C. Schmid, Ph.D. thesis, Ludwig-Maximilian-Universität, 2008; D. F. V. James *et al.*, *Phys. Rev. A* **64**, 052312 (2001).
- [18] B. Jungnitsch *et al.*, *Phys. Rev. Lett.* **104**, 210401 (2010).
- [19] J. A. Miszczak *et al.*, *Quantum Inf. Comput.* **9**, 0103 (2009).
- [20] G. Tóth *et al.*, *New J. Phys.* **11**, 083002 (2009).
- [21] Expectation values are obtained directly from the measured data, rather than from ρ_{PI} .
- [22] Values without error are deduced from fitted matrices obtained via maximum likelihood estimation [17].

Permutationally Invariant Quantum Tomography - Supplementary Material

 Géza Tóth,^{1,2,3} Witłef Wieczorek,^{4,5,*} David Gross,⁶ Roland Krischek,^{4,5} Christian Schwemmer,^{4,5} and Harald Weinfurter^{4,5}
¹Department of Theoretical Physics, The University of the Basque Country, P.O. Box 644, E-48080 Bilbao, Spain

²IKERBASQUE, Basque Foundation for Science, E-48011 Bilbao, Spain

³Research Institute for Solid State Physics and Optics,

Hungarian Academy of Sciences, P.O. Box 49, H-1525 Budapest, Hungary

⁴Max-Planck-Institut für Quantenoptik, Hans-Kopfermann-Strasse 1, D-85748 Garching, Germany

⁵Fakultät für Physik, Ludwig-Maximilians-Universität, D-80797 München, Germany

⁶Institute for Theoretical Physics, Leibniz University Hannover, D-30167 Hannover, Germany

(Dated: December 23, 2010)

The supplement contains some derivations to help to understand the details of the proofs of the main text. It also contains some additional experimental results.

PACS numbers: 03.65.Wj, 42.50.Dv, 03.65.Ud

Proof of that we have to measure the same operator on all qubits. From the proof of Observation 1, we know that at least \mathcal{D}_N measurements are needed to get the expectation values of all the \mathcal{D}_N independent symmetric full N -particle correlations. What if we measure \mathcal{D}_N settings, but several of them are not $\{A_j, A_j, \dots, A_j\}$ -type, but $\{A_j^{(1)}, A_j^{(2)}, \dots, A_j^{(N)}\}$ -type, i.e., we do not measure the same operator on all qubits? Each setting makes it possible to get a single operator containing full N -qubit correlations. Let us denote this operator by M_k for $k = 1, 2, \dots, \mathcal{D}_N$. Then, we know the expectation value of any operator of the space defined by the M_k operators. However, not all M_k 's are permutationally invariant. Thus, the size of the PI subspace of the space of the M_k operators is less than \mathcal{D}_N . We do not have \mathcal{D}_N linearly independent symmetric operators in this space. Thus, \mathcal{D}_N measurement settings are sufficient to measure ϱ_{PI} only if we have settings of the type $\{A_j, A_j, \dots, A_j\}$.

Derivation of Eq. (7). The eigen-decomposition of the correlation term is

$$(A_j^{\otimes(N-n)} \otimes \mathbf{1}^{\otimes n})_{\text{PI}} = \sum_k \Lambda_{j,n,k} |\Phi_{j,k}\rangle \langle \Phi_{j,k}|. \quad (\text{S1})$$

The individual counts $N_C(A_j)_k$ follow a Poissonian distribution $f(n_c, \lambda_{j,k})$, where $\lambda_{j,k}$ are the parameters of the Poissonian distributions and $\sum_k \lambda_{j,k} = \lambda_j$. The conditional variance, knowing that the total count is $N_C(A_j)$, is

$$\mathcal{E}^2[(A_j^{\otimes(N-n)} \otimes \mathbf{1}^{\otimes n})_{\text{PI}} | N_C(A_j)] = \frac{[\Delta(A_j^{\otimes(N-n)} \otimes \mathbf{1}^{\otimes n})_{\text{PI}}]^2}{N_C(A_j)}. \quad (\text{S2})$$

After straightforward algebra, the variance is obtained as

$$\begin{aligned} \mathcal{E}^2[(A_j^{\otimes(N-n)} \otimes \mathbf{1}^{\otimes n})_{\text{PI}}] &= \sum_m f(m, \lambda_j) \mathcal{E}^2[(A_j^{\otimes(N-n)} \otimes \mathbf{1}^{\otimes n})_{\text{PI}} | N_C(A_j) = m] \\ &= \frac{[\Delta(A_j^{\otimes(N-n)} \otimes \mathbf{1}^{\otimes n})_{\text{PI}}]^2}{\lambda_j - 1}. \end{aligned} \quad (\text{S3})$$

Similar results can be obtained through assuming Poissonian measurement statistics and Gaussian error propagation

[S1, S2]. If $\varrho_0 = \mathbf{1}/2^N$, then $\Delta(A_j^{\otimes(N-n)} \otimes \mathbf{1}^{\otimes n})_{\text{PI}}$ is independent from the choice of A_j . By substituting $A_j = Z$, straightforward calculations gives

$$\mathcal{E}^2[(A_j^{\otimes(N-n)} \otimes \mathbf{1}^{\otimes n})_{\text{PI}}] = \frac{\binom{N}{n}^{-1}}{\lambda_j - 1}. \quad (\text{S4})$$

Obtaining the formula for $c_j^{(k,l,m)}$ for the smallest error.

We look for $c_j^{(k,l,m)}$ for which the squared uncertainty given in Eq. (6) is the smallest. In the following, we use the definition given in the main text for \vec{c} , \vec{v} , V and E . Thus, V is matrix mapping a large space \mathbb{R}^l to a small space \mathbb{R}^s . Let E be a non-singular diagonal matrix in the small space. We have to solve

$$\min_{\vec{c}} \|E\vec{c}\|^2 \quad \text{s.t.} \quad V\vec{c} = \vec{v}, \quad (\text{S5})$$

where $\|\vec{a}\|$ is the Euclidean norm of \vec{a} . Using Lagrangian multipliers, we write down the condition for a minimum fulfilling the constraints $V\vec{c} = \vec{v}$

$$\nabla_{\vec{c}} \left\{ \vec{c}^T E^2 \vec{c} + \sum_{i=1}^s \lambda_i [(V\vec{c})_i - v_i] \right\} = 0. \quad (\text{S6})$$

Hence, the condition for a local (and, due to convexity, global) minimum is

$$\vec{c} = \frac{1}{2} E^{-2} V^T \vec{\lambda}, \quad (\text{S7})$$

where $\lambda \in \mathbb{R}^s$ is the vector of multipliers. In other words, we have a minimum if and only if $\vec{c} \in \text{range } E^{-2} V^T$. Because the range of V^T is an s -dimensional subspace in \mathbb{R}^l , there is a *unique* \vec{c} in that range such that $V\vec{c} = \vec{v}$. A solution in a closed form can be obtained as

$$\vec{c} = E^{-2} V^T (V E^{-2} V^T)^{-1} \vec{v}. \quad (\text{S8})$$

Simple calculation shows that the $V\vec{c} = \vec{v}$ condition holds

$$V\vec{c} = V E^{-2} V^T (V E^{-2} V^T)^{-1} \vec{v} = \vec{v}. \quad (\text{S9})$$

4. Quantum state tomography

2

Table S1: Fidelities to the 4-qubit Dicke states.

measurement	$ D_4^{(0)}\rangle$	$ D_4^{(1)}\rangle$	$ D_4^{(2)}\rangle$	$ D_4^{(3)}\rangle$	$ D_4^{(4)}\rangle$	Σ
full tomography	-0.001 ± 0.002	0.023 ± 0.004	0.873 ± 0.005	0.026 ± 0.004	0.002 ± 0.002	0.922
full tomography (max-like)	0.001	0.021	0.869	0.023	0	0.914
PI tomography	-0.001 ± 0.002	0.040 ± 0.007	0.852 ± 0.009	0.036 ± 0.007	-0.002 ± 0.002	0.925
PI tomography (max-like)	0.003	0.038	0.850	0.037	0	0.928
PI tomography (ran)	0.000 ± 0.002	0.055 ± 0.027	0.814 ± 0.059	0.023 ± 0.027	0.001 ± 0.002	0.893
PI tomography (ran,max-like)	0.004	0.050	0.816	0.020	0.007	0.897

Proof of Observation 2. The eigenstates of $\vec{J}^2 = J_x^2 + J_y^2 + J_z^2$ are usually labelled by $|j, m, \alpha\rangle$, where $\vec{J}^2|j, m, \alpha\rangle = j(j+1)|j, m, \alpha\rangle$, $J_z|j, m, \alpha\rangle = m|j, m, \alpha\rangle$, and α is used to label the different eigenstates having the same j and m [S3]. Let $P_{j,\alpha}$ denote the projector to the subspace of a given j and α . The number of subspaces is denoted by N_{SS} , and, for a given N , it can be calculated from group theory. Moreover, $P_s \equiv P_{N/2,1}$. Using this notation, $\varrho_{PI} = \sum_{j,\alpha} P_{j,\alpha} \varrho P_{j,\alpha} = (P_s \varrho P_s) + \sum_{j < N/2, \alpha} (P_{j,\alpha} \varrho P_{j,\alpha})$. In the basis of \vec{J}^2 eigenstates, ϱ_{PI} can be written as a block diagonal matrix

$$\varrho_{PI} = \bigoplus_{j,\alpha} (\langle P_{j,\alpha} \rangle \hat{\varrho}_{j,\alpha}), \quad (S10)$$

where $\hat{\varrho}_{j,\alpha}$ are density matrices of size $(2j+1) \times (2j+1)$. In another context,

$$\varrho_{PI} = \sum_{j,\alpha} \langle P_{j,\alpha} \rangle \varrho_{j,\alpha}, \quad (S11)$$

where $\varrho_{j,\alpha} = P_{j,\alpha} \varrho P_{j,\alpha} / \text{Tr}(P_{j,\alpha} \varrho P_{j,\alpha})$. Based on that, we obtain

$$F(\varrho, \varrho_{j,\alpha}) = \langle P_{j,\alpha} \rangle \varrho. \quad (S12)$$

Then, due to the separate concavity of the fidelity, i.e., $F(\varrho, p_1 \varrho_1 + p_2 \varrho_2) \geq p_1 F(\varrho, \varrho_1) + p_2 F(\varrho, \varrho_2)$, we obtain $F(\varrho, \varrho_{PI}) \geq \langle P_s \rangle \varrho F(\varrho, \varrho_s) + \sum_{j < N/2, \alpha} \langle P_{j,\alpha} \rangle \varrho F(\varrho, \varrho_{j,\alpha})$. Substituting Eq. (S12) into this inequality, we obtain $F(\varrho, \varrho_{PI}) \geq \langle P_s \rangle \varrho^2 + \sum_{j < N/2, \alpha} \langle P_{j,\alpha} \rangle \varrho^2$. Using the fact that $\langle P_s \rangle \varrho + \sum_{j < N/2, \alpha} \langle P_{j,\alpha} \rangle \varrho = 1$, we obtain

$$F(\varrho, \varrho_{PI}) \geq \langle P_s \rangle \varrho^2 + \frac{(1 - \langle P_s \rangle \varrho)^2}{N_{SS} - 1}. \quad (S13)$$

In many practical situations, the state ϱ is almost symmetric and N is large. In such cases the second term in Eq. (S13) is negligible. Thus, a somewhat weaker bound presented in Observation 2 can be used.

Numerical optimization used to minimize $\mathcal{E}_{\text{total}}$. The measurement directions minimizing $\mathcal{E}_{\text{total}}$ can be obtained as follows. Let us represent the measurement directions by three-dimensional vectors $\{\vec{a}_j\}_{j=1}^{D_N}$. The operators can be obtained as $A_j = a_{j,x}X + a_{j,y}Y + a_{j,z}Z$.

First, we need an initial guess. This can come from a set of randomly chosen vectors representing the measurement directions. One can also use the result of a minimization for

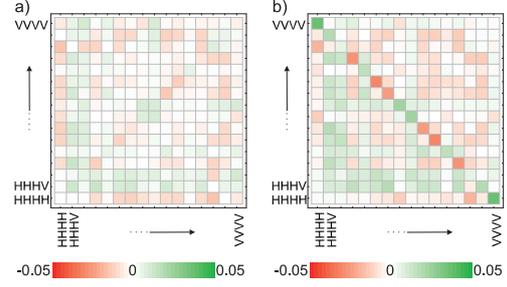


Figure S1: (a) The difference of the real part of the density matrices from optimized settings and the one of full tomography. (b) The difference of the density matrices from random settings and the one of full tomography. For the former, no clear structure is observed, whereas for the latter the largest difference is observed for the anti-diagonal elements.

some measure that characterizes how equally the vectors are distributed. Such a measure is defined by

$$\mathcal{F}(\{v_j\}) = \sum_{k,l} (\vec{v}_k \cdot \vec{v}_l)^{2m}, \quad (S14)$$

where \vec{v}_k represent the measurement directions and \cdot is the scalar product and m is an integer. Such cost functions, called frame potentials, appear in the theory of t -designs essentially for the same purpose.

After we obtain the initial guess from such a procedure, we start an optimization for decreasing $\mathcal{E}_{\text{total}}$. At each iteration of the method, we change the measurement directions by rotating them with a small random angle around a randomly chosen axis. If the change decreases $\mathcal{E}_{\text{total}}$, then we keep the new measurement directions, while if it does not then we discard it. We repeat this procedure until $\mathcal{E}_{\text{total}}$ does not change significantly.

Three-setting witness for estimating the fidelity The three-setting witness for detecting genuine multipartite entanglement in the vicinity of the Dicke state is [S4]

$$\mathcal{W}_{D(4,2)}^{(P3)} = 2 \cdot \mathbb{1} + \frac{1}{6}(J_x^2 + J_y^2 - J_x^4 - J_y^4) + \frac{3}{12}J_z^2 - \frac{7}{12}J_z^4. \quad (S15)$$

For this witness we have [S4]

$$\mathcal{W}_{D(4,2)}^{(P3)} - 3\mathcal{W}_{D(4,2)}^{(P)} \geq 0, \quad (S16)$$

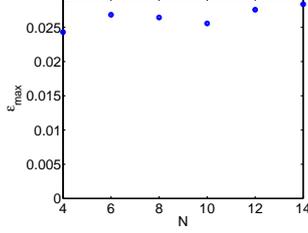


Figure S2: The maximum uncertainty of the Bloch vector elements defined in Eq. (S21) for the optimal measurement settings as a function of the number of qubits, N , for $N = 4, 6, 8, 10, 12$ and 14 .

where the projector witness is defined as

$$\mathcal{W}_{D(4,2)}^{(P)} = \frac{2}{3} \cdot \mathbb{1} - |D_4^{(2)}\rangle\langle D_4^{(2)}|. \quad (\text{S17})$$

Hence, the fidelity with respect to the state $|D_4^{(2)}\rangle$ is bounded from below as [S4]

$$F_{D(4,2)} \geq \frac{2}{3} - \frac{1}{3} \langle \mathcal{W}_{D(4,2)}^{(P3)} \rangle. \quad (\text{S18})$$

Fidelities with respect to the four-qubit Dicke states. In Table S1 we summarize the results for full tomography (full) and for permutationally invariant tomography (pi) for random (ran) and optimized (opt) directions. To obtain a physical density matrix with non-negative eigenvalues we perform a maximum-likelihood fit (max-like) of the measured data. In Fig. S1, the differences between the density matrix obtained from full tomography and the ones obtained from permutationally invariant tomography can be seen.

Efficient representation of permutationally invariant operators on a digital computer. Every PI operator O can be decomposed as

$$O = \sum_{k+l+m+n=N} c_{k,l,m,n}^{(O)} (X^{\otimes k} \otimes Y^{\otimes l} \otimes Z^{\otimes m} \otimes \mathbb{1}^{\otimes n})_{\text{PI}}. \quad (\text{S19})$$

Such a decomposition for operators of the form $(A^{\otimes(N-n)} \otimes \mathbb{1}^{\otimes n})_{\text{PI}}$ with $A = a_x X + a_y Y + a_z Z$ is given by

$$\sum_{k,l,m} a_x^k a_y^l a_z^m \frac{(k+l+m)!}{k!l!m!} (X^{\otimes k} \otimes Y^{\otimes l} \otimes Z^{\otimes m} \otimes \mathbb{1}^{\otimes n})_{\text{PI}}, \quad (\text{S20})$$

where the summation is carried out such that $k+l+m+n=N$.

Results for larger systems. We determined the optimal A_j for PI tomography for $N = 4, 6, \dots, 14$. In Fig. S2, we plot the maximal uncertainty of the Bloch vector elements

$$\epsilon_{\max} = \max_{k,l,m,n} \mathcal{E}[(X^{\otimes k} \otimes Y^{\otimes l} \otimes Z^{\otimes m} \otimes \mathbb{1}^{\otimes n})_{\text{PI}}] \quad (\text{S21})$$

for the total count realized in the experiment $\lambda_j = \lambda = 2050$ as a function of N , when the state of the system is $\varrho_0 = \mathbb{1}/2^N$. It increases slowly with N . Thus, for large N the number of

counts per measurement setting does not have to increase very much in order to keep the maximal uncertainty of the Bloch vector elements the same as for the $N = 4$ case. In particular, for $N = 14$, a total count of 2797 per setting yields the same maximal uncertainty as we had for the $N = 4$ case.

An upper bound on the uncertainty of PI tomography for ϱ_0 different from the white noise can be obtained by using $[\Delta(A_j^{\otimes(N-n)} \otimes \mathbb{1}^{\otimes n})_{\text{PI}}]_{\varrho_0}^2 = 1$ for error calculations. According to numerics, for optimal A_j for $N = 4, 6, \dots, 14$, ϵ_{\max} remains the same as in the case of white noise, since for the full correlation terms with $n = 0$ the upper bound equals the value for white noise, and the full correlations terms contribute to the noise of the Bloch vector elements with the largest uncertainty. Thus, the total count per setting will not increase more with the number of qubits even for states different from the completely mixed state.

The operators that give a bound on $\langle P_s \rangle$ with three settings for $N = 6$ and 8 are the following

$$\begin{aligned} P_s^{(6)} &\geq \frac{2}{225}(Q_2 + J_z^2) - \frac{1}{90}(Q_4 + J_z^4) + \frac{1}{450}(Q_6 + J_z^6), \\ P_s^{(8)} &\geq -0.001616Q_2 + 0.002200Q_4 - 0.0006286Q_6 \\ &\quad + 0.00004490Q_8 + 0.003265J_z^2 - 0.004444J_z^4 \\ &\quad + 0.001270J_z^6 - 0.00009070J_z^8, \end{aligned} \quad (\text{S22})$$

where $Q_n = J_x^n + J_y^n$. They were determined using semi-definite programming, with a method similar to one used for obtaining three-setting witnesses in Ref. [S4]. They have an expectation value $+1$ for the Dicke states $|D_6^{(3)}\rangle$ and $|D_8^{(4)}\rangle$, respectively. Moreover, their expectation value give the highest possible lower bound on $\langle P_s \rangle$ for states of the form

$$\varrho_{\text{noisy}}(p) = p \frac{\mathbb{1}}{2^N} + (1-p) |D_N^{(N/2)}\rangle\langle D_N^{(N/2)}|. \quad (\text{S23})$$

among the operators that are constructed as a linear combination of the operators J_i^n . The validity of the relations in Eq. (S22) can easily be checked by direct calculation.

Bounding the differences between elements of ϱ and ϱ_{PI} based on the fidelity. For any pure state $|\Psi\rangle$, it is possible to bound the difference between $|\langle \Psi | \varrho_{\text{PI}} | \Psi \rangle|$ and $|\langle \Psi | \varrho | \Psi \rangle|$ as

$$|\langle \Psi | \varrho | \Psi \rangle - \langle \Psi | \varrho_{\text{PI}} | \Psi \rangle| \leq \sqrt{1 - F(\varrho, \varrho_{\text{PI}})}. \quad (\text{S24})$$

Thus, if the fidelity is close to 1, then $\langle \Psi | \varrho | \Psi \rangle \approx \langle \Psi | \varrho_{\text{PI}} | \Psi \rangle$, even if $|\Psi\rangle$ is non-symmetric. If $|\Psi\rangle$ is an element of the product basis, e.g., $|0011\rangle$, then Eq. (S24) is a bound on the difference between the corresponding diagonal elements of ϱ and ϱ_{PI} .

Eq. (S24) can be proved as follows: There is a well-known relation between the trace norm and the fidelity [S5]

$$\frac{1}{2} \|\varrho - \varrho_{\text{PI}}\|_{\text{tr}} \leq \sqrt{1 - F(\varrho, \varrho_{\text{PI}})}. \quad (\text{S25})$$

Moreover, for a projector P and density matrices ϱ_k we have [S6]

$$|\text{Tr}(P\varrho_1) - \text{Tr}(P\varrho_2)| \leq \frac{1}{2} \|\varrho_1 - \varrho_2\|_{\text{tr}}. \quad (\text{S26})$$

4. Quantum state tomography

4

Combining Eq. (S25) and Eq. (S26), leads to Eq. (S24).

* Present address: Faculty of Physics, University of Vienna, Boltzmanngasse 5, A-1090 Wien, Austria

[S1] C. Schmid, Ph.D. Thesis, Ludwig-Maximilian-Universität, Munich, Germany, 2008; D.F.V. James *et al.*, Phys. Rev. A **64**,

052312 (2001).

[S2] B. Jungnitsch *et al.*, Phys. Rev. Lett. **104**, 210401 (2010).

[S3] See, for example, J.I. Cirac *et al.*, Phys. Rev. Lett. **82**, 4344 (1999).

[S4] G. Tóth *et al.*, New J. Phys. **11**, 083002 (2009).

[S5] J.A. Miszczak *et al.*, J. Quant. Inf. Comp. **9**, 0103 (2009).

[S6] See Eqs. (9.18) and (9.22) of M.A. Nielsen and I.L. Chuang, *Quantum computation and quantum information* (Cambridge University Press, Cambridge, 2000).

P4.2

PRL 113, 040503 (2014)

PHYSICAL REVIEW LETTERS

week ending
25 JULY 2014

Experimental Comparison of Efficient Tomography Schemes for a Six-Qubit State

Christian Schwemmer,^{1,2} Géza Tóth,^{3,4,5} Alexander Niggebaum,⁶Tobias Moroder,⁷ David Gross,⁸ Otfried Gühne,⁷ and Harald Weinfurter^{1,2}¹Max-Planck-Institut für Quantenoptik, Hans-Kopfermann-Straße 1, D-85748 Garching, Germany²Department für Physik, Ludwig-Maximilians-Universität, D-80797 München, Germany³Department of Theoretical Physics, University of the Basque Country UPV/EHU, P.O. Box 644, E-48080 Bilbao, Spain⁴IKERBASQUE, Basque Foundation for Science, E-48011 Bilbao, Spain⁵Wigner Research Centre for Physics, Hungarian Academy of Sciences, P.O. Box 49, H-1525 Budapest, Hungary⁶School of Physics and Astronomy, University of Birmingham, B15 2TT Birmingham, United Kingdom⁷Naturwissenschaftlich-Technische Fakultät, Universität Siegen, Walter-Flex-Straße 3, D-57068 Siegen, Germany⁸Physikalisches Institut & FDM, Universität Freiburg, Rheinstraße 10, D-79104 Freiburg, Germany

(Received 29 January 2014; published 24 July 2014)

Quantum state tomography suffers from the measurement effort increasing exponentially with the number of qubits. Here, we demonstrate permutationally invariant tomography for which, contrary to conventional tomography, all resources scale polynomially with the number of qubits both in terms of the measurement effort as well as the computational power needed to process and store the recorded data. We demonstrate the benefits of combining permutationally invariant tomography with compressed sensing by studying the influence of the pump power on the noise present in a six-qubit symmetric Dicke state, a case where full tomography is possible only for very high pump powers.

DOI: 10.1103/PhysRevLett.113.040503

PACS numbers: 03.67.Mn, 03.65.Wj, 42.50.Dv

Introduction.—The number of controllable qubits in quantum experiments is steadily growing [1,2]. Yet, to fully characterize a multiqubit state via quantum state tomography (QST), the measurement effort scales exponentially with the number of qubits. Moreover, the amount of data to be saved and the resources to process them scale exponentially, too. Thus, the limit of conventional QST will soon be reached. The following question arises: how much information about a quantum state can be inferred without all the measurements a full QST would require? Protocols have been proposed which need significantly fewer measurement settings if one has additional knowledge about a state, e.g., that it is of low rank, a matrix product state, or a permutationally invariant (PI) state [3–8]. Some of these approaches only require a polynomially increasing number of measurements and even offer scalable postprocessing algorithms [5,8]. Yet, it is important to test the different approaches and evaluate their results for various quantum states.

Here, we implement and compare four different QST schemes in a six-photon experiment. In detail, we perform the largest QST of a photonic multiqubit state so far. We use these data as a reference for a detailed evaluation of different tomography schemes, which enable the state determination with significantly fewer measurements. The recently proposed, scalable PI analysis is implemented

here and thus enables us, for the first time, to also perform the numerical evaluation with polynomial resources only. We evaluate the convergence of compressed sensing (CS) schemes and show that the combination of PI and CS can further reduce the measurement effort, without sacrificing performance. We demonstrate the usability of these significantly improved methods to characterize the effects of higher-order emission in spontaneous parametric down-conversion (SPDC), an analysis which would not have been possible without the novel tomography schemes.

Scalable scheme for measurements.—Let us first consider the measurement effort needed for tomography. For full QST, each N -qubit state is associated with a normalized non-negative Hermitian matrix ρ with $4^N - 1$ real free parameters. Since all free parameters have to be determined, any scheme that is suitable to fully analyzing an arbitrary state, such as, e.g., the standard Pauli tomography scheme, suffers from an exponentially increasing measurement effort [9,10]. PI states, in contrast, are described by only $\binom{N+3}{N} - 1 = O(N^3)$ free parameters. Tomography in the PI subspace can be performed by measuring (global) operators of the form $A_i^{\otimes N}$ with $A_i = \vec{n}_i \vec{\sigma}$, i.e., measurements of the polarization along the same direction \vec{n}_i for every photon [7]. Here, $|\vec{n}_i| = 1$ and $\vec{\sigma} = (\sigma_x, \sigma_y, \sigma_z)$ with Pauli operators σ_i ($i = x, y, z$). Each single measurement setting $A_i^{\otimes N}$ delivers N expectation values of the operators $M_i^n = (1/N!) \sum_k \Pi_k [|0\rangle_i \langle 0|^{\otimes(N-n)} \otimes |1\rangle_i \langle 1|^{\otimes n}] \Pi_k^\dagger$, where the summation is over all permutations Π_k and i refers to the eigenbasis of A_i . This reduces the number of necessary

Published by the American Physical Society under the terms of the Creative Commons Attribution 3.0 License. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.

4. Quantum state tomography

settings to $\mathcal{D}_N = \binom{N+2}{N} = \frac{1}{2}(N^2 + 3N + 2) = O(N^2)$.

Note that if one allows global entangled measurements, this number can be further reduced [11]. Most importantly, whether an unknown N -qubit state is close to being PI can be checked in advance by measuring the settings $\sigma_x^{\otimes N}$, $\sigma_y^{\otimes N}$, and $\sigma_z^{\otimes N}$. These measurements are already sufficient to give a lower bound for the overlap with the symmetric subspace [7,12].

Scalable representation of states and operators.—The above approach not only reduces the experimental effort, it also offers the possibility to efficiently store and process the measured data. Describing states in the PI subspace enables an efficient representation with only polynomial scaling of the storage space and processing time [8,13].

Consider the angular momentum basis states $|j, j_z, \alpha\rangle$ for the N -qubit Hilbert space, with $\vec{J}^2|j, j_z, \alpha\rangle = j(j+1)|j, j_z, \alpha\rangle$ and $J_z|j, j_z, \alpha\rangle = j_z|j, j_z, \alpha\rangle$, where the total spin numbers are restricted to be $j = j_{\min}, j_{\min} + 1, \dots, N/2$ starting from $j_{\min} = 0$ for N even and $j_{\min} = \frac{1}{2}$ for N odd, while $j_z = -(N/2), -(N/2) + 1, \dots, N/2$. Here, $\alpha = 1, 2, \dots, d_j$ is a label to remove the degeneracy (of degree d_j [14]) of the eigenstates of \vec{J}^2 and J_z . In this basis, PI states can be written in a simple block diagonal form

$$\rho_{\text{PI}} = \bigoplus_{j=j_{\min}}^{N/2} \frac{\mathbb{1}_{d_j}}{d_j} \otimes p_j \rho_j, \quad (1)$$

with ρ_j being the density operators of the spin- j subspace and p_j a probability distribution. Hence, it is sufficient to consider only the $N/2$ blocks $\tilde{\rho}_j = p_j \rho_j / d_j$ (of which each has a multiplicity of d_j ; see Fig. 1) with the largest block—the symmetric subspace—being of dimension $(N+1) \times (N+1)$ and multiplicity $d_{N/2} = 1$. Consequently, a PI state can be stored efficiently.

Even if the state to be analyzed is not PI, as long as the observable to be measured is PI, one can hugely benefit from the scheme, since a similarly scalable decomposition can be found for any PI operator O , i.e., $O = \bigoplus_j \mathbb{1}_{d_j} \otimes O_j$.

Together with Eq. (1), this yields an efficient way to also calculate the expectation values $\langle O \rangle = \text{Tr}(\rho O) =$

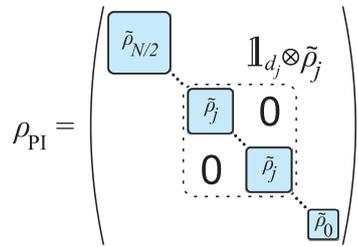


FIG. 1 (color online). Every PI state can be decomposed into a block diagonal form. Exemplarily shown is the combination of d_j block matrices $\tilde{\rho}_j$ which are all identical.

$\sum_j p_j \text{Tr}(\rho_j O_j)$ for non-PI states. Note that while, in the regular case, the trace has to be taken over the product of two 2^N -dimensional matrices, now we only have about $\frac{N}{2}$ terms with traces of at most $(N+1)$ -dimensional matrices. Again, the effort reduces from exponential to polynomial. For the six-qubit case ($j \in j_{\min} = 0, 1, 2, N/2 = 3$), this means that the state to be analyzed as well as each measurement operator can be described by only four Hermitian matrices of size 7×7 , 5×5 , 3×3 , and 1×1 , respectively, reducing the number of parameters from $4^6 - 1 = 4095$ to $\binom{9}{6} - 1 = 83$ only.

Data analysis starts with the counts c_i^n observed measuring M_i^n and the frequencies $f_i^n = c_i^n / \sum_k c_i^k$, respectively. Solving the system of linear equations $f_i^n \approx \langle M_i^n \rangle = \text{Tr}(\rho M_i^n)$ for the free parameters of ρ usually results in a nonpositive and thus unphysical density matrix ($\rho \not\geq 0$) due to statistical errors and misalignment. Here, typically, a maximum likelihood (ML) fitting algorithm is used to find the physical state that optimally agrees with the measured data [9,15,16]. We use convex optimization [8,19], which guarantees a unique minimum and fast convergence. The performance of our algorithm is illustrated best by the fact that a 20-qubit PI state can be reconstructed in fewer than 10 min on a standard desktop computer.

State reconstruction of low rank states and compressed sensing.—As shown recently, low rank states, i.e., states with only a few nonzero eigenvalues, enable state reconstruction even if the underlying set of data obtained from random Pauli measurements is incomplete [3]. There, the measurement effort to analyze a state of rank r with $r2^N$ free parameters scales like $O(r2^N \log 2^N)$ —clearly achieving optimal scaling up to a log factor. Despite the still exponential scaling, the square root improvement can be considerable. Since, in many cases, the state to be experimentally prepared is at the same time PI and of low rank, we demonstrate here for the first time that combining the two methods is possible [16,20].

Experimental state tomography.—Let us now compare the various QST schemes. In particular, we evaluate the number of settings necessary to obtain (almost) full knowledge about the state. As a reference, we perform, for the first time, full QST of a six-photon state. This is possible only at very high pump power (8.4 W) of the down-conversion source where we collect data for the complete set of Pauli settings. PI tomography is performed to test it against full QST and to analyze states emitted for lower pump powers. For both strategies, we also analyze the convergence of CS tomography for incomplete data.

The six-photon state observed in this work is the symmetric Dicke state $|D_6^{(3)}\rangle$. In general, symmetric Dicke states are defined as

$$|D_N^{(n)}\rangle = \binom{N}{n}^{-1/2} \sum_i \mathcal{P}_i(|H^{\otimes(N-n)}\rangle \otimes |V^{\otimes n}\rangle), \quad (2)$$

TABLE I. Overlap with the symmetric Dicke states determined from full tomography, PI tomography with 28 settings, CS with 270 settings, and CS in the PI subspace (PI,CS) with 16 settings. The fidelities for all tomography schemes were determined from the respective ML reconstructed states. Nonparametric bootstrapping [23] was performed from which the corresponding standard deviations were determined as < 0.005 , < 0.015 , < 0.008 , and < 0.020 for full tomography, PI tomography, CS, and CS in the PI subspace, respectively.

State	Full	PI	CS	PI, CS
$ D_6^{(0)}\rangle$	0.001	0.001	0.001	0.002
$ D_6^{(1)}\rangle$	0.005	0.008	0.011	0.006
$ D_6^{(2)}\rangle$	0.197	0.222	0.181	0.207
$ D_6^{(3)}\rangle$	0.604	0.590	0.615	0.592
$ D_6^{(4)}\rangle$	0.122	0.127	0.118	0.119
$ D_6^{(5)}\rangle$	0.003	0.004	0.003	0.005
$ D_6^{(6)}\rangle$	0.000	0.003	0.001	0.004
Σ	0.933	0.954	0.929	0.935

where $|H/V\rangle_i$ denotes horizontal or vertical polarization in the i th mode and the \mathcal{P}_i represent all the distinct permutations. In order to experimentally observe $|D_6^{(3)}\rangle$, we distribute an equal number of H and V polarized photons over six output modes and apply conditional detection (for details, see the Supplemental Material [16] and Ref. [21]). The setup uses cavity enhanced SPDC [22] with special care taken to further reduce losses of all components and to optimize the yield of $|D_6^{(3)}\rangle$.

Data are recorded at a pump power of 8.40 ± 0.56 W over 4 min for each of the $3^6 = 729$ Pauli settings. The six-photon count rate was 58 events per minute on average, leading to about 230 events per basis setting within a total measurement time of approximately 50 h [24]. The reconstructed density matrix can be seen in Fig. 2(a). Table I lists the fidelity [25] with all the various Dicke states. Their sum reaches high values, proving that the state is close to the symmetric subspace.

Evidently, the experimental state is a mixture of mainly $|D_6^{(2)}\rangle$, $|D_6^{(3)}\rangle$, and $|D_6^{(4)}\rangle$, and thus CS might be used beneficially. The following question arises: how many settings are required for CS for a faithful reconstruction of the state? We chose random subsets of up to 300 settings from the 729 settings for full tomography. Figure 2(d) gives the probability distribution of the fidelity of the reconstructed matrix for a bin size of 0.01 with respect to the results of full tomography. While, for a low number of settings (< 10), the results are randomly spread out, the overlap is already, on average, ≥ 0.800 for 20 settings. We find that to reach a fidelity of ≥ 0.950 , one requires about 270 settings. Figure 2(c) shows the density matrix obtained from 270 settings [$F(q_{CS}, q_{full}) = 0.950$].

PI tomography should be clearly more efficient. To test its applicability, we first determined the lower bound for the

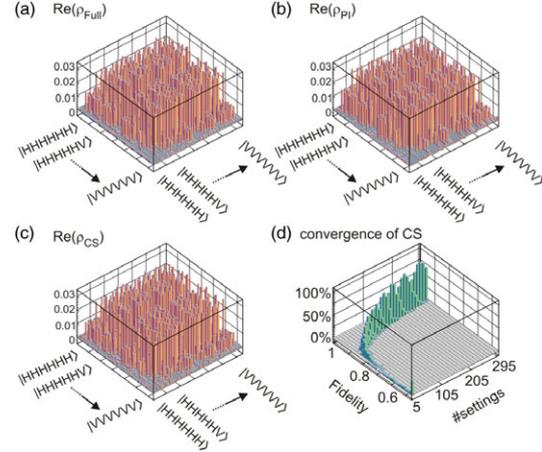


FIG. 2 (color online). ML reconstruction of the state $|D_6^{(3)}\rangle$ obtained from (a) full and (b) PI tomography and (c) CS with 270 settings performed at a pump power of 8.4 W. The respective fidelities are 0.604, 0.590, and 0.615 with mutual overlaps of $F(q_{full}, q_{PI}) = 0.922$, $F(q_{full}, q_{CS}) = 0.950$, and $F(q_{PI}, q_{CS}) = 0.908$. (d) Probability to obtain a certain fidelity for CS with a certain number of randomly chosen settings in comparison with full tomography.

projection of the state onto the symmetric subspace, i.e., the largest block in Fig. 1, $\langle P_s^{(6)} \rangle$ from the settings $\sigma_x^{\otimes 6}$, $\sigma_y^{\otimes 6}$, and $\sigma_z^{\otimes 6}$ by analyzing all photons under $\pm 45^\circ$, right- or left-circular, and H/V polarization. We found that $\langle P_s^{(6)} \rangle \geq 0.922 \pm 0.055$, indicating that it is legitimate to use PI tomography, which for six qubits only requires 25 more settings [16]. Under the same experimental conditions as before and 4 min of data collection per setting, we performed the experiment within 2 h only. The density matrix q_{PI} obtained is shown in Fig. 2(b), with its symmetric subspace shown in Fig. 3(a). The fidelities with the symmetric Dicke states for PI tomography can be found again in Table I. (For the projector to the Dicke state $|D_N^{(n)}\rangle$, all $\{O_j\}_{kl} = 0$, except for $\{O_{N/2}\}_{n+1, n+1} = 1$.) The overlap between the

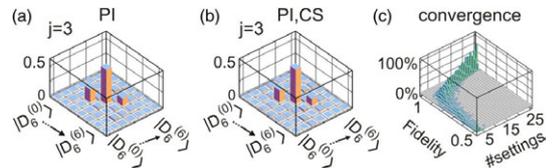


FIG. 3 (color online). Symmetric subspaces ($j = 3$) obtained with (a) PI tomography and (b) CS in the PI subspace with 16 settings. The central bars can be associated with the target state $|D_6^{(3)}\rangle$ and the small bars next to it with $|D_6^{(2)}\rangle$ and $|D_6^{(4)}\rangle$ originating from higher-order noise. (c) Probability to observe a certain fidelity for arbitrarily chosen tomographically incomplete sets of settings in comparison with PI tomography from 28 settings. For 16 settings, the overlap is ≥ 0.950 on average.

4. Quantum state tomography

reconstructed states using either full or PI tomography is 0.922, which is equivalent to the fidelity of 0.923 between full tomography and its PI part. Clearly, PI tomography rapidly and precisely determines the PI component of the state.

PI tomography with CS.—To speed up analysis even further, based on subsets of the data used for PI tomography, we derived the density matrix $\rho_{\text{PI,CS}}$; see Fig. 3(b). Here, the fidelity averaged over a series of different samples is above 0.950 for 16 or more settings [Fig. 3(c)]. Again, both methods are compatible within 1 standard deviation. In summary, our results prove that PI tomography (with CS) enables precise state reconstruction with minimal experimental and computational effort.

Application to noise analysis.—As the count rates for six-photon states depend on the cube of the pump power, full QST is not possible for lower pump power within reasonable time and thus does not allow us to analyze the features of multiphoton states obtained from SPDC. As SPDC is a spontaneous process, with certain probability, there are cases where eight photons have been emitted but only six have been detected, leading to an admixture of $\rho_{D_6^{(2)}}$ and $\rho_{D_6^{(4)}}$. Ideally, the amplitude of the two admixtures should be the same, but, due to polarization dependent coupling efficiencies of H and V photons [26,27], this is not the case. Therefore, we extended the noise model [28] to better specify the experimental state using $\rho_{\text{exp}}^{\text{noise}}(q, \lambda) = (1 - q)\rho_{D_6^{(3)}} + q\rho_6^{\text{asym}}(\lambda)$, with $\rho_6^{\text{asym}}(\lambda) = \frac{4}{7}\rho_{D_6^{(3)}} + 3/14[(1 + \lambda)\rho_{D_6^{(2)}} + (1 - \lambda)\rho_{D_6^{(4)}}]$, the noise q , and the asymmetry parameter λ . Both q and λ can be determined from the fidelities to the Dicke states (see also the Supplemental Material [16]). At 8.4 W, noise parameters of $q = 0.807 \pm 0.013$ and $\lambda = 0.234 \pm 0.015$ were obtained from full tomography, which agree well with those from PI tomography ($q = 0.867 \pm 0.041$ and $\lambda = 0.273 \pm 0.059$). After convincing ourselves that (CS) PI tomography is in excellent agreement with full QST, we can now also perform tomography for low pump powers.

We performed PI analysis at 3.7, 5.1, 6.4, and 8.6 W [see Fig. 4(a)] with sampling times of 67, 32, 18, and 15 h and average counts per setting of 340, 390, 510, and 610, respectively. PI tomography shows an increase of the noise parameter q from 0.677 ± 0.029 for 3.7 W to 0.872 ± 0.023 for 8.6 W due to the increasing probability of eight-photon emission for high pump power [29]. Note that the ratio between six-photon detection from eight-photon emission relative to detection from six-photon emission is given by $q/(1 - q)$; i.e., for a pump power of 8.6 W, we obtain sixfold detection events with 90% probability from eight photon emissions, of which two photons were lost. Although fluctuating, the asymmetry parameter λ does not show significant dependence on the pump power and lies in the interval $[0.136 \pm 0.042, 0.200 \pm 0.053]$ for PI tomography (within $[0.101 \pm 0.116, 0.190 \pm 0.071]$ for CS in the PI subspace). This confirms that the difference in the coupling efficiency of H and V does not change with the pump power [see Fig. 4(b)]. The fidelity between the ML fits and the noise model $\rho_{\text{exp}}^{\text{noise}}(p, \lambda)$ is > 0.925 for all pump levels, and, for CS in the PI subspace, it is > 0.897 . The high values indicate that our noise model adequately describes the experimental results.

As an example where full knowledge of ρ is necessary, let us consider the quantum Fisher information F_Q which measures the suitability of ρ to estimate the phase θ in an evolution $U(\theta, \mathcal{H}) = e^{-i\theta\mathcal{H}}$ [30]. Here, we want to test whether, in spite of the higher-order noise, the reconstructed states still exhibit sub-shot-noise phase sensitivity. For \mathcal{H} , we choose the collective spin operator $J_x = 1/2 \sum_{i=1}^N \sigma_x^{(i)}$, where $\sigma_x^{(i)}$ is σ_x acting on the i th particle. In the case of $N = 6$, a value $F_Q > 6$ indicates sub-shot-noise phase sensitivity. We observed 11.858 ± 0.576 , 10.904 ± 0.528 , 10.289 ± 0.468 , and 9.507 ± 0.411 for the corresponding pump powers from 3.7 to 8.6 W [29] [see Fig. 4(b)]; i.e., sub-shot-noise phase sensitivity is maintained for high pump powers.

Conclusions.—We compared standard quantum state tomography with the significantly more efficient permutationally invariant tomography and also with compressed sensing in the permutationally invariant subspace. For this purpose, we used data of the symmetric Dicke state $|D_6^{(3)}\rangle$ obtained from spontaneous parametric down-conversion of very high pump power. All methods give compatible results within their statistical errors. The number of measurement settings was gradually reduced from 729 for full tomography, to 270 for compressed sensing, to 28 for permutationally invariant tomography, and to only 16 for compressed sensing in the permutationally invariant subspace, giving, in total, a reduction of about a factor of 50 without significantly changing the quantities specifying the state. We applied this highly efficient state reconstruction scheme to study the dependence of higher-order noise on the pump power, clearly demonstrating its benefits for the analysis of multiqubit states required for future quantum computation and quantum simulation applications.

We thank R. Krischek, W. Wieczorek, Z. Zimborás, S. Neuhaus, and L. Knips for stimulating discussions. We

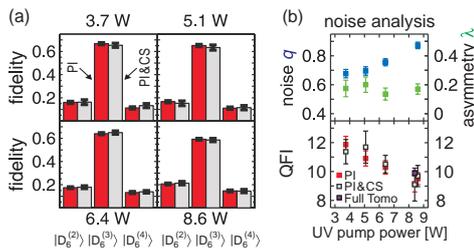


FIG. 4 (color online). (a) Observed fidelities with the states $|D_6^{(2)}\rangle$, $|D_6^{(3)}\rangle$, and $|D_6^{(4)}\rangle$ at different ultraviolet (UV) pump powers for PI tomography and CS in the PI subspace from 12 settings. The error bars were determined by nonparametric bootstrapping [23]. (b) The influence of the pump power on the higher-order noise expressed via the noise q and the asymmetry parameter λ (upper part) and the phase estimation sensitivity expressed via the quantum Fisher information (QFI) (lower part).

acknowledge the support of this work by the EU (QWAD, ERC StG GEDENTQOPT, ERC QOLAPS, CHIST-ERA QUASAR, Marie Curie CIG 293993/ENFOQI), the Excellence Initiative of the German Federal and State Governments (ZUK 43), the DFG, FQXi Fund (Silicon Valley Community Foundation), the MINECO (Project No. FIS2012-36673-C03-03), the Basque Government (Project No. IT4720-10), and the National Research Fund of Hungary OTKA (Contract No. K83858). D. G. acknowledges Grants No. W911NF-14-1-0098 and No. W911NF-14-1-0133 from the ARO. C. S. thanks QCCC of the Elite Network of Bavaria for support.

- [1] 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, and R. Blatt, *Nature (London)* **438**, 643 (2005); T. Monz, P. Schindler, J. T. Barreiro, M. Chwalla, D. Nigg, W. A. Coish, M. Harlander, W. Hänsel, M. Hennrich, and R. Blatt, *Phys. Rev. Lett.* **106**, 130506 (2011).
- [2] Y.-F. Huang, B.-H. Liu, L. Peng, Y.-H. Li, L. Li, C.-F. Li, and G.-C. Guo, *Nat. Commun.* **2**, 546 (2011); X.-C. Yao, T.-X. Wang, P. Xu, H. Lu, G.-S. Pan, X.-H. Bao, C.-Z. Peng, C.-Y. Lu, Y.-A. Chen, and J.-W. Pan, *Nat. Photonics* **6**, 225 (2012).
- [3] 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); M. Ohliger, V. Nesme, D. Gross, Y.-K. Liu, and J. Eisert, *arXiv:1111.0853*; M. Guță, T. Kypraios, and I. Dryden, *New J. Phys.* **14**, 105002 (2012); M. Ohliger, V. Nesme, and J. Eisert, *New J. Phys.* **15**, 015024 (2013); A. Smith, C. A. Riofrio, B. E. Anderson, H. Sosa-Martinez, I. H. Deutsch, and P. S. Jessen, *Phys. Rev. A* **87**, 030102(R) (2013).
- [4] 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).
- [5] T. Baumgratz, D. Gross, M. Cramer, and M. B. Plenio, *Phys. Rev. Lett.* **111**, 020401 (2013); T. Baumgratz, A. Nüßeler, M. Cramer, and M. B. Plenio, *New J. Phys.* **15**, 125004 (2013).
- [6] J. O. S. Yin and S. J. van Enk, *Phys. Rev. A* **83**, 062110 (2011).
- [7] G. Tóth, W. Wieczorek, D. Gross, R. Krischek, C. Schwemmer, and H. Weinfurter, *Phys. Rev. Lett.* **105**, 250403 (2010).
- [8] T. Moroder, P. Hyllus, G. Tóth, C. Schwemmer, A. Niggebaum, S. Gaile, O. Gühne, and H. Weinfurter, *New J. Phys.* **14**, 105001 (2012).
- [9] M. Paris and J. Řeháček, *Quantum State Estimation* (Springer-Verlag, Berlin, Heidelberg, 2004).
- [10] N. Kiesel, 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).
- [11] A. B. Klimov, G. Björk, and L. L. Sánchez-Soto, *Phys. Rev. A* **87**, 012109 (2013).
- [12] For six qubits, we have $P_i^{(6)} \geq 2/225(J_x^2 + J_y^2 + J_z^2) - 1/90(J_x^4 + J_y^4 + J_z^4) + 1/450(J_x^6 + J_y^6 + J_z^6)$ with $J_i = \frac{1}{2} \sum_k \sigma_i^{(k)}$ and $\sigma_i^{(k)}$ the application of σ_i on the k th qubit.
- [13] R. B. A. Adamson, P. S. Turner, M. W. Mitchell, and A. M. Steinberg, *Phys. Rev. A* **78**, 033832 (2008).
- [14] J. I. Cirac, A. K. Ekert, and C. Macchiavello, *Phys. Rev. Lett.* **82**, 4344 (1999).
- [15] Z. Hradil, *Phys. Rev. A* **55**, R1561 (1997).
- [16] See Supplemental Material at <http://link.aps.org/supplemental/10.1103/PhysRevLett.113.040503> for additional derivations, which includes Refs. [17,18].
- [17] G. Tóth, W. Wieczorek, R. Krischek, N. Kiesel, P. Michelberger, and H. Weinfurter, *New J. Phys.* **11**, 083002 (2009).
- [18] For efficient entanglement detection methods for PI density matrices, see L. Novo, T. Moroder, and O. Gühne, *Phys. Rev. A* **88**, 012305 (2013); M. Bergmann and O. Gühne, *J. Phys. A* **46**, 385304 (2013).
- [19] S. Boyd and S. Vandenberghe, *Convex Optimization* (Cambridge University Press, Cambridge, England, 2004).
- [20] In order to apply CS and PI tomography together, it suffices that the ρ_j matrices in Eq. (1) are of low rank. This also covers the case of global density matrices of comparatively high rank, such as the PI multiqubit singlet state discussed by I. Urizar-Lanz, P. Hyllus, I. L. Egusquiza, M. W. Mitchell, and G. Tóth, *Phys. Rev. A* **88**, 013626 (2013). For this highly mixed state, for even N , $\rho_0 = 1$, $p_0 = 1$, and all other p_j 's are 0.
- [21] W. Wieczorek, R. Krischek, N. Kiesel, P. Michelberger, G. Tóth, and H. Weinfurter, *Phys. Rev. Lett.* **103**, 020504 (2009); R. Prevedel, G. Cronenberg, M. S. Tame, M. Paternostro, P. Walther, M. S. Kim, and A. Zeilinger, *Phys. Rev. Lett.* **103**, 020503 (2009).
- [22] R. Krischek, W. Wieczorek, A. Ozawa, N. Kiesel, P. Michelberger, T. Udem, and H. Weinfurter, *Nat. Photonics* **4**, 170 (2010).
- [23] B. Efron and R. J. Tibshirani, *An Introduction to the Bootstrap* (Chapman and Hall, London, 1994). The error bars computed via this commonly employed method are a mere quantitative estimate about the fluctuations of the reconstructed state. One should keep in mind that the state reconstruction unavoidably induces systematic errors; see C. Schwemmer, L. Knips, D. Richart, T. Moroder, M. Kleinmann, O. Gühne, and H. Weinfurter, *arXiv:1310.8465*, which are not recognizable with bootstrap analysis.
- [24] The overall runtime of the experiment was almost two weeks, from which 50 h of useful data could be extracted.
- [25] We use the Uhlmann fidelity $F(\rho_1, \rho_2) = \text{Tr}(\sqrt{\sqrt{\rho_1}\rho_2\sqrt{\rho_1}})^2$ between the states ρ_1 and ρ_2 which simplifies to $F(\rho_1, \rho_2) = \text{Tr}(\rho_1\rho_2)$ if one of the two states is pure.
- [26] R. S. Bennink, Y. Liu, D. D. Earl, and W. P. Grice, *Phys. Rev. A* **74**, 023802 (2006).
- [27] P. Trojek, Ph.D. thesis, Ludwig-Maximilians-Universität München, 2007.
- [28] W. Wieczorek, Ph.D. thesis, Ludwig-Maximilians-Universität München, 2009.
- [29] The corresponding noise parameter q determined from CS in the PI subspace with 12 randomly chosen settings is 0.725 ± 0.052 for 3.7 W and 0.886 ± 0.040 for 8.6 W. The expectation values for the Fisher information are 11.371 ± 0.842 , 11.685 ± 1.113 , 10.500 ± 0.613 , and 9.688 ± 0.733 for CS in the PI subspace.
- [30] The quantum Fisher information is defined as $F_Q(\theta, \mathcal{R}) = 2 \sum_{i,j} (\lambda_i - \lambda_j)^2 / (\lambda_i + \lambda_j) |\langle i | \mathcal{R} | j \rangle|^2$ with $\{\lambda_i, |i\rangle\}$ the eigenspectrum of Q ; see P. Hyllus, W. Laskowski, R. Krischek, C. Schwemmer, W. Wieczorek, H. Weinfurter, L. Pezzé, and A. Smerzi, *Phys. Rev. A* **85**, 022321 (2012).

4. Quantum state tomography

1

EXPERIMENTAL COMPARISON OF EFFICIENT TOMOGRAPHY SCHEMES FOR A SIX-QUBIT STATE (SUPPLEMENTAL MATERIAL)

Christian Schwemmer,^{1,2} Géza Tóth,^{3,4,5} Alexander Niggebaum,⁶
Tobias Moroder,⁷ David Gross,⁸ Otfried Gühne,⁷ and Harald Weinfurter^{1,2}

¹Max-Planck-Institut für Quantenoptik, Hans-Kopfermann-Straße 1, D-85748 Garching, Germany

²Department für Physik, Ludwig-Maximilians-Universität, D-80797 München, Germany

³Department of Theoretical Physics, University of the Basque Country UPV/EHU, P.O. Box 644, E-48080 Bilbao, Spain

⁴IKERBASQUE, Basque Foundation for Science, E-48011 Bilbao, Spain

⁵Wigner Research Centre for Physics, Hungarian Academy of Sciences, P.O. Box 49, H-1525 Budapest, Hungary

⁶School of Physics and Astronomy, University of Birmingham, B15 2TT Birmingham, United Kingdom

⁷Naturwissenschaftlich-Technische Fakultät, Universität Siegen, Walter-Flex-Straße 3, D-57068 Siegen, Germany

⁸Physikalisches Institut & FDM, Universität Freiburg, Rheinstraße 10, D-79104 Freiburg, Germany

In the Supplemental Material, we present further experimental results and calculations.

THE SETUP

The photon source is based on a femtosecond enhancement cavity in the UV with a 1 mm thick β -barium borate (BBO) crystal cut for type II phase matching placed inside [22] (Fig. S1). In order to compensate for walk off effects a half-wave plate (HWP) and a second BBO crystal of 0.5 mm are applied. Spatial filtering is achieved by coupling the photons into a single mode fiber (SM) and an interference filter (IF) ($\Delta\lambda = 3$ nm) enables spectral filtering. Distributing the photons into six spatial modes is realized by 3 beam splitters with a splitting ratio of 50:50 (BS₁, BS₃, BS₄) and two beam splitters with a ratio of 66:33 (BS₂, BS₅). Yttrium-vanadate (YVO₄) crystals are used to compensate for unwanted phase shifts. State analysis is realized by half-wave and quarter-wave plates (QWP) and polarizing beam splitters (PBS). The photons are detected by fiber-coupled single photon counting modules connected to a FPGA-based coincidence logic.

In Fig. S1 (lower right corner) a visualization of the measurement directions on the Bloch sphere is depicted. Each point (a_x, a_y, a_z) on the sphere corresponds to a measurement operator of the form $a_x\sigma_x + a_y\sigma_y + a_z\sigma_z$. In order to perform PI tomography for 6 qubits 28 operators have to be measured.

STATE RECONSTRUCTION

The target function to be minimized is the *logarithmic likelihood* which is given by $\sum_{k,s} \frac{n_{k,s}}{N_{\max}} \log(p_{k,s})$ where $n_{k,s}$ labels the number of counts for the outcome k when measuring setting s with the corresponding probability $p_{k,s}$ for the guess $\hat{\rho}$. In order to take into account slightly different total count numbers per setting, the $n_{k,s}$ have to be divided by the maximum count number observed in one setting $N_{\max} = \max(N_s)$.

For CS exactly the same target function has to be minimized with the only difference that the underlying set of

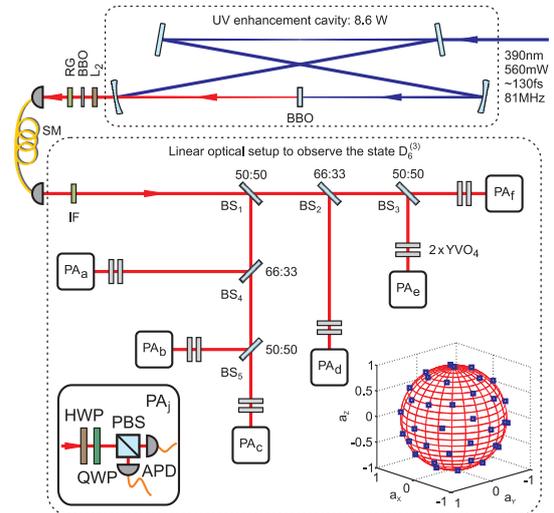


FIG. S1. Schematic drawing of the experimental setup to observe the symmetric Dicke state $|D_6^{(3)}\rangle$. For a description, see text.

measurement data is tomographically incomplete.

CONVERGENCE OF CS IN THE PI SUBSPACE

As described in the main text, we performed PI tomography together with CS in the PI subspace at different UV pump powers. In order to investigate the convergence of CS, series of different samples were randomly chosen from the full set of measurements. For all pump powers, the average fidelity with respect to all PI settings is above 0.950 as soon as the number of settings is ≥ 12 (out of 28), see Fig. S2.

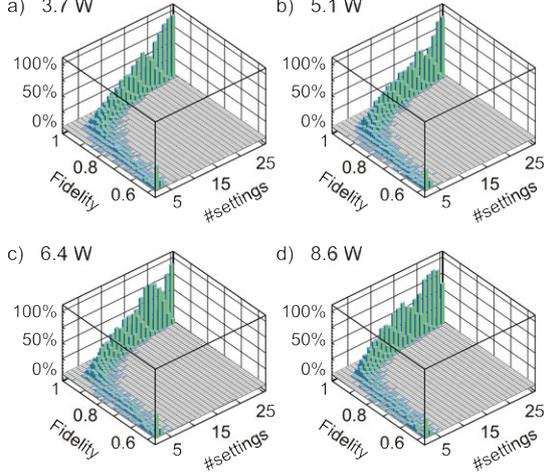


FIG. S2. Probability to observe a certain fidelity for arbitrarily chosen tomographically incomplete sets of settings in comparison with PI tomography from 28 settings at different pump levels. As soon as the number of settings surpasses 12, the state is almost perfectly determined, i.e., the overlap with respect to the states reconstructed from all settings ≥ 0.950 .

NOISE MODEL

As already explained in the main part of this paper, SPDC is a spontaneous process and therefore with a certain probability 8 photons are emitted from the source. The loss of two of these 8 photons in the linear optical setup and subsequent detection leads to an admixture of the states $\varrho_{D_6^{(2)}}$ and $\varrho_{D_6^{(4)}}$ for the case that either two H or two V polarized photons are not detected, respectively. However, in the case that one H and one V polarized photon remain undetected a considerable amount of this higher-order noise consists of the target state $\varrho_{D_6^{(3)}}$ thus preserving genuine multipartite entanglement even at high UV pump powers. The probabilities of the respective states to occur can be deduced from simple combinatorics, see Fig. S3. From this simple noise model, an experimental state of the form

$$\varrho_{\text{exp}}^{\text{noise}}(q, \lambda) = (1 - q)\varrho_{D_6^{(3)}} + q\varrho_6 \quad (\text{S1})$$

with

$$\varrho_6 = \frac{4}{7}\varrho_{D_6^{(3)}} + \frac{3}{14}[\varrho_{D_6^{(2)}} + \varrho_{D_6^{(4)}}] \quad (\text{S2})$$

would be expected. However, this is not observed experimentally since the emission angles of down-conversion photons are polarization dependent [26, 27] leading to an asymmetry in the coupling into the single mode fiber

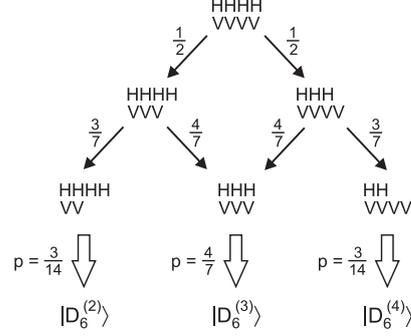


FIG. S3. The loss of two photons in a 8 photon event leads to an admixture of the state $\varrho_{D_6^{(2)}}$ and $\varrho_{D_6^{(4)}}$ to the target state. The respective probabilities p can be determined by simple combinatorics.

used. Therefore, the noisemodel was extended by the asymmetry parameter λ . Both q and λ can be deduced from the fidelities F with respect to the Dicke states $|D_6^{(2)}\rangle$, $|D_6^{(3)}\rangle$ and $|D_6^{(4)}\rangle$

$$\begin{aligned} q &= \frac{7}{3} \cdot \frac{F_{|D_6^{(2)}\rangle} + F_{|D_6^{(4)}\rangle}}{F_{|D_6^{(2)}\rangle} + F_{|D_6^{(3)}\rangle} + F_{|D_6^{(4)}\rangle}}, \\ \lambda &= \frac{F_{|D_6^{(2)}\rangle} - F_{|D_6^{(4)}\rangle}}{F_{|D_6^{(2)}\rangle} + F_{|D_6^{(4)}\rangle}}. \end{aligned} \quad (\text{S3})$$

ENTANGLEMENT WITNESS

Entanglement witnesses with respect to symmetric states are PI operators and thus can be determined efficiently. For detecting genuine multipartite entanglement, we used the entanglement witness

$$\begin{aligned} \mathcal{W} &= 0.420 \cdot \mathbb{1} - 0.700|D_6^{(3)}\rangle\langle D_6^{(3)}| \\ &\quad - 0.160|D_6^{(2)}\rangle\langle D_6^{(2)}| - 0.140|D_6^{(4)}\rangle\langle D_6^{(4)}|, \end{aligned} \quad (\text{S4})$$

where an expectation value $\langle \mathcal{W} \rangle < 0$ rules out any biseparability. In order to obtain \mathcal{W} we take an operator of the form

$$\begin{aligned} A_\alpha &= \alpha|D_6^{(3)}\rangle\langle D_6^{(3)}| + \beta|D_6^{(2)}\rangle\langle D_6^{(2)}| \\ &\quad + (1 - \alpha - \beta)|D_6^{(4)}\rangle\langle D_6^{(4)}|. \end{aligned} \quad (\text{S5})$$

An entanglement witness can be obtained as

$$W_\alpha = \max_{PPT} \langle A_\alpha \rangle \cdot \mathbb{1} - A_\alpha \quad (\text{S6})$$

where the maximum for bipartite PPT states can be obtained with semidefinite programming [17]. For $\alpha = 0.700$, $\beta = 0.160$ we have for PPT states over all parti-

4. Quantum state tomography

3

tions $\max_{PPT} \langle A_\alpha \rangle = 0.420$. It is important that semidefinite programming always finds the global optimum. A systematic generalization to construct witnesses for Dicke states can be found in Ref. [18].

Here, we want use this witness to test whether, in spite of the higher-order noise, the observed states are still genuinely six-partite entangled. For the corresponding pump powers from 3.7 W to 8.6 W, we determined the expectation value of \mathcal{W} as -0.088 ± 0.006 , -0.078 ± 0.006 , -0.075 ± 0.006 and -0.048 ± 0.005 for PI tomography and -0.082 ± 0.011 , -0.064 ± 0.013 , -0.083 ± 0.009 and -0.044 ± 0.009 for CS in the PI subspace. Clearly, due to the high probability of $\varrho_{D_6^{(3)}}$ states in the higher-order noise the entanglement is maintained also for high pump powers.

P4.3



Permutationally invariant state reconstruction

Tobias Moroder^{1,2,9}, Philipp Hyllus³, Géza Tóth^{3,4,5},
Christian Schwemmer^{6,7}, Alexander Niggebaum^{6,7},
Stefanie Gaile⁸, Otfried Gühne^{1,2} and Harald Weinfurter^{6,7}

¹ Naturwissenschaftlich-Technische Fakultät, Universität Siegen,
Walter-Flex-Straße 3, D-57068 Siegen, Germany

² Institut für Quantenoptik und Quanteninformation, Österreichische Akademie
der Wissenschaften, Technikerstraße 21A, A-6020 Innsbruck, Austria

³ Department of Theoretical Physics, University of the Basque Country
UPV/EHU, PO Box 644, E-48080 Bilbao, Spain

⁴ IKERBASQUE, Basque Foundation for Science, E-48011 Bilbao, Spain

⁵ Wigner Research Centre for Physics, Hungarian Academy of Sciences,
PO Box 49, H-1525 Budapest, Hungary

⁶ Max-Planck-Institut für Quantenoptik, Hans-Kopfermann-Straße 1,
D-85748 Garching, Germany

⁷ Fakultät für Physik, Ludwig-Maximilians-Universität, D-80797 München,
Germany

⁸ Technical University of Denmark, Department of Mathematics,
Matematiktorvet Building 303 B, 2800 Kgs. Lyngby, Denmark
E-mail: moroder@physik.uni-siegen.de

New Journal of Physics **14** (2012) 105001 (25pp)

Received 21 May 2012

Published 1 October 2012

Online at <http://www.njp.org/>

doi:10.1088/1367-2630/14/10/105001

Abstract. Feasible tomography schemes for large particle numbers must possess, besides an appropriate data acquisition protocol, an efficient way to reconstruct the density operator from the observed finite data set. Since state reconstruction typically requires the solution of a nonlinear large-scale optimization problem, this is a major challenge in the design of scalable tomography schemes. Here we present an efficient state reconstruction scheme

⁹ Author to whom any correspondence should be addressed.



Content from this work may be used under the terms of the [Creative Commons Attribution-NonCommercial-ShareAlike 3.0 licence](https://creativecommons.org/licenses/by-nc-sa/3.0/). Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

New Journal of Physics **14** (2012) 105001
1367-2630/12/105001+25\$33.00

© IOP Publishing Ltd and Deutsche Physikalische Gesellschaft

4. Quantum state tomography

2

IOP Institute of Physics Φ DEUTSCHE PHYSIKALISCHE GESELLSCHAFT

for permutationally invariant quantum state tomography. It works for all common state-of-the-art reconstruction principles, including, in particular, maximum likelihood and least squares methods, which are the preferred choices in today's experiments. This high efficiency is achieved by greatly reducing the dimensionality of the problem employing a particular representation of permutationally invariant states known from spin coupling combined with convex optimization, which has clear advantages regarding speed, control and accuracy in comparison to commonly employed numerical routines. First prototype implementations easily allow reconstruction of a state of 20 qubits in a few minutes on a standard computer.

Contents

1. Introduction	2
2. Background	4
2.1. Permutationally invariant tomography	4
2.2. Statistical state reconstruction	5
3. Method	6
3.1. Reduction of the dimensionality	6
3.2. Optimization	9
4. Examples	11
5. Details	13
5.1. Reduction	13
5.2. Convex optimization	17
5.3. Additional tools	19
6. Conclusion and outline	22
Acknowledgments	23
References	23

1. Introduction

Full information about the experimental state of a quantum system is naturally highly desirable because it enables one to determine the mean value of each observable and thus also of every other property of the quantum state. Abstractly, such a complete description is given, for example, by the density operator, a positive semidefinite matrix ρ with unit trace. Quantum state tomography [1] refers to the task to determine the density operator for a previously unknown quantum state by means of appropriate measurements. Via the respective outcomes, more and more information about the true state generating the data is collected up to the point where this information uniquely specifies the particular state. Quantum state tomography has successfully been applied in many experiments using different physical systems, including trapped ions [2] or photons [3], as prominent instances.

Unfortunately, tomography comes with a very high price due to the exponential scaling of the number of parameters required to describe composed quantum systems. For an N -qubit system the total number of parameters of the associated density operator is $4^N - 1$ and any standard tomography protocol is naturally designed to determine all these variables. The

New Journal of Physics **14** (2012) 105001 (<http://www.njp.org/>)

most common scheme used in experiments [4] consists of locally measuring in the basis of all different Pauli operators and requires an overall amount of 3^N different measurement settings with 2^N distinct outcomes each. Other schemes, e.g. using an informationally complete measurement [5] locally, would require just one setting but, nevertheless, the statistics for 4^N give different outcomes. Hence the important figure of merit to compare different methods is given by the combination of settings and outcomes.

For such a scaling, the methods rapidly become intractable, already for present state-of-the-art experiments: recording, for example, the data of 14 trapped ions [6], currently the record for quantum registers, would require about 150 days, although 100 measurement outcomes can be collected for a single setting in only about 3 s. In photonic experiments this scaling is even worse because count rates are typically much lower; for example, in recent eight-photon experiments [7, 8], a coincidence of single click events occurs only on the order of minutes; hence, it would require about 7 years to collect an adequate data set. This directly shows that more sophisticated tomography techniques are mandatory.

New tomography protocols equipped with better scaling behaviour exploit the idea that the measurement scheme is explicitly optimized only for particular kinds of states rather than for all possible ones. If the true unknown state lies within this designed target class, then full information about the state can be obtained with much less effort, and if the underlying density operator is not a member, then a certificate signals that tomography is impossible in this given case. Recent results along this direction include tomography schemes designed for states with a low rank [9–11], particularly for important low rank states such as matrix product [12, 13] or multi-scale entanglement renormalization ansatz states [14]. Other schemes include a tomography scheme based on information criteria [15, 16] or—the topic of this paper—states with permutation invariance [17].

However, it needs to be stressed that in any real experiment all these tomography schemes must cope up with another, purely statistical challenge: since only a finite number of measurements can be carried out in any experiment, one cannot access the true probabilities predicted by quantum mechanics $p_k = \text{tr}(\rho_{\text{true}} M_k)$, operator M_k describing the measurements, but merely relative frequencies f_k . Although these deviations might be small, the approximation $f_k \approx p_k$ causes severe problems in the actual state reconstruction process, i.e. the task to determine the density operator from the observed data. If one naïvely uses the frequencies according to Born's rule $f_k = \text{tr}(\hat{\rho}_{\text{lin}} M_k)$ and solves for the unknown operator $\hat{\rho}_{\text{lin}}$, then, apart from possible inconsistencies in the set of linear equations, the reconstructed operator $\hat{\rho}_{\text{lin}} \not\geq 0$ is often not a valid density operator anymore because some of its eigenvalues are negative. Hence, in such cases, this reconstruction called linear inversion delivers an unreasonable answer. It should be kept in mind that the inconsistencies can also be due to systematic errors, e.g. if the true measurements are aligned wrongly relative to the respective operator representation, but such effects are typically ignored [18].

Therefore, statistical state reconstruction relies on principles other than linear inversion. In general, these methods require the solution of a nonlinear optimization problem, which is much harder to solve than just a system of linear equations. For large system sizes this becomes, besides the exponential scaling of the number of settings and outcomes, a second major problem, again due to the exponential scaling of the number of parameters of the density operator. In fact for the current tomography record of eight ions in a trap [2], this actual reconstruction took even longer than the experiment itself (one week versus a couple of hours). Hence, feasible quantum state tomography schemes for large systems must, in addition to an efficient

measurement procedure, also possess an efficient state reconstruction algorithm; otherwise they are not scalable.

In this paper, we develop a scalable reconstruction algorithm for the proposed permutationally invariant tomography scheme [17]. It works for common reconstruction principles, including, among others, maximum likelihood and least squares methods. This scheme becomes possible once more by taking advantage of the symmetry of permutationally invariant states, which provides an efficient and operational way to store, characterize and even process those states. This method enables a large dimension reduction in the underlying optimization problem such that it gets into the feasible regime. The final low dimensional optimization is performed via nonlinear convex optimization which offers a great advantage in contrast to commonly used numerical routines, in particular regarding numerical stability and accuracy. Already a first prototype implementation of this algorithm allows state reconstruction for 20 qubits in a few minutes on a standard desktop computer.

The outline of the paper is as follows. Section 2 summarizes the background on permutationally invariant tomography and on statistical state reconstruction. The key method is explained in section 3 and highlighted via examples in section 4, which are generated by our current implementation. Section 5 collects all the technical details: the mentioned toolbox, more notes about convex optimization, additional information about the pretest or certificate and the measurement optimization, both addressed for large qubit numbers. Finally, in section 6 we conclude and provide an outlook on directions for further research.

2. Background

2.1. Permutationally invariant tomography

Permutationally invariant tomography has been introduced as a scalable reconstruction protocol for multi-qubit systems in [17]. It is designed for all states of the system that remain invariant under all possible interchanges of its different particles. Abstractly, such a permutationally invariant state ρ_{PI} of N qubits can be expressed in the form

$$\rho_{\text{PI}} = [\rho]_{\text{PI}} = \frac{1}{N!} \sum_{p \in S_N} V(p) \rho V(p)^\dagger, \quad (1)$$

where $V(p)$ is the unitary operator which permutes the N different subsystems according to the particular permutation p and the summation runs over all possible elements of the permutation group S_N . Many important states, such as the Greenberger–Horne–Zeilinger states or Dicke states, fall within this special class.

As shown in [17], full information on such states can be obtained by using in total $\binom{N+2}{N} = (N^2 + 3N + 2)/2$ different local binary measurement settings, while for each setting only the count rates of $(N + 1)$ different outcomes need to be registered. This finally leads to a cubic scaling in contrast to the exponential scaling of standard tomography schemes.

The measurement strategy that attains this number runs as follows: each setting is described by a unit vector $\hat{a} \in \mathbb{R}^3$ which defines associated eigenstates $|i\rangle_a$ of the trace-less operator $\hat{a} \cdot \vec{\sigma}$. Each part locally measures in this basis and registers the outcomes ‘0’ or ‘1’, respectively. The permutationally invariant part can be reconstructed from the collective outcomes, i.e. only the number of ‘0’ or ‘1’ results at the different parties matters but not the individual site information.

The corresponding coarse-grained measurements are given by

$$M_k^a = \sum_{p'} V(p') |0\rangle_a \langle 0|^{\otimes k} \otimes |1\rangle_a \langle 1|^{\otimes N-k} V(p')^\dagger, \quad (2)$$

$$= \binom{N}{k} [|0\rangle_a \langle 0|^{\otimes k} \otimes |1\rangle_a \langle 1|^{\otimes N-k}]_{\text{PI}} \quad (3)$$

with $k = 0, \dots, N$, and where the summation p' is over all permutations that give distinct terms. In total, one needs the above stated number of different settings \hat{a} . These settings can be optimized in order to minimize the total variance which provides an advantage in contrast to random selection.

In addition to this measurement strategy there is also a pretest which estimates the ‘closeness’ of the true, unknown state with respect to all permutationally invariant states from just a few measurement results [17]. This provides a way to test in advance whether permutationally invariant tomography is a good method for the unknown state.

Restricting oneself to the permutationally invariant part of a density operator has already been discussed in the literature: for example, for spins in a Stern–Gerlach experiment [19] or in terms of the polarization density operator [20–22]. Here, due to the restricted class of possible measurements, only the permutationally invariant part of, in principle distinguishable, particles is accessible [21, 22]. This is a strong conceptual difference compared to permutationally invariant tomography where one intentionally constrains oneself to this invariant part. Nevertheless, it should be emphasized that the employed techniques are similar.

2.2. Statistical state reconstruction

Since standard linear inversion of the observed data typically results in unreasonable estimates as explained in the introduction, one employs other principles for actual state reconstruction. In general, one uses a certain fit function $F(\rho)$ that penalizes deviations between the observed frequencies f_k and the true probabilities predicted by quantum mechanics $p_k(\rho) = \text{tr}(\rho M_k)$ if the state of the system is ρ . The reconstructed density operator $\hat{\rho}$ is then given by the (often unique) state that minimizes this fit function,

$$\hat{\rho} = \arg \min_{\rho \geq 0} F(\rho); \quad (4)$$

hence the reconstructed state is precisely the one which best fits the observed data. Since the optimization explicitly restricts itself to physical density operators this ensures validity of the final estimate $\hat{\rho}$ in contrast to linear inversion. Depending on the functional form of the fit function, different reconstruction principles are distinguished. A list of the most common choices is provided in table 1.

The presumably best-known and most widely employed method is called the maximum likelihood principle [23]. Given a set of measured frequencies, f_k , the maximum likelihood state, $\hat{\rho}_{\text{ml}}$, is exactly the one with the highest probability of generating these data. Other common fit functions, usually employed in photonic state reconstruction, are different variants of least squares [4, 24] that originate from the likelihood function using Gaussian approximations for the probabilities. There, this is often called maximum likelihood principle, but we distinguish these, indeed different, functions here explicitly. Typically, the weights in the least squares function are set to $w_k = 1/f_k$ because f_k represents an estimate of the variance in a multinomial distribution,

4. Quantum state tomography

6

IOP Institute of Physics Φ DEUTSCHE PHYSIKALISCHE GESELLSCHAFT

Table 1. Common reconstruction principles and their corresponding fit functions $F(\rho)$ used in the optimization given by equation (4); see text for further details.

Reconstruction principle	Fit function $F(\rho)$
Maximum likelihood [23]	$-\sum_k f_k \log[p_k(\rho)]$
Least squares [24]	$\sum_k w_k [f_k - p_k(\rho)]^2, w_k > 0$
Free least squares [4]	$\sum_k 1/p_k(\rho) [f_k - p_k(\rho)]^2$
Hedged maximum likelihood [25]	$-\sum_k f_k \log[p_k(\rho)] - \beta \log[\det(\rho)], \beta > 0$

of the free least squares principle. However, this leads to a strong bias if the count rates are extremal; for example, if one of the outcomes is never observed, this method naturally leads to difficulties. A method to circumvent this is given by the free least squares function [4] or using improved error analysis for rare events. Let us stress that all these principles have the property that if linear inversion delivers a valid estimate $\hat{\rho}_{\text{lin}} \geq 0$, it is also the estimate given by these reconstruction principles¹⁰.

Finally, hedged maximum likelihood [25] represents a method that circumvents low rank state estimates. Via this, one obtains more reasonable error bars using parametric bootstrapping methods [26]; for other error estimates, see the recently introduced confidence regions for quantum states [27, 28]. In principle, many more fit functions are possible, such as generic loss functions [29], but considering these is beyond the scope of this work.

3. Method

From the above, it is apparent that permutationally invariant state reconstruction requires the solution of

$$\hat{\rho}_{\text{PI}} = \arg \min_{\rho_{\text{PI}} \geq 0} F(\rho_{\text{PI}}) \quad (5)$$

for the preferred fit function. This large-scale optimization becomes feasible along the following lines.

Firstly, one reduces the dimensionality of the underlying optimization problem because one cannot work with full density operators anymore. This requires an operational way to characterize permutationally invariant states $\rho_{\text{PI}} \geq 0$ over which the optimization is performed, and additionally, demands an efficient way to compute probabilities $p_k(\rho_{\text{PI}})$ which appear in the fit function. Secondly, one needs a method for performing the final optimization. We employ convex optimization for this.

3.1. Reduction of the dimensionality

This reduction relies on an efficient toolbox to handle permutationally invariant states, which exploits the particular symmetry. Here we will explain this method and the final structure; for

¹⁰ For the least squares fit functions this follows because $F = 0$ in this case and clearly $F \geq 0$ for those functions. In the case of the likelihood it follows from positivity of the classical relative entropy between probability distributions.

more details see section 5.1. These techniques are well proven and established; we employ and adapt them here for the permutationally invariant tomography scheme such that we finally reach state reconstruction of larger qubits.

The methods of this toolbox are obtained via the concept of spin coupling that describes how individual, distinguishable spins couple to a combined system if they become indistinguishable. Since we deal with qubits we only need to focus on spin-1/2 particles. In the simplest case, two spin-1/2 particles can couple to a spin-1 system, called the triplet, if both spins are aligned symmetric, or to a spin-0 state, the singlet, if the spins are aligned anti-symmetric. Abstractly, this can be denoted as $\mathbb{C}^2 \otimes \mathbb{C}^2 = \mathbb{C}^3 \oplus \mathbb{C}^1$. If one considers now three spins, then of course all spins can point in the same direction giving a total spin-3/2 system. There is also a spin-1/2 system possible if two particles form already a spin-0 and the remaining one stays untouched. This can be achieved, however, by more than one possibility, in fact by two inequivalent choices¹¹, and is expressed by $\mathbb{C}^2 \otimes \mathbb{C}^2 \otimes \mathbb{C}^2 = \mathbb{C}^4 \oplus (\mathbb{C}^2 \otimes \mathbb{C}^2)$.

This scheme can be extended to N spin-1/2 particles to obtain the following decomposition of the total Hilbert space,

$$\mathcal{H} = (\mathbb{C}^2)^{\otimes N} = \bigoplus_{j=j_{\min}}^{N/2} \mathcal{H}_j \otimes \mathcal{K}_j, \quad (6)$$

where the summation runs over different total spin numbers $j = j_{\min}, j_{\min} + 1, \dots, N/2$ starting from $j_{\min} \in \{0, 1/2\}$, depending on whether N is even or odd. Here, \mathcal{H}_j are called the spin Hilbert spaces with dimensions $\dim(\mathcal{H}_j) = 2j + 1$, while \mathcal{K}_j are referred to as multiplicative spaces that account for the different possibilities to obtain a spin- j state. They are, generally, of a much larger dimension, cf equation (22).

Permutationally invariant states have a simpler form on this Hilbert space decomposition, namely

$$\rho_{\text{PI}} = \bigoplus_{j=j_{\min}}^{N/2} p_j \rho_j \otimes \frac{\mathbb{1}}{\dim(\mathcal{K}_j)}, \quad (7)$$

with density operators ρ_j called spin states and according probabilities p_j . Thus a permutationally invariant density operator only contains non-trivial parts on the spin Hilbert spaces while carrying no information on the multiplicative spaces. Note, further, that there are no coherences between different spin states. This means that any permutationally invariant state can be parsed into a block structure as schematically represented in figure 1. The main diagonal is made up of unnormalized spin states $\tilde{\rho}_j = p_j \rho_j / \dim(\mathcal{K}_j)$, which appear several times, the number being equal to the dimension of the corresponding multiplicative space. This block-decomposition represents a natural way to treat permutationally invariant states and has, for example, already been employed in the aforementioned related works on permutationally invariant tomography [19–22] but also in other contexts [30–33].

This structure shows that if we work with permutationally invariant states, we do not need to consider the full density operator but rather that it is sufficient to deal only with this ensemble

¹¹ More precisely, all states of the form $|\psi\rangle = V(p)|\psi^-\rangle \otimes |0\rangle$, with p being any possible permutation, are states of total spin $j = 1/2$ and projection $m = 1/2$ to the collective spin operators $J_i = \sum_{n=1}^3 \sigma_{i,n}/2$, $\sigma_{i,n}$ being the corresponding Pauli operator on qubit n . However, as can be checked, these states only span a two-dimensional subspace.

4. Quantum state tomography

8

IOP Institute of Physics Φ DEUTSCHE PHYSIKALISCHE GESELLSCHAFT

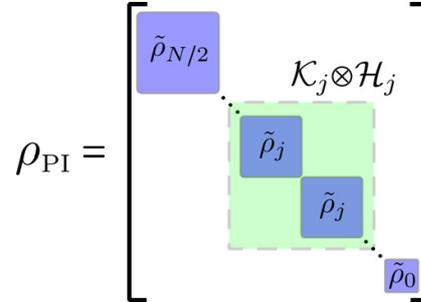


Figure 1. Block decomposition for a generic permutationally invariant state as given by equation (7) with $\tilde{\rho}_j = p_j \rho_j / \dim(\mathcal{K}_j)$. Note that we interchanged the spin- and multiplicative spaces.

of spin states. Therefore we identify from now on

$$\rho_{\text{PI}} \iff p_j \rho_j, \quad j = j_{\min}, j_{\min} + 1, \dots, N/2. \quad (8)$$

This provides already an efficient way to store and to visualize such states. More importantly, it also provides an operational way to characterize valid states, since any permutationally invariant operator ρ_{PI} represents a true state if and only if all these spin operators ρ_j are density operators and p_j a probability distribution. This is in contrast to the generalized Bloch vector employed in the original proposal of permutationally invariant tomography [17] given by equation (52), which also provides efficient storage and processing of permutationally invariant states, but where Bloch vectors of physical states are not as straightforward to characterize.

Via this identification one can demonstrate once more the origin of the cubic scaling of the permutationally invariant tomography scheme. The largest spin state is of dimension $N + 1$ which requires parameters of the order of N^2 for characterization. Since one has of the order of N of these states, this results in a cubic scaling.

Fixing the ensemble of spin states as parameterization, it is now necessary to obtain an efficient procedure to compute the probabilities $p_k^a(\rho_{\text{PI}})$ for the optimized measurement scheme. This is achieved as follows: let us first stress that a similar block decomposition to that given by equation (7) holds for all permutationally invariant operators. Hence also the measurements M_k^a given by equation (2) can be cast into this form. Using the convention

$$M_k^a = \bigoplus_{j=j_{\min}}^{N/2} M_{k,j}^a \otimes \mathbb{1} \quad (9)$$

leads to

$$\text{tr}(\rho_{\text{PI}} M_k^a) = \sum_{j=j_{\min}}^{N/2} p_j \text{tr}(\rho_j M_{k,j}^a). \quad (10)$$

Therefore the problem is shifted to the computation of the spin- j operators $M_{k,j}^a$ for each setting \hat{a} . As we show in proposition 5.2 below, using the idea that measurements can be transformed into each other by a local operation $U_a|i\rangle = |i\rangle_a$, this provides the relation

$$M_{k,j}^a = W_j^a M_{k,j}^{e_3} W_j^{a,\dagger}. \quad (11)$$

New Journal of Physics **14** (2012) 105001 (<http://www.njp.org/>)

Here $M_{k,j}^{e_3}$ corresponds to the measurement in the standard basis (that need to be computed once) and W_j^a is a unitary transformation determined by the rotation U_a . This connection is given by

$$U_a = \exp\left(-i \sum_l t_l \sigma_l / 2\right) \implies W_j^a = \exp\left(-i \sum_l t_l S_{l,j}\right), \quad (12)$$

where $S_{l,j}$ stands for the spin operators in dimension $2j + 1$. This finally provides an efficient way to compute probabilities.

3.2. Optimization

As a second step one still needs to cope up with the optimization itself. Although there are different numerical routines for statistical state reconstruction like maximum likelihood [34] or least squares [4, 35], we prefer nonlinear convex optimization [36] to obtain the final solution. Quantum state reconstruction problems are known to be convex [35, 37], but convex optimization has hardly been used for this task. However, convex optimization has several advantages: first of all it is a systematic approach that works for any convex fit function, including maximum likelihood and least squares. In contrast to other algorithms such as the fixed-point algorithm proposed in [34], it gives a precise stopping condition via an appropriate error control (see, however, [38]) and still exploits all the favourable, convex, structure in comparison with re-parameterization ideas as in [4]. Moreover, it is guaranteed to find the global optimum and the accuracy obtained is typically much higher than that obtained with other methods.

Quantum state reconstruction as defined via equation (4) can be formulated as a convex optimization problem as follows: all fit functions listed in table 1 are convex on the set of states. Via a linear parameterization of the density operator $\rho(x) = \mathbb{1} / \dim(\mathcal{H}) + \sum x_i B_i$, using an appropriate operator basis B_i such that normalization is fulfilled directly, the required optimization problem becomes

$$\begin{aligned} \min_x F[\rho(x)] \\ \text{s.t. } \rho(x) = \frac{\mathbb{1}}{\dim(\mathcal{H})} + \sum_i x_i B_i \geq 0, \end{aligned} \quad (13)$$

with a convex objective function $F(x) = F[\rho(x)]$ and a linear matrix inequality as the constraint, i.e. precisely the structure of a nonlinear convex optimization problem [36]. For permutationally invariant states, one uses $\rho(x) = \oplus_j \bar{\rho}_j(x)$ with $\bar{\rho}_j = p_j \rho_j$ by using an appropriate block-diagonal operator basis B_i ; therefore we continue this discussion with the more general form.

The optimization given by equation (13) can be performed, for instance, with the help of a barrier function [36]¹². Rather than considering the constrained problem, one solves the unconstrained convex task given by

$$\min_x F[\rho(x)] - t \log[\det \rho(x)], \quad (14)$$

¹² Let us stress that both least squares options can be parsed into a simpler convex problem, called a semidefinite program, as shown in, for instance, [24, 35], but that this does not work with the true maximum likelihood function to our best knowledge.

4. Quantum state tomography

where the constraint is now directly included in the objective function. This so-called barrier term acts precisely as its name suggests: if one tries to leave the strictly feasible set, i.e. all parameters x that satisfy $\rho(x) > 0$, one always reaches a point where at least one of the eigenvalues vanish. Since the barrier term is large within this neighbourhood, in fact singular at the crossing, it penalizes points close to the border and thus ensures that one searches for an optimum well inside the region where the constraint is satisfied. The penalty parameter $t > 0$ plays the role of a scaling factor. If it becomes very small the effect of the barrier term becomes negligible within the strictly feasible set and only remains at the border. Therefore a solution of equation (14) with a very small value of t provides an excellent approximation to the real solution. As shown in section 5.2 this statement can be made more precise by

$$F[\rho(x_{\text{sol}}^t)] - F[\rho(x_{\text{sol}})] \leq t \dim(\mathcal{H}), \quad (15)$$

which follows from convexity and which relates the true solution x_{sol} of the original problem given by equation (13) to the solution x_{sol}^t of the unconstrained problem with penalty parameter t . This condition represents the above-mentioned error control and serves as a stopping condition, i.e. as a quantitative error bound for a given t . Note that for permutationally invariant tomography $\dim(\mathcal{H})$ is not the dimension of the true N -qubit Hilbert space but instead the dimension of $\rho(x) = \bigoplus_j \rho_j(x)$, i.e. $\sum_{j=j_{\text{min}}} (2j+1) = (N+1)(N+2j_{\text{min}}+1)/4$ which increases only quadratically.

Small penalty parameters are approached by an iterative process: for a given starting point x_{start}^n and a certain value of the parameter $t_n > 0$, one solves equation (14). Its solution will be the starting point for $x_{\text{start}}^{n+1} = x_{\text{sol}}^n$ for the next unconstrained optimization with a lower penalty parameter $t_{n+1} < t_n$. As the starting point for the first iteration, we employ $t_0 = 1$ and the point x_{start}^0 corresponding to the totally mixed state. This procedure is repeated until one has reached small enough penalty parameters. The penalty parameter is decreased step-wise. Then each unconstrained problem can be solved very efficiently since one starts already quite close to the true solution.

Let us point out that via the above-mentioned barrier method, one additionally obtains solutions to the hedged state reconstruction with $\beta = t$ since the unconstrained problem given by equation (14) is precisely the fit function of the hedged version of table 1.

Finally, for comparison purposes, we would also like to mention the iterative fixed-point algorithm of [39]; for a modification see [40]. It is designed for maximum likelihood estimation and is straightforward to implement, since it only requires matrix multiplication; however, it has deficits regarding control and accuracy. For permutationally invariant tomography the algorithm can be adapted as follows: given a valid iterate ρ_{PI}^n characterized by the ensemble of spin states $\bar{\rho}_j^n = p_j^n \rho_j^n$, one evaluates the probabilities $p_k^n(\rho_{\text{PI}}^n)$ using equation (10). Next, one computes the operators

$$R_j^n = \sum_{a,k} \frac{f_k^a}{p_k^n(\rho_{\text{PI}}^n)} M_{k,j}^a, \quad (16)$$

which determine the next iterate $\bar{\rho}_j^{n+1} = R_j^n \bar{\rho}_j^n R_j^{n\dagger} / \mathcal{N}$ with $\mathcal{N} = \sum_j \text{tr}(R_j^n \bar{\rho}_j^n R_j^{n\dagger})$. This iteration is started, for example, from the totally mixed state and repeated until a sufficiently good solution is obtained.

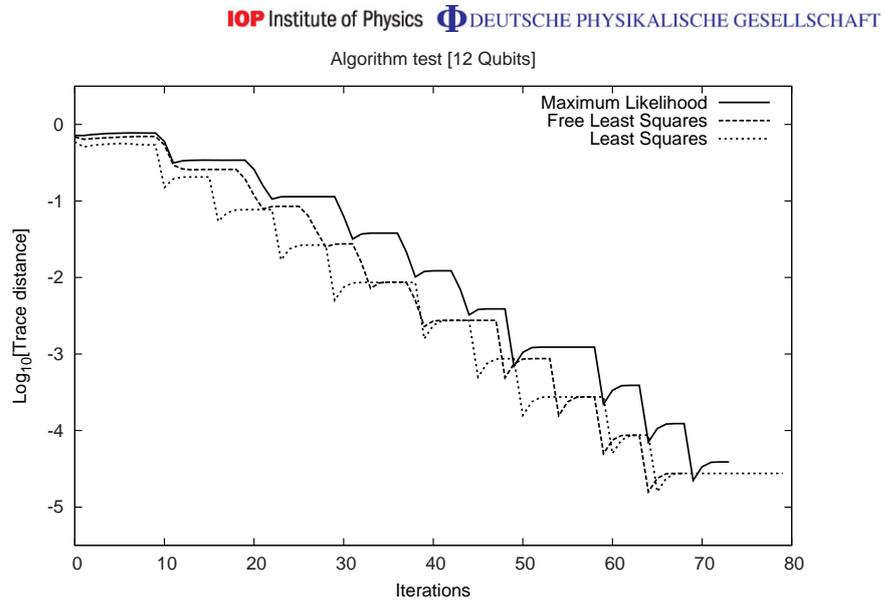


Figure 2. Trace distance between the analytic solution and the estimate after n algorithm steps for the three most common reconstruction principles from table 1.

4. Examples

The two methods from the previous section are employed in a prototype implementation under MATLAB, which already enables state reconstruction of about 20 qubits on a standard desktop computer.

The current algorithm is tested along the following lines: for a randomly generated permutationally invariant state $\rho_{\text{PI}}^{\text{true}}$ we compute the true probabilities $p_{k,\text{true}}^a$ for the chosen measurement settings. Rather than sampling we set the observed frequencies equal to this distribution, i.e. $f_k^a = p_{k,\text{true}}^a$. In this way linear inversion would return the original state; hence also every other reconstruction principle from table 1 has this state as the solution. We now start the algorithm and compare the trace distance between the analytic solution $\rho_{\text{PI}}^{\text{true}}$ and the state after n iterations ρ_{PI}^n . This distance $\frac{1}{2}|\text{tr}[\rho_{\text{PI}}^{\text{true}} - \rho_{\text{PI}}^n]|$ quantifies the probability with which the two states, the true analytic solution and the iterate after n steps in the algorithm, could be distinguished [41].

A typical representative of this process is depicted in figure 2 for 12 qubits using optimized settings. The randomly generated state $\rho_{\text{PI}}^{\text{true}}$ was chosen to lie at the boundary of the state space since such rank-reduced solutions better resemble the case of state reconstruction of real data. More precisely, each spin state of the true density operator is given by a pure state $\rho_j^{\text{true}} = |\psi_j\rangle\langle\psi_j|$ chosen according to the Haar measure, while the p_j are selected by the symmetric Dirichlet distribution with concentration measure $\alpha = 1/2$ [42]. As is apparent the algorithm behaves similarly for all three reconstruction principles and rapidly obtains a good solution after about 70 iterations. The steps in this plot are points where the penalty parameter is reduced by a factor of 10 starting from $t = 1$ and decreased down to $t = 10^{-10}$. The slight rise

4. Quantum state tomography

12

IOP Institute of Physics Φ DEUTSCHE PHYSIKALISCHE GESELLSCHAFT

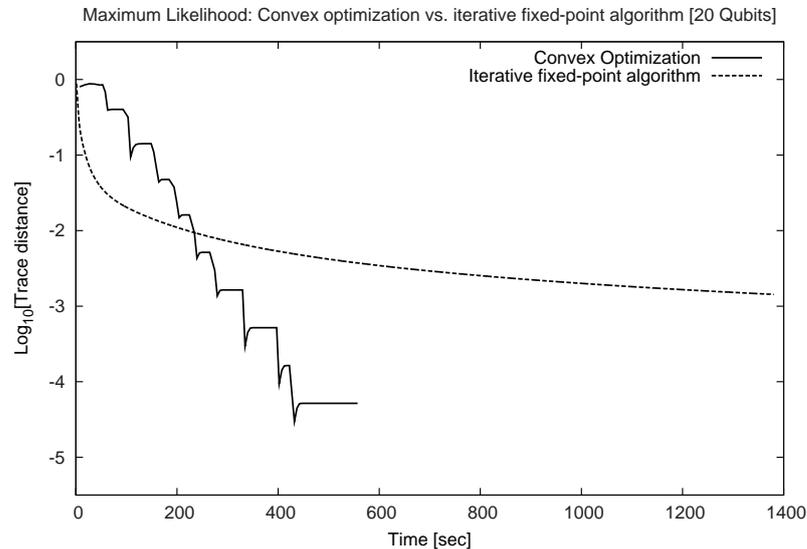


Figure 3. Comparison of the maximum likelihood principle between convex optimization and the iterative fixed-point algorithm for the described testing procedure with respect to accuracy and algorithm time.

after these points comes from the fact that we plot the trace distance and not the actual function (fit function plus penalty term) that is minimized.

Figure 3 shows a similar comparison for the maximum likelihood reconstruction of 20 qubits but now plotted versus algorithm time¹³. For comparison we include the performance of the iterative fixed-point algorithm, which requires much more iterations in general (3000 in this case versus about 90 for convex optimization). Let us emphasize that a similar behaviour between these two algorithms appears also for smaller qubit numbers. As one can see, convex optimization delivers a faster and in particular more accurate solution. In contrast, the iterative fixed-point algorithm shows a bad convergence rate although it initially starts off better. This was one of the main reasons for us to switch to convex optimization.

The current algorithm times of this test are listed in table 2 which are averaged over 50 randomly generated states. Thus already this prototype implementation enables state reconstruction of larger qubit numbers in moderate times. The small time difference between reconstruction principles is because least squares as a quadratic fit function provides some advantages in the implementation. More details of this difference are given in section 5.

Table 2 also contains the algorithm times for reconstructions using simulated frequencies $f_k^a = n_k^a / N_r$. For each setting they are deduced from the count rates n_k^a sampled from a multinomial distribution using the true event distribution $p_{k,\text{true}}^a$ and $N_r = 1000$ repetitions. The true probabilities correspond to the same states as already employed in the algorithm test. From the table, one observes that state reconstruction for data with count statistics requires only slightly more time than the algorithm test with the correct probabilities. We attribute this to the fact that a few more iterations are typically required in order to achieve the desired accuracy.

¹³ All simulations were performed on an Intel Core i5-650, 3.2 GHz, 8 GB RAM using MATLAB 7.12.

Table 2. Current performance of the convex optimization algorithm on the described test procedure and on frequencies from simulated experiments; free least squares provides similar results to the maximum likelihood principle.

	$N = 8$	$N = 12$	$N = 16$	$N = 20$
Maximum likelihood				
Algorithm test	8.5 s	47 s	2.7 min	9.2 min
Simulated experiment	9.2 s	48 s	2.9 min	9.3 min
Least squares				
Algorithm test	8.4 s	39 s	2.5 min	6 min
Simulated experiment	9.2 s	43 s	2.7 min	6.7 min

Finally, let us perform the reconstruction of a simulated experiment of $N = 14$ qubits. Suppose that one intends to create a Dicke state $|D_{k,N}\rangle$ as given by equation (26) for some k and N , but that the preparation suffers from some imperfections such that at best one can prepare states of the form

$$\rho_{\text{dicke mix}} = \sum_{l=0}^N \binom{N}{l} p^l (1-p)^{N-l} |D_{l,N}\rangle \langle D_{l,N}|, \quad (17)$$

where $p = 0.6$ characterizes some asymmetry. As the true state prepared in the experiment we now model some further imperfection in the form of an additional small misalignment $U^{\otimes N}$, with $U = \exp(-i\theta\sigma_y/2)$, $\theta = 0.2$, and some permutationally invariant noise σ_{PI} (chosen via the aforementioned method but using Hilbert–Schmidt instead of the Haar measure), i.e.

$$\rho_{\text{true}} = 0.6 U^{\otimes N} \rho_{\text{dicke mix}} U^{\dagger \otimes N} + 0.4 \sigma_{\text{PI}}. \quad (18)$$

The frequencies are obtained via sampling from the state given by equation (18) using intentionally only $N_r = 200$ repetitions per setting (to see some differences). Finally, we reconstruct the state according to the maximum likelihood principle. Figure 4 shows the tomography bar plots of one of these examples for the largest spin state $p_j \rho_j$, $j = N/2 = 7$ for both states. Although this state might be artificial, this example should highlight once more that this state reconstruction algorithm works also for realistic data and for qubit sizes where clearly any non-tailored state reconstruction scheme would fail. Moreover, it demonstrates that the spin ensemble $p_j \rho_j$ represents a very convenient graphical representation of such states (compared to a $2^{14} \times 2^{14}$ matrix in this case).

5. Details

5.1. Reduction

Let us first give more details regarding the reduction step. This starts by recalling a group theoretic result summarized in the next section, which is then used to show how the stated simplifications with respect to states and measurements are obtained.

4. Quantum state tomography

14

IOP Institute of Physics Φ DEUTSCHE PHYSIKALISCHE GESELLSCHAFT

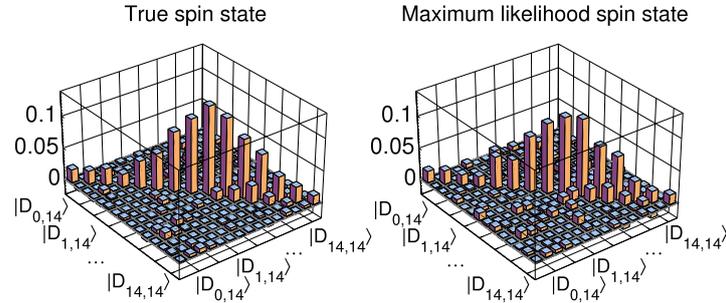


Figure 4. The real part of the true and reconstructed (according to maximum likelihood) largest spin state ensemble $p_j \rho_j$, $j = N/2 = 7$ using the optimal measurement setting. The basis is given by the Dicke basis $|D_{k,14}\rangle$, cf equation (26).

5.1.1. Background. Consider the following two unitary representations defined on the N qubit Hilbert space: the permutation or symmetric group $V(p)$ which is defined by their action onto a standard tensor product basis by $V(p)|i_1, \dots, i_N\rangle = |i_{p^{-1}(1)}, \dots, i_{p^{-1}(N)}\rangle$ according to the given permutation p , and the tensor product representation $W(U) = U^{\otimes N}$ of the special unitary group. A result known as the Schur–Weyl duality [43, 44] states that the action of these two groups is dual, which means that the total Hilbert space can be divided into blocks on which the two representations commute. More precisely, one has

$$(\mathbb{C}^2)^{\otimes N} = \bigoplus_{j=j_{\min}}^{N/2} \mathcal{H}_j \otimes \mathcal{K}_j, \quad (19)$$

$$V(p) = \bigoplus_{j=j_{\min}}^{N/2} \mathbb{1} \otimes V_j(p), \quad (20)$$

$$W(U) = \bigoplus_{j=j_{\min}}^{N/2} W_j(U) \otimes \mathbb{1}. \quad (21)$$

Here V_j and W_j are the respective irreducible representations, and $j_{\min} \in \{0, 1/2\}$ depending on whether N is even or odd. The dimensions of the appearing Hilbert spaces are $\dim(\mathcal{H}_j) = 2j + 1$ and

$$\dim(\mathcal{K}_j) = \binom{N}{N/2-j} - \binom{N}{N/2-j-1} \quad (22)$$

for all $j < N/2$ and $\dim(\mathcal{K}_{N/2}) = 1$. Let us note that equation (20) already ensures the block-diagonal structure of permutationally invariant operators, while equation (21) becomes important for the measurement computation.

A basis of the Hilbert space $\mathcal{H}_j \otimes \mathcal{K}_j$ is formed by the spin states $|j, m, \alpha\rangle = |j, m\rangle \otimes |\alpha_j\rangle$ with $m = -j, -j + 1, \dots, j$ and $\alpha_j = 1, \dots, \dim(\mathcal{K}_j)$. These are obtained by starting with the states having the largest spin number $m = j$, which are given by

$$|j, j, 1\rangle = |0\rangle^{\otimes 2j} \otimes |\psi^-\rangle^{\otimes N/2-j}, \quad (23)$$

$$|j, j, \alpha\rangle = \sum_p c_{j,p} V(p) |j, j, 1\rangle, \quad (24)$$

for all $\alpha \geq 2$. The coefficients $c_{j,p}$ must ensure that the states $|j, j, \alpha\rangle$ are orthogonal; otherwise their choice is completely free since the detailed structure of different α 's is not important. The full basis is obtained by subsequently applying the ladder operator $J_- = \sum_{n=1}^N \sigma_{-,n}$ to decrease the spin number m . Here $\sigma_{-,n}$ refers to the operator with $\sigma_- = (\sigma_x - i\sigma_y)/2$ on the n th qubit and identity on the rest. Thus in total the basis becomes

$$|j, m, \alpha\rangle = \mathcal{N} J_-^{j-m} |j, j, \alpha\rangle, \quad (25)$$

with appropriate normalizations \mathcal{N} . Note that the subspace corresponding to the highest spin number $j = N/2$ is also called the symmetric subspace, which contains many important states such as the Greenberger–Horne–Zeilinger or Dicke states, which using the spin states read as

$$|\text{GHZ}\rangle = \frac{1}{\sqrt{2}} (|0\rangle^{\otimes N} + |1\rangle^{\otimes N}) = \frac{1}{\sqrt{2}} (|N/2, N/2, 1\rangle + |N/2, -N/2, 1\rangle), \quad (26)$$

$$|D_{k,N}\rangle = \mathcal{N} [|1\rangle^{\otimes k} \otimes |0\rangle^{\otimes N-k}]_{\text{PI}} = |N/2, N/2 - k, 1\rangle. \quad (27)$$

5.1.2. Permutationally invariant states and measurement operators. Let us now employ this result in order to derive a generic form for permutationally invariant states; we give the proof for completeness.

Proposition 5.1 (Permutationally invariant states). *Any permutationally invariant state of N qubits ρ_{PI} defined via equation (1) can be written as*

$$\rho_{\text{PI}} = \bigoplus_{j=j_{\min}}^{N/2} p_j \rho_j \otimes \frac{\mathbb{1}}{\dim(\mathcal{K}_j)}; \quad (28)$$

hence it is fully characterized by the ensemble $p_j \rho_j$. Moreover, ρ_{PI} is a density operator if and only if all ρ_j are density operators and p_j a probability distribution.

Proof. The proposition follows using the representation $V(p)$ given by equation (20) in the definition of the states, cf equation (1), and then applying Schur's lemma [43, 44]. This lemma states that any linear operator A from \mathcal{K}_j to \mathcal{K}_i which commutes with all elements p of the group $V_i(p)A = AV_j(p)$ must either be zero if i and j label different irreducible representations or $A = c\mathbb{1}$ if they are unitarily equivalent. Since $A_{\text{PI}} = 1/N! \sum_p V_i(p)A V_j(p)^\dagger$ fulfils this relation, one obtains

$$\frac{1}{N!} \sum_p V_i(p)A V_j(p)^\dagger = \delta_{ij} \text{tr}(A) \frac{\mathbb{1}}{\dim(\mathcal{K}_j)}. \quad (29)$$

The normalization can be checked taking the trace on both sides. Adding appropriate identities provides

$$\frac{1}{N!} \sum_p \mathbb{1} \otimes V_i(p)B\mathbb{1} \otimes V_j(p)^\dagger = \delta_{ij} \text{tr}_{\mathcal{K}_j}(B) \otimes \frac{\mathbb{1}}{\dim(\mathcal{K}_j)}, \quad (30)$$

where B should now be a linear operator from $\mathcal{H}_j \otimes \mathcal{K}_j$ to $\mathcal{H}_i \otimes \mathcal{K}_i$.

4. Quantum state tomography

16

IOP Institute of Physics Φ DEUTSCHE PHYSIKALISCHE GESELLSCHAFT

Finally, let P_j denote the projector onto $\mathcal{H}_j \otimes \mathcal{K}_j$, and using equation (30) yields

$$\rho_{\text{PI}} = \frac{1}{N!} \sum_p V(p) \rho V(p)^\dagger = \sum_{i,i',j,j'} \frac{1}{N!} \sum_p P_i V(p) P_{i'} \rho P_j V(p)^\dagger P_{j'} \quad (31)$$

$$= \sum_{i,j} \left\{ \frac{1}{N!} \sum_p P_i [\mathbb{1} \otimes V_i(p)] P_i \rho P_j [\mathbb{1} \otimes V_j(p)]^\dagger P_j \right\} \quad (32)$$

$$= \sum_{i,j} P_i \left[\delta_{ij} \text{tr}_{\mathcal{K}_j}(P_i \rho P_j) \otimes \frac{\mathbb{1}}{\dim(\mathcal{K}_j)} \right] P_j \quad (33)$$

$$= \bigoplus_j \text{tr}_{\mathcal{K}_j}(P_j \rho P_j) \otimes \frac{\mathbb{1}}{\dim(\mathcal{K}_j)}, \quad (34)$$

which provides the general structure.

The state characterization part follows because positivity of a block matrix is equivalent to positivity of each block. \square

Next let us concentrate on the measurement part. Although the block decomposition follows already from the previous proposition, it is here more important to obtain an efficient computation of each measurement block for the selected setting.

Proposition 5.2 (Measurement operators). *The POVM elements M_k^a as defined in equation (2) for any local setting $\hat{a} \in \mathbb{R}^3$ can be decomposed as $M_k^a = \bigoplus_j M_{k,j}^a \otimes \mathbb{1}$ with*

$$M_{k,j}^a = W_j(U_a) M_{k,j}^{e_3} W_j(U_a)^\dagger. \quad (35)$$

The unitary is given by $W_j(U_a) = \exp(-i \sum_l t_l S_{l,j})$ using the spin operators $S_{l,j}$ in dimension $2j+1$, while the coefficients t_l are determined by $U_a = \exp(-i \sum_l t_l \sigma_l / 2)$ which satisfies $\hat{a} \cdot \vec{\sigma} = U_a \sigma_z U_a^\dagger$. For the measurement in the standard basis $\hat{a} = \hat{e}_3$ one obtains

$$M_{k,j}^{e_3} = |j, N/2 - k\rangle \langle j, N/2 - k| \quad (36)$$

if $-j \leq N/2 - k \leq j$ and zero otherwise.

Proof. The basic idea is to consider the measurement in an arbitrary local basis \hat{a} by a rotation followed by the collective measurement in the standard basis. The block decomposition is obtained as follows:

$$M_k^a = U_a^{\otimes N} M_k^{e_3} U_a^{\dagger \otimes N} = W(U_a) \left[\bigoplus_j M_{k,j}^{e_3} \otimes \mathbb{1} \right] W(U_a)^\dagger \quad (37)$$

$$= \bigoplus_j W_j(U_a) M_{k,j}^{e_3} W_j(U_a)^\dagger \otimes \mathbb{1}. \quad (38)$$

The first step holds because U_a satisfies $|i\rangle_a \langle i| = U_a |i\rangle \langle i| U_a^\dagger$, while the block decomposition of the standard basis measurement $M_k^{e_3}$ is employed afterwards. In the last part one uses the tensor product representation given by equation (21).

Since one knows that W_j is irreducible it can be uniquely written in terms of its Lie algebra representation dW_j as $W_j(U_a) = W_j(e^{-iX}) = e^{-idW_j(X)}$, which is given by the spin operators in this case, i.e. $dW_j(\sigma_l/2) = S_{l,j}$ [45].

Thus we are left to compute the measurement blocks $M_{k,j}^{e_3}$ for the standard basis. Note it is sufficient to evaluate $M_{k,j}^{e_3} \otimes |1_j\rangle\langle 1_j|$ such that one can employ the spin basis states $|j, m, 1\rangle$ as introduced in section 5.1.1. At first, note that $M_k^{e_3}$ exactly contains k projections onto $|0\rangle$, while each basis state $|j, m, 1\rangle$ possesses $(N/2+m)$ zeros. Therefore one obtains $M_{k,j}^{e_3}|j, m, 1\rangle \propto \delta_{k, N/2+m}|j, m, 1\rangle$. Since each POVM has to resolve the identity this is only possible if each $M_{k,j}^{e_3}$ is the stated rank-1 projector on the basis states. \square

Finally one still needs to express $U_a = \exp(-i \sum_l t_l \sigma_l/2)$ for the chosen setting $\hat{a} \in \mathbb{R}^3$. Since this can be related to a familiar rotation [45] these coefficients can be expressed as $t_l = (\theta \hat{n})_l$ via a rotation about an angle θ around the axis \hat{n} . Since this rotation should turn \hat{e}_3 into \hat{a} , its parameters are given by

$$\hat{n} = \frac{\hat{e}_3 \times \hat{a}}{\|\hat{e}_3 \times \hat{a}\|_2}, \tag{39}$$

$$\theta = \arccos(\hat{e}_3 \cdot \hat{a}), \tag{40}$$

and $\hat{n} = \hat{e}_1$ (or any other orthogonal vector) if $\hat{a} = \pm \hat{e}_3$.

5.2. Convex optimization

In this part, we collect some more details of the described convex optimization algorithm; for complete coverage see the book [36].

Each unconstrained optimization given by equation (14) is solved via a damped Newton algorithm. The minimization of $f(x) = F[\rho(x)] - t \log[\det \rho(x)]$ is obtained by an iterative process. In order to determine a search direction at a given iterate x^n , one minimizes the quadratic approximation

$$f(x^n + \Delta x) \approx f(x^n) + \nabla f(x^n)^T \Delta x + \frac{1}{2} \Delta x^T \nabla^2 f(x^n) \Delta x. \tag{41}$$

This reduces to solving a linear set of equations called the Newton equation

$$\nabla^2 f(x^n) \Delta x_{\text{nt}} = -\nabla f(x^n), \tag{42}$$

which determines the search direction Δx_{nt} . The step length s for the next iterate $x^{n+1} = x^n + s \Delta x_{\text{nt}}$ is chosen by a backtracking line search. Here one picks the largest $s = \max_{k \in \mathbb{N}} \beta^k$ with $\beta \in (0, 1)$ such that the iterate stays feasible $\rho(x^{n+1}) > 0$ and that the function value decreases sufficiently, i.e. $f(x^{n+1}) \leq f(x^n) + \alpha s \nabla f(x^n)^T \Delta x_{\text{nt}}$ with $\alpha \in (0, 0.5)$. The process is stopped if one has reached an appropriate solution, which can be identified by $\|\nabla f(x^n)\|_2 \leq \epsilon$. If the initial point x_{start} is already sufficiently close to the true solution then the whole algorithm converges quadratically, i.e. the precision gets doubled at each step.

At this point, let us give the gradient and Hessian of the appearing functions. For the barrier term $\psi(x) = -\log[\det \rho(x)]$ restricted to the positive domain $\rho(x) = \mathbb{1} / \dim(\mathcal{H}) + \sum_i x_i B_i > 0$, one obtains [36]

$$\frac{\partial \psi(x)}{\partial x_i} = -\text{tr}[\rho(x)^{-1} B_i], \tag{43}$$

4. Quantum state tomography

$$\frac{\partial^2 \psi(x)}{\partial x_j \partial x_i} = \text{tr}[\rho(x)^{-1} B_j \rho(x)^{-1} B_i]. \quad (44)$$

Equation (44) shows that the Hessian of the penalty term $\nabla^2 \psi(x) > 0$ is positive definite, such that $\psi(x)$ is indeed convex. The derivatives of the preferred fit function can be computed directly. For instance, using the likelihood function $F_{\text{ml}}(x) = -\sum_k f_k \log p_k(x)$ with $p_k(x) = \text{tr}[\rho(x) M_k]$, they read

$$\frac{\partial F_{\text{ml}}(x)}{\partial x_i} = -\sum_k \frac{f_k}{p_k(x)} \text{tr}(B_i M_k), \quad (45)$$

$$\frac{\partial^2 F_{\text{ml}}(x)}{\partial x_j \partial x_i} = \sum_k \frac{f_k}{p_k^2(x)} \text{tr}(B_j M_k) \text{tr}(B_i M_k). \quad (46)$$

The bottleneck of such an algorithm is the actual computation of the second derivatives. Although the expansion coefficients of each measurement $\text{tr}(B_j M_k)$ can be computed in advance, it is still necessary to compute equation (46) anew at each point x due to the dependence of $p_k(x)$. With respect to that, the least squares fit function bears a great advantage since its Hessian is constant, i.e. $\partial_j \partial_i F_{\text{ls}}(x) = 2 \sum_k w_k \text{tr}(B_j M_k) \text{tr}(B_i M_k)$, such that one saves time on this part.

Finally, let us comment on the optimality conditions, known as the Karush–Kuhn–Tucker conditions [36]. A given x^* is the global solution of the convex problem given by equation (13) if and only if¹⁴ there exists an additional Lagrange multiplier Z^* (as given in the equations beneath) such that the pair (x^*, Λ^*) satisfies

$$\frac{\partial}{\partial x_i} F(x^*) - \text{tr}[\Lambda^* B_i] = 0, \quad \forall i, \quad (47)$$

$$\Lambda^* \geq 0, \quad \rho(x^*) \geq 0, \quad (48)$$

$$\text{tr}[\Lambda^* \rho(x^*)] = 0. \quad (49)$$

Given the solution x_{sol}^t of the corresponding unconstrained problem with penalty parameter t , it follows from $\nabla f(x_{\text{sol}}^t) = 0$ using equation (43) that the gradient conditions are satisfied with $\Lambda_t = t \rho(x_{\text{sol}}^t)^{-1}$ being the Lagrange multiplier. This pair $(x_{\text{sol}}^t, \Lambda_t)$ also satisfies equation (48); only the duality gap condition $\text{tr}[\Lambda_t \rho(x_{\text{sol}}^t)] = t \dim(\mathcal{H}) > 0$ does not hold exactly. However, this quantity appears in the following inequality:

$$F(x_{\text{sol}}^t) - \text{tr}[\Lambda_t \rho(x_{\text{sol}}^t)] = \min_{x: \rho(x) \geq 0} F(x) - \text{tr}[\Lambda_t \rho(x)] \quad (50)$$

$$\leq \min_{x: \rho(x) \geq 0} F(x) = F(x_{\text{sol}}). \quad (51)$$

Here one used that x_{sol}^t is the solution of $F(x) - \text{tr}[\Lambda_t \rho(x)]$ because the gradient vanishes (and the solution is not at the border), and $\text{tr}[\Lambda_t \rho(x)] \geq 0$ for the inequality. This is the stated accuracy given by equation (15) which relates the function value of x_{sol}^t to the true solution x_{sol} .

¹⁴ Sufficiency holds under the Slater regularity condition that demands a strictly feasible point $\rho(x) > 0$, which naturally holds for state reconstruction problems.

5.3. Additional tools

5.3.1. *Optimization of measurement settings.* Measurement settings, each described by a unit vector $\hat{a}_i \in \mathbb{R}^3$ as explained in section 2.1, are chosen to optimize a figure of merit characterizing how well a given permutationally invariant target state ρ_{tar} can be reconstructed. As such a quality measure we use the sum of errors for the tomographically complete operator set of all tensor products of Pauli operators. Note that a permutationally invariant state ρ_{PI} is already uniquely determined by its generalized Bloch vector [17] defined as

$$b_{klmn} = \text{tr} \left(\left[\sigma_x^{\otimes k} \otimes \sigma_y^{\otimes l} \otimes \sigma_z^{\otimes m} \otimes \mathbb{1}^{\otimes n} \right]_{\text{PI}} \rho_{\text{PI}} \right) \quad (52)$$

with natural numbers satisfying $k + l + m + n = N$. Consequently, the total error of all Bloch vector elements is given by

$$\mathcal{E}_{\text{total}}^2(\hat{a}_i, \rho_{\text{tar}}) = \sum_{k,l,m,n} \binom{N}{k,l,m,n} \mathcal{E}_{b_{klmn}}^2(\hat{a}_i, \rho_{\text{tar}}), \quad (53)$$

where the multinomial coefficient weights the error of each generalized Bloch vector by its number of appearance in a generic Pauli product decomposition.

The error of each Bloch vector element must now be related to the carried out measurements. For that, note that each Bloch vector element can be expressed as

$$b_{klmn} = \sum_i c_i^{klmn} \text{tr} \left([(\hat{a}_i \cdot \vec{\sigma})^{\otimes N-n} \otimes \mathbb{1}^{\otimes n}]_{\text{PI}} \rho_{\text{tar}} \right) \quad (54)$$

using appropriate coefficients c_i^{klmn} and the expectation values of $[(\hat{a}_i \cdot \vec{\sigma})^{\otimes N-n} \otimes \mathbb{1}^{\otimes n}]_{\text{PI}}$ which can be computed from the coarse-grained measurement outcomes M_k^i of setting \hat{a}_i as given by equation (2) using linear combinations. Assuming independent errors, one obtains

$$\mathcal{E}_{b_{klmn}}^2(\hat{a}_i, \rho_{\text{tar}}) = \sum_i c_i^{klmn} \mathcal{E}_{\rho_{\text{tar}}}^2 \left([(\hat{a}_i \cdot \vec{\sigma})^{\otimes N-n} \otimes \mathbb{1}^{\otimes n}]_{\text{PI}} \right). \quad (55)$$

The detailed form of the error expression $\mathcal{E}_{\rho_{\text{tar}}}^2 \left([(\hat{a}_i \cdot \vec{\sigma})^{\otimes N-n} \otimes \mathbb{1}^{\otimes n}]_{\text{PI}} \right)$ may depend on the actual physical realization, but we assume the following form:

$$\mathcal{E}_{\rho_{\text{tar}}}^2 \left([(\hat{a}_i \cdot \vec{\sigma})^{\otimes N-n} \otimes \mathbb{1}^{\otimes n}]_{\text{PI}} \right) = K \Delta_{\rho_{\text{tar}}} \left([(\hat{a}_i \cdot \vec{\sigma})^{\otimes N-n} \otimes \mathbb{1}^{\otimes n}]_{\text{PI}} \right), \quad (56)$$

where $\Delta_{\rho}[A] = \text{tr}(\rho A^2) - [\text{tr}(\rho A)]^2$ is the standard variance and K is a state-independent factor. This form fits well, for example, to the common error model in photonic experiments where count rates are assumed to follow a Poissonian distribution. For more details of this derivation, see [17].

For large qubit numbers, N , this optimization is carried out iteratively. Starting from randomly chosen measurement directions or from vectors which are uniformly distributed according to some frame potential [46], one searches for updates according to

$$\hat{a}'_i = \frac{p \hat{a}_i^n + (1-p) \hat{r}_i}{\|p \hat{a}_i^n + (1-p) \hat{r}_i\|}. \quad (57)$$

Here \hat{a}_i^n is the current iterate, $p < 1$ a probability close to 1 and \hat{r}_i are randomly chosen unit vectors. If this new set of directions \hat{a}'_i leads to a smaller total error $\mathcal{E}_{\text{total}}^2(\hat{a}'_i, \rho_{\text{tar}})$ than the previous set, then these new measurement settings form the next iterate $\hat{a}_i^{n+1} = \hat{a}'_i$; otherwise this procedure is carried out once more. This process is repeated until the total error does not decrease anymore. This way of optimizing the measurements requires a method for computing

4. Quantum state tomography

20

IOP Institute of Physics Φ DEUTSCHE PHYSIKALISCHE GESELLSCHAFT

the total error $\mathcal{E}_{\text{total}}^2(\hat{a}_i, \rho_{\text{tar}})$ for a given set of measurements \hat{a}_i . Using the generalized Bloch vector or the spin ensemble this computation can be carried out again efficiently for larger qubit numbers N .

Although this algorithm is not proven to attain the true global optimum it is still a straightforward technique to obtain good settings. In the end this is often sufficient, recalling that the true unknown state can deviate from the target state and that one considers ‘just’ an overall single figure of merit. For our simulations we always use the optimized settings for the totally mixed state; a reasonable guess if no extra information is available [47].

Regarding this point, we finally like to add that if one does not employ the minimal number of measurement settings, but rather an over-complete set, e.g. four times as many settings but four times fewer measurements per setting, then the procedure is quite insensitive to the chosen measurement directions. Hence, in many practical situations the search for optimal directions might not even be necessary and randomly chosen measurement directions suffice equally well.

5.3.2. Statistical pretest. Via the pretest one estimates the fidelity between the true ρ_{true} and the best permutationally invariant state $F_{\text{PI}}(\rho_{\text{true}}) = \max_{\rho_{\text{PI}} \geq 0} \text{tr}(\sqrt{\sqrt{\rho_{\text{true}}}\rho_{\text{PI}}\sqrt{\rho_{\text{true}}}})$ using only measurement results from a few settings $\hat{a} \in T$, e.g. employing only $T = \{\hat{e}_1, \hat{e}_2, \hat{e}_3\}$. Depending on this quantity, one decides on whether permutationally invariant tomography is worth continuing. As explained in detail in [17], this fidelity can be bounded by

$$F_{\text{PI}}(\rho_{\text{true}}) \geq [\text{tr}(\rho_{\text{true}}Z)]^2 \quad (58)$$

with an operator $Z = \sum z_k^a M_k^a$ being built up by the carried out measurements M_k^a given by equation (2) and satisfying $Z \leq P_{\text{sym}}$, where P_{sym} denotes the projector onto the symmetric subspace.

The expansion coefficients z_k^a should be optimized to attain the best lower bound. For a given target state ρ_{tar} this problem can be cast into a semidefinite program [36, 48] that can be solved efficiently using standard numerical routines. However, for larger qubit numbers one must again employ the block structure of the measurement operators as given by equation (9) to handle the operator inequality. Note that the projector on the symmetric subspace has a Block structure $P_{\text{sym},j} = \delta_{j,N/2}\mathbb{1}$. Then the final problem reads as

$$\begin{aligned} \max_z \quad & \sum_{a \in T, k} z_k^a \text{tr}(\rho_{\text{tar}} M_k^a) \\ \text{s.t.} \quad & \sum_{a \in T, k} z_k^a M_{k,j=N/2}^a \leq \mathbb{1}, \quad \sum_{a \in T, k} z_k^a M_{k,j}^a \leq 0, \quad \forall j < N/2. \end{aligned} \quad (59)$$

If one experimentally implements this pretest, one must account for additional statistical fluctuations. For the chosen Z , one can pretest the sample mean $\bar{Z} = \sum z_k^a f_k^a$ using the observed frequencies $f_k^a = n_k^a / N_{\text{R}}$ in N_{R} repetitions of setting \hat{a} , as an estimate of the true expectation value $\text{tr}(\rho_{\text{true}}Z)$. This sample mean \bar{Z} will fluctuate around the true mean but large deviations will become less likely, such that for an appropriately chosen ϵ the quantity $\text{sign}(\bar{Z} - \epsilon)(\bar{Z} - \epsilon)^2$ is a lower bound to the true fidelity at the desired confidence level. The proof essentially uses the techniques employed in [49, 50].

Proposition 5.3 (Statistical pretest). *For any $Z = \sum z_k^a M_k^a \leq P_{\text{sym}}$ let $\bar{Z} = \sum z_k^a n_k^a / N_{\text{R}}$ denote the sample mean using N_{R} repetitions for setting $\hat{a} \in T$. If the data are generated by*

the state ρ_{true} , then

$$\text{Prob} [F_{\text{PI}}(\rho_{\text{true}}) \geq \text{sign}(\bar{Z} - \epsilon)(\bar{Z} - \epsilon)^2] \geq 1 - \exp(-2N_{\text{R}}\epsilon^2/C_z^2) \quad (60)$$

with $C_z^2 = \sum_a (z_{\text{max}}^s - z_{\text{min}}^s)^2$ where $z_{\text{max/min}}^a$ are the respective optima for setting \hat{a} over all outcomes k .

Proof. The given statement follows along

$$\text{Prob} [F_{\text{PI}}(\rho_{\text{true}}) \geq \text{sign}(\bar{Z} - \epsilon)(\bar{Z} - \epsilon)^2] \geq \text{Prob} [\text{tr}(\rho_{\text{true}}Z) \geq \bar{Z} - \epsilon], \quad (61)$$

$$\geq 1 - \text{Prob} [\text{tr}(\rho_{\text{true}}Z) \leq \bar{Z} - \epsilon] \geq 1 - \exp(-2N_{\text{R}}\epsilon^2/C_z^2). \quad (62)$$

Here the first inequality holds because the set of outcomes satisfying $\{n_k^a : \text{tr}(\rho_{\text{true}}Z) \geq \bar{Z} - \epsilon\}$ is a subset of $\{n_k^a : F_{\text{PI}}(\rho_{\text{true}}) \geq \text{sign}(\bar{Z} - \epsilon)(\bar{Z} - \epsilon)^2\}$ using equation (58). In the last inequality we use the Hoeffdings tail inequality [51] to bound $\text{Prob}[\text{tr}(\rho_{\text{true}}Z) \leq \bar{Z} - \epsilon]$. \square

Note that this pretest can also be applied after the whole tomography scheme in which case the projector $P_{\text{sym}} = \sum z_k^a M_k^a$ becomes accessible. Moreover, let us point out that a strong statistical significance, or a low ϵ respectively, might not be achieved with the best expectation value as given by equation (59) [52]; hence, optimizing Z for a rather mixed state is often better.

Finally, let us remark that the pretest can be improved by additional projectors, see the supplementary material of [17]. This leads to the bound $F_{\text{PI}}(\rho_{\text{true}}) \geq \sum_j p_j^2$ with p_j being the weight of the corresponding spin- j state of the permutationally invariant part of ρ_{true} as given in equation (7). From this expression one sees that this test only works well for states having a rather large weight on one of these spin states. Others, like the totally mixed state, while clearly being permutationally invariant, are not identified as states close to the permutationally invariant subspace. This is in stark contrast to compressed sensing where the certificate succeeds for the whole class of low-rank target states [10].

5.3.3. Entanglement and the MaxLik–MaxEnt principle. Following the last comment from the previous subsection, we want to argue that even in the case of an inefficient certificate, permutationally invariant state reconstruction as given by equation (5) is meaningful. Firstly we would like to emphasize that the permutationally invariant part of any density operator corresponds to a good representative to investigate the entanglement properties of the true, unknown state. This is because the transformation given by equation (1) can be achieved by permutations or multiple swap operations together with classical mixing, whereby entanglement cannot be created [53]. Thus if the permutationally invariant part of the density operator is entangled this holds true also for the real state. In addition, any symmetric entanglement measure, i.e. all commonly known ones, can only decrease under this operation, thereby even quantification is faithful [54].

Secondly, permutationally invariant state reconstruction also represents the solution of the maximum-likelihood maximum-entropy principle as introduced in [55], which goes as follows: if the carried out measurements are not tomographically complete, then there is, in general, not a single state $\hat{\rho}_{\text{ml}}$ that maximizes the likelihood but rather a complete set of them. In order to single out a ‘good’ representative, the authors of [55] propose to choose the state that has the largest entropy, which, according to the Jaynes principle [47], is the special state for which one has the fewest information.

4. Quantum state tomography

Proposition 5.4 (Permutationally invariant MaxLik–MaxEnt principle). *Using the described permutationally invariant tomography scheme, the reconstructed permutationally invariant state given by equation (5) (with the likelihood function) is also the solution of the maximum-likelihood maximum-entropy principle.*

Proof. Since the measurements given by equation (2) are invariant and tomographically complete for permutationally invariant states, all density operators with the same spin ensemble as $\hat{\rho}_{\text{PI}}$ have the same maximum likelihood. According to Ref. [47], the state with maximal entropy and consistent with a given set of expectation values for operators M_k^a has the form $\rho \propto \exp(\sum_{a,k} \lambda_k^a M_k^a)$. The Lagrange multipliers $\lambda_k^a \in \mathbb{R}$ must be chosen such that the given expectation values match. However, because all M_k^a are permutationally invariant we can employ the block decomposition given by equation (9) and finally obtain $\exp(\sum_{a,k} \lambda_k^a M_k^a) = \exp(\oplus_j \sum_{a,k} \lambda_k^a M_{k,j}^a \otimes \mathbb{1}) = \oplus_j \exp(\sum_k \lambda_k^a M_{k,j}^a) \otimes \mathbb{1}$. Hence we obtain the same structure as $\hat{\rho}_{\text{PI}}$, which therefore is also the state with maximum entropy. \square

6. Conclusion and outline

In this paper, we have provided all the necessary ingredients to carry out permutationally invariant tomography [17] in experiments with large qubit numbers. This includes, besides scheme-specific tasks such as the statistical pretest and the optimization of the measurement settings, in particular the state reconstruction part. Accounting for statistical fluctuations due to a finite number of data, this reconstruction demands the solution of a nonlinear large-scale optimization problem. We achieve this, firstly, by using a convenient toolbox to store, characterize and process permutationally invariant states, which largely reduces the dimension of the underlying problem and, secondly, by using convex optimization, which is superior compared to commonly used numerical routines in many respects. This makes permutationally invariant tomography a complete tomography method requiring only moderate measurement and data analysis effort.

There are many questions one may pursue in this direction: firstly, let us stress that the current prototype implementation is still not optimal. As explained, the bottleneck is the computation of the second derivatives; hence, we strongly believe that Hessian free optimization, like quasi-Newton or conjugate gradients [56], or the use of other barrier functions more tailored to linear matrix inequalities [57] are likely to push the reconstruction limit further. Secondly, it is natural to try to exploit other symmetries in the development of ‘symmetry’ tomography protocols; that is, tomography should work for all states that remain invariant under the action of a specific group. Clearly, any symmetry decreases the number of state-dependent parameters, but the challenge is to devise efficient local measurement strategies. Interesting classes here are graph-diagonal [58] or, more general, locally maximally entangleable states [59], translation or shift-invariant states [60], or $U^{\otimes N}$ invariant states [61, 62]. Thirdly, it is worth investigating to what extent particularly designed state tomography protocols are useful for further tasks, such as process tomography for quantum gates. For instance, permutationally invariant tomography might be unable to resolve the Toffoli gate [63] directly, but since the operation on all N target qubits is symmetric, a permutationally invariant resolution of this subspace (and the additional control qubit) might be sufficient. Finally, let us point out that permutationally invariant tomography can be further restricted to the symmetric

subspace, which often contains the desired states. This is reasonable since we have seen that the pretest is only good if the unknown state has a large weight in one of the spin states. However, since the symmetric subspace grows only linearly with the number of particles, this tomography scheme can analyse many more qubits efficiently.

Acknowledgments

We thank M Kleinmann, B Kraus, T Monz, L Novo, P Schindler and J Wehr for stimulating discussions on the topic and technicalities. This work was supported by the FWF (START prize Y376-N16), the EU (Project Q-ESSENCE and Marie Curie CIG 293993/ENFOQI), the BMBF (Chist-Era Project QUASAR), the ERC (Starting Grant GEDENTQOPT), the QCCC of the Elite Network of Bavaria, the Spanish MICINN (project no. FIS2009-12773-C02-02), the Basque Government (project no. IT4720-10) and the National Research Fund of Hungary OTKA (contract no. K83858).

References

- [1] Paris M G A and Řeháček J (ed) 2004 *Quantum State Estimation* (Berlin: Springer)
- [2] Häffner H *et al* 2005 *Nature* **438** 643
- [3] Kiesel N, Schmid C, Tóth G, Solano E and Weinfurter H 2007 *Phys. Rev. Lett.* **98** 063604
- [4] James D F V, Kwiat P G, Munro W J and White A G 2001 *Phys. Rev. A* **64** 052312
- [5] Renes J M, Blume-Kohout R, Scott A J and Caves C M 2004 *J. Math. Phys.* **45** 2171
- [6] Monz T, Schindler P, Barreiro J T, Chwalla M, Nigg D, Coish W A, Harlander M, Haensel W, Hennrich M and Blatt R 2011 *Phys. Rev. Lett.* **106** 130506
- [7] Yao X C, Wang T X, Xu P, Lu H, Pan G S, Bao X H, Peng C Z, Lu C Y, Chen Y A and Pan J W 2012 *Nature Photon.* **6** 225
- [8] Huang Y F, Liu B H, Peng L, Li Y H, Li L, Li C F and Guo G C 2011 *Nature Commun.* **2** 546
- [9] Gross D, Liu Y K, Flammia S T, Becker S and Eisert J 2010 *Phys. Rev. Lett.* **105** 150401
- [10] Gross D 2011 *IEEE Trans. Inf. Theory* **57** 1548
- [11] Flammia S T, Gross D, Liu Y K and Eisert J 2012 Quantum tomography via compressed sensing: error bounds, sample complexity and efficient estimators arXiv:1205.2300
- [12] Cramer M, Plenio M B, Flammia S T, Gross D, Bartlett S D, Somma R, Landon-Cardinal O, Liu Y K and Poulin D 2010 *Nature Commun.* **1** 9
- [13] Baumgratz T, Gross D, Cramer M and Plenio M B 2012 Scalable reconstruction of density matrices arXiv:1207.0358
- [14] Landon-Cardinal O and Poulin D 2012 Practical learning method for multi-scale entangled states arXiv:1204.0792
- [15] Yin J O S and van Enk S J 2011 *Phys. Rev. A* **83** 062110
- [16] Guta M, Kypraios T and Dryden I 2012 Rank-based model selection for multiple ions quantum tomography arXiv:1206.4032
- [17] Tóth G, Wieczorek W, Gross D, Krischek R, Schwemmer C and Weinfurter H 2010 *Phys. Rev. Lett.* **105** 250403
- [18] Moroder T, Kleinmann M, Schindler P, Monz T, Gühne O and Blatt R 2012 Detection of systematic errors in quantum experiments arXiv:1204.3644
- [19] D'Ariano G M, Maccone L and Painsi M 2003 *J. Opt. B: Quantum Semiclass. Opt.* **5** 77
- [20] Karassiov V P 2005 *J. Russ. Laser Res.* **26** 484
- [21] Adamson R B A, Shalm L K, Mitchell M W and Steinberg A M 2007 *Phys. Rev. Lett.* **98** 043601

4. Quantum state tomography

- [22] Adamson R B A, Turner P S, Mitchell M W and Steinberg A M 2008 *Phys. Rev. A* **78** 033832
- [23] Hradil Z 1997 *Phys. Rev. A* **55** R1561
- [24] Langford N K 2007 Encoding, manipulating and measuring quantum information in optics *PhD Thesis* School of Physical Sciences, University of Queensland
- [25] Blume-Kohout R 2010 *Phys. Rev. Lett.* **105** 200504
- [26] Efron B and Tibshirani R J 1994 *An Introduction to the Bootstrap* (London: Chapman and Hall)
- [27] Christandl M and Renner R 2011 Reliable quantum state tomography arXiv:1108.5329
- [28] Blume-Kohout R 2012 Robust error bars for quantum tomography arXiv:1202.5270
- [29] Mood A F 1974 *Introduction to the Theory of Statistics* (New York: McGraw-Hill)
- [30] Cirac J I, Ekert A K and Macchiavello C 1999 *Phys. Rev. Lett.* **82** 4344
- [31] Demkowicz-Dobrzański R 2005 *Phys. Rev. A* **71** 062321
- [32] Chase B A and Geremia J M 2008 *Phys. Rev. A* **78** 052101
- [33] Baragiola B Q, Chase B A and Geremia J M 2010 *Phys. Rev. A* **81** 032104
- [34] Hradil Z, Řeháček J, Fiurášek J and Ježek M 2004 *Lect. Notes Phys.* **649** 59
- [35] Reimpell M 2008 Quantum information and convex optimization *PhD Thesis* Technische Universität Braunschweig
- [36] Boyd S and Vandenberghe S 2004 *Convex Optimization* (Cambridge: Cambridge University Press)
- [37] Kosut R L, Walmsley I A and Rabitz H 2004 Optimal experiment design for quantum state and process tomography and Hamiltonian parameter estimation arXiv:quant-ph/0411093
- [38] Glancy S, Knill E and Girard M 2012 Gradient-based stopping rules for maximum-likelihood quantum-state tomography arXiv:1205.4043
- [39] Řeháček J, Hradil Z and Ježek M 2001 *Phys. Rev. A* **63** 040303
- [40] Řeháček J, Hradil Z, Knill E and Lvovsky A I 2007 *Phys. Rev. A* **75** 042108
- [41] Nielsen M A and Chuang I L 2000 *Quantum Computation and Quantum Information* (Cambridge: Cambridge University Press)
- [42] Życzkowski K and Sommers H J 2001 *J. Phys. A: Math. Gen.* **34** 7111
- [43] Simon B 1996 *Representations of Finite and Compact Groups (Graduate Studies in Mathematics vol 10)* (Providence, RI: American Mathematical Society)
- [44] Christandl M 2006 The structure of bipartite quantum states—insights from group theory and cryptography arXiv:0604183
- [45] Hall B C 2003 *Lie Groups, Lie Algebras and Representations: An Elementary Introduction* (New York: Springer)
- [46] Gross D, Audenaert K and Eisert J 2007 *J. Math. Phys.* **48** 052104
- [47] Jaynes E 1957 *Phys. Rev.* **108** 171
- [48] Tóth G, Wieczorek W, Krischek R, Kiesel N, Michelberger P and Weinfurter H 2009 *New J. Phys.* **11** 083002
- [49] Flammia S T and Liu Y K 2011 *Phys. Rev. Lett.* **106** 230501
- [50] da Silva M P, Landon-Cardinal O and Poulin D 2011 *Phys. Rev. Lett.* **107** 210404
- [51] Hoeffding W 1963 *J. Am. Stat. Assoc.* **58** 301
- [52] Jungnitsch B, Niekamp S, Kleinmann M, Gühne O, Lu H, Gao W B, Chen Y A, Chen Z B and Pan J W 2010 *Phys. Rev. Lett.* **104** 210401
- [53] Busch P 2003 *Int. J. Theor. Phys.* **42** 937
- [54] Horodecki K, Horodecki M and Horodecki P 2010 *Quantum Inform. Comput.* **10** 0901
- [55] Teo Y S, Zhu H B-G, Englert B G, Řeháček J and Hradil Z 2011 Knowledge and ignorance in incomplete quantum state tomography arXiv:1102.2662
- [56] Geiger C and Kanzow C 1999 *Numerische Verfahren zur Lösung unrestringierter Optimierungsaufgaben* (Berlin: Springer)
- [57] Kočvara M and Stingl M 2003 *Optim. Methods Softw.* **18** 317
- [58] Hein M, Dür W, Eisert J, Raussendorf R, Van den Nest M and Briegel H J 2005 *Proc. Int. School of Physics 'Enrico Fermi' on 'Quantum Computers, Algorithms and Chaos', Varenna, Italy*

- [59] Kruszynska C and Kraus B 2009 *Phys. Rev. A* **79** 052304
- [60] O'Connor K M and Wootters W K 2001 *Phys. Rev. A* **63** 052302
- [61] Eggeling T and Werner R F 2001 *Phys. Rev. A* **63** 042111
- [62] Cabello A 2003 *J. Mod. Opt.* **50** 10049
- [63] Monz T, Kim K, Hänsel W, Riebe M, Villar A S, Schindler P, Chwalla M, Hennrich M and Blatt R 2009 *Phys. Rev. Lett.* **102** 040501

4. Quantum state tomography

5. Systematic errors of standard quantum state estimation

Applying the tomography schemes described in section 4.1.1 and 4.1.2 often results in a density matrix which fails to be physical, i.e., some of its eigenvalues are negative. Since the eigenvalues are associated with probabilities, and are thus constrained to lie within the interval $[0, 1]$, this poses a serious conceptual problem. One might argue that as long as a value of 0 lies within the error, there is no reason to bother. Nevertheless, the question arises how to deduce entanglement measures or mixed state fidelities for such states, as they might be mathematically ill-defined or just meaningless.

Typically, in such a situation, maximum likelihood or least squares fitting algorithms as discussed in section 4.1.3 are applied which guarantee to deliver a physical result. Recently, various alternative approaches have been proposed which all restrict the result to lie within the physical regime [281, 282] and even, although numerically expensive, allow to calculate meaningful confidence and credibility regions for the estimate itself [263, 283–287]. Furthermore, quantum state tomography and the closely related quantum process tomography could even be adopted to detect systematic errors in the alignment of an experiment [246, 288–292].

In this chapter, several reasons for seemingly unphysical results in tomography experiments will be discussed. In section 5.1, it will be shown that the origin of such unphysical results can be manifold, ranging from experimental imperfections to the orientation of the state with respect to the measurement axis. Then, in section 5.2, it will be shown that implying additional constraints in quantum state estimation, such as non-negativity of a quantum state, can introduce systematic errors. In estimation theory, such errors are well-known under the term bias [80, 293] (see publication P5.1). The size of these errors for derived quantities like, e.g., the fidelity with respect to a target state is investigated on the example of a four-qubit GHZ state mixed with white noise. Furthermore, it will be proven that an estimator which always delivers a state obeying the physicality constraint cannot be unbiased. In publication P5.1, the way towards an, at least partial, solution of this problem is presented. There, it is shown that any quantity that is given by a convex function, like, e.g., many entanglement measures, can be approximated by a linear function for which meaningful bounds together with simple to calculate errors can be given. However, as these bounds are derived with the aid of Hoeffding’s tail inequality [294], they are relatively weak.

5.1. Origin of unphysical density matrices

It is known that misalignment in an experimental apparatus can cause inconsistencies in the data evaluation process when it remains undetected [288]. Here, by means of two illustrative examples, it will be shown that there are several reasons why the states as obtained from a tomographic analysis together with linear inversion (see section 4.1) can lead to seemingly unphysical results, i.e., states with at least one negative eigenvalue. More precisely, the misalignment of the experimental apparatus is just one possible reason for such unphysical results and, in view of the quality of today's experiments, it is by far not the most relevant one. Below, a few examples are given, where it is the statistical noise which suffices to cause unphysical estimates.

As a first example, let us consider Pauli tomography of a pure single-qubit state $|z^+\rangle$ and measurements in the standard bases σ_x , σ_y , and σ_z . When performing measurements in the σ_z basis, the result is always $+1$ and thus $T_z = 1$, independently of the measurement statistics! Consequently, the remaining correlations T_x

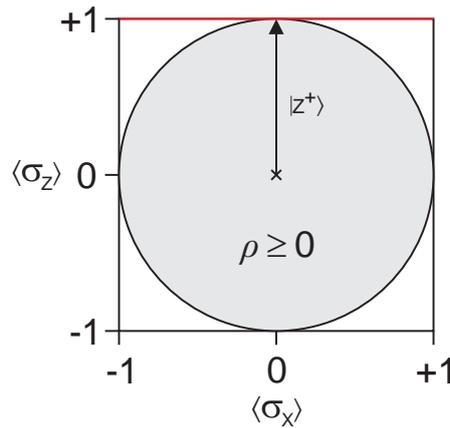


Figure 5.1.: Experimental Pauli tomography of the single-qubit state $|z^+\rangle$. For simplicity, only the x - z plane is shown. The square represents all results that can, at least in principle, be observed in an experiment. All points inside the gray circle represent valid quantum states ρ , i.e., $\text{Tr}(\rho) = 1$ and $\rho \geq 0$, with the pure states lying on the rim of the circle. The states represented by the white corners of the square correspond to states where the physicality condition $\rho \geq 0$ is violated. Measuring in the σ_z basis delivers a correlation value of $T_z = 1$ independently of the number of repetitions of the experiment. Then, when measuring in the σ_x basis afterwards, all resultant states except the one for $T_x = 0$ lie outside the physical regime as indicated by the red line in the figure. However, when the number of measurements in the σ_x basis is odd, $T_x \neq 0$. Please note that this peculiar behavior cannot be overcome by simply improving the measurement statistics!

5.1. Origin of unphysical density matrices

and T_y both have to be 0 because otherwise the state would lie outside the Bloch sphere. This can easily be seen by expressing the purity of a single-qubit state in terms of its Bloch vector elements, $\text{Tr}(\rho^2) = \frac{1}{2}(1 + T_x^2 + T_y^2 + T_z^2)$ and the fact that the purity is bounded from above by 1. Now, when measuring in the σ_x basis, the correlation T_x can only vanish if the results $+1$ and -1 are observed with the same frequency. However, this is not possible if the number of measurements in the σ_x basis is odd! For more details, see Fig. 5.1.

The second example to be considered is Pauli tomography of the one parameter family of four-qubit states $p \frac{\mathbb{1}^{\otimes 4}}{16} + (1-p)|GHZ_4\rangle\langle GHZ_4|$ with $p \in [0, 1]$. The two extremal cases $p = 0$ and $p = 1$ correspond to a GHZ state and to white noise, respectively. Since all the eigenvalues of white noise are $\frac{1}{2^N}$, i.e., it is a full rank state, at first glance, one would expect, that it is fairly unlikely to observe negative eigenvalues in a tomographic experiment. This view is further supported by the fact that the overlap between white noise and *any* pure state $|\phi\rangle$ is $\langle \phi | \frac{\mathbb{1}^{\otimes N}}{2^N} | \phi \rangle = \frac{1}{2^N}$ which means that, in some sense, white noise can be considered to lie in the middle of the state space. However, as can be seen in Fig. 5.2, the opposite turns out to be the case, namely, that on the order of 1000 counts per basis setting (a measurement statistics hardly found in multiqubit experiments) are necessary until one can expect to observe no negative eigenvalues. This counterintuitive behavior can be explained by the fact that determining the eigenvalues of a state comes with some statistical error and that the observed eigenvalues scatter around their true value. Obviously, the smaller the sample size, the larger the scattering is expected to be. Hence, for small sample sizes, an eigenvalue < 0 lies within the scatter of the respective eigenvalue. Thus, it is likely that at least one of the 2^N eigenvalues is negative and it is therefore normal to observe negative eigenvalues even for white noise.

As explained in section 4.1.3, when determining ρ by means of linear inversion, one interprets the experimental data from a frequentistic point of view. The observed relative frequencies are treated as an approximation to the unknown underlying probabilities. In the asymptotic limit, i.e., for infinitely many repetitions of the experiment, the relative frequencies converge towards the probabilities. Therefore, one might expect that the probability to observe a state with at least one negative eigenvalue decreases for better and better statistics. In principle, this view is correct, as can be seen in Fig. 5.2, for values of $p > 0$. However, for pure states (i.e. $p = 0$) increasing the sample size does not reduce the number of states with at least one negative eigenvalue, as can be seen in Fig. 5.2. This peculiar behavior comes from the fact that for pure N -qubit states $2^N - 1$ eigenvalues are exactly 0 and thus every infinitesimal perturbation, as induced for example by statistical noise, will generally lead to at least one negative eigenvalue.

To show that the above described characteristics of the Pauli tomography scheme together with linear inversion are no artifact of the considered family of states, the simulation was repeated for a second family of states, $p \frac{\mathbb{1}^{\otimes 4}}{16} + (1 -$

5. Systematic errors of standard quantum state estimation

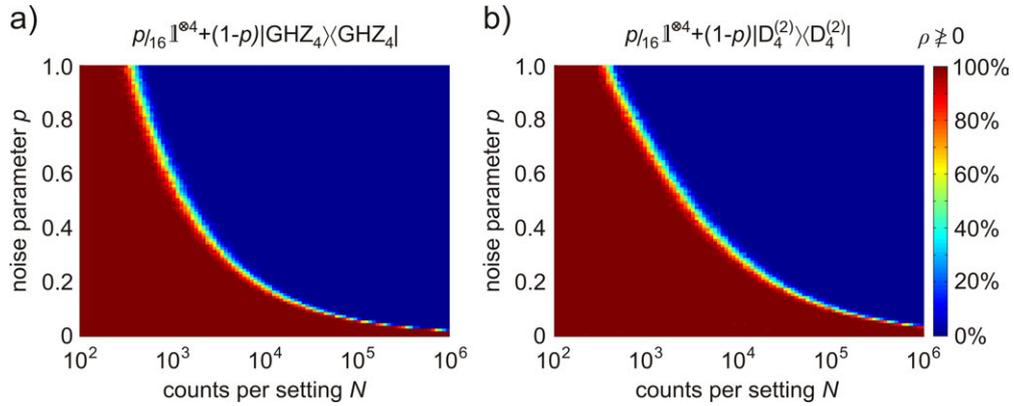


Figure 5.2.: Simulation of Pauli tomography for two one parameter families a) $p\frac{\mathbb{1}^{\otimes 4}}{16} + (1-p)|GHZ_4\rangle\langle GHZ_4|$ and b) $p\frac{\mathbb{1}^{\otimes 4}}{16} + (1-p)|D_4^{(2)}\rangle\langle D_4^{(2)}|$ for different measurement statistics under the assumption of a multinomial distribution of the counts. As can be seen in both cases, the percentage of states violating the physicality condition $\rho \geq 0$, is significant for small sample sizes and high purities.

$p|D_4^{(2)}\rangle\langle D_4^{(2)}|$. Moreover, it has to be excluded that the orientation of the state with respect to the measurement axes plays a role. Therefore, the Dicke state $D_4^{(2)}$ was not expressed in the computational basis ($|z^+\rangle$ and $|z^-\rangle$) but rotated such that $|z^+\rangle \rightarrow |x^+\rangle + |y^+\rangle + |z^+\rangle$ and $|z^-\rangle \rightarrow |x^-\rangle + |y^-\rangle + |z^-\rangle$. The result of the simulation can be seen in Fig. 5.2b). As the similarity between Fig. 5.2a) and Fig. 5.2b) is remarkable, it is safe to say that the observed characteristics can really be attributed to the Pauli tomography scheme together with linear inversion.

Furthermore, in Fig. 5.2, it can be seen that there is a sharp border between the regime where the reconstructed state is almost certainly positive-semidefinite and where it normally fails to be positive-semidefinite. More precisely, this means that when the number of counts surpasses a certain limit, the reconstructed state is almost for sure physical. By means of random matrix theory, it can be shown that the eigenvalues of the reconstructed density matrix follow a Wigner semicircle distribution. Additionally, random matrix theory even allows to determine the number of counts where the transition between the physical and the unphysical regime occurs. A detailed discussion of these findings will be given in a future publication by Lukas Knips.

5.2. Constrained state estimation

5.2.1. Flaws of constrained state estimation

In order to ensure physicality of an experimentally determined density matrix, the fitting algorithms described in section 4.1.3 are commonly applied. However, at this point, one might ask if this approach really solves the problem of negative eigenvalues or whether it suffers from other shortcomings? To answer this question, a simple numerical experiment shall be performed. In detail, the Pauli tomography scheme is simulated on the example of a test state together with different algorithms for data post processing, namely, maximum likelihood, least squares, and, for comparison, linear inversion. As a test state, let us consider a four-qubit GHZ state that is mixed with white noise

$$\varrho_{\text{GHZ}}(p) = (1 - p) |\text{GHZ}_4\rangle\langle\text{GHZ}_4| + p \frac{\mathbb{1}^{\otimes 4}}{16} \quad (5.1)$$

where the parameter p is chosen such that the fidelity with the ideal GHZ state is 0.8. The reason to consider a mixed state is to avoid possible boundary effects from the border of the state space. The numerical experiment consists of 4 steps:

- i)* Calculate the expectation values p_i for all measurement operators M_i .
- ii)* Toss coins to generate frequencies f_i according to a multinomial distribution with \mathcal{N} counts for every measurement basis.
- iii)* Reconstruct the state with either procedure for data post processing mentioned above and, from the result, calculate any quantity of interest.
- iv)* Repeat steps *i*)-*iv*) 500 times

In Fig. 5.3, the result of this numerical experiment for $\mathcal{N} = 100$ with the fidelity as quantity of interest is shown. Ideally, on average, one would expect that all three methods return a value of 0.8 for the fidelity. However, as can be clearly seen, both maximum likelihood and, even worse, least squares fitting lead to a systematic underestimation of the fidelity. In detail, maximum likelihood delivers a value of 0.788 ± 0.010 for the fidelity and least squares fitting delivers 0.749 ± 0.010 . Even if one takes the widths of the distributions into consideration, the values are off by one and five standard deviations, respectively. In contrast, the fidelity value of 0.799 ± 0.012 delivered by linear inversion is in agreement with the true value of 0.8. From an engineering point of view, one might argue that underestimating the fidelity is more an inconvenience than an actual problem. However, much worse, other quantities that are used to detect entanglement, like the negativity, are overestimated, see Fig. 5.4 which, as a consequence, can lead to false entanglement detection! In general, this systematic offset depends

5. Systematic errors of standard quantum state estimation

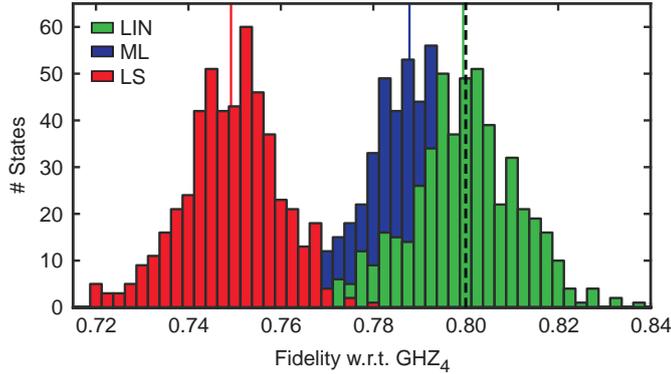


Figure 5.3.: Histogram of the fidelity as obtained from 500 independent simulations of the Pauli tomography scheme together with three different schemes for data post processing. The average values delivered by maximum likelihood (ML) and least squares (LS) clearly fluctuate around a value well below the true fidelity of 0.8. This systematic error called bias is not observed for linear inversion (LIN).

on various parameters like, e.g., the system size, the quantity of interest or the state itself (for details see publication P5.1). Therefore, a later manual correction of the offset seems unjustifiable!

In mathematical statistics and in the theory of point estimates, such systematic offsets are well-known under the term bias [295]. Adopted to quantum state tomography, an estimator $\hat{\varrho}$ is called unbiased if it fulfills the relation

$$\mathbb{E}_{\varrho_0}(\hat{\varrho}) \equiv \sum_f P_{\varrho_0}(f) \hat{\varrho}(f) = \varrho_0. \quad (5.2)$$

Otherwise it is called biased. In order to obtain a more intuitive understanding of the origin of the bias in quantum state estimation consider Fig. 5.5. In subfigure a), possible results of tomography of a pure state before fitting are shown. Due to statistical noise, many of these results lie outside the physical regime. However, the center of mass of the whole distribution coincides with the true (unknown) state. Subfigure b) shows how maximum likelihood or least squares fitting change the distribution. All states that were positive-semidefinite remain unchanged whereas all states that were outside the physical regime are mapped on the border of the state space. Consequently, the center of mass of the whole distribution changes which, as a result, leads to biased state estimation. Please note that linear inversion does not suffer from this problem, i.e., it is unbiased because, according to Eq. 4.12, ϱ_{lin} is a linear function of the frequencies f_i and for all frequencies $\mathbb{E}(f_i) = p_i$ holds. In estimation theory it is common knowledge that applying additional constraints can lead to biased estimation. In section 5.2.2, it is proven that by adding the physicality constraint *any* quantum state estimation scheme necessarily becomes biased.

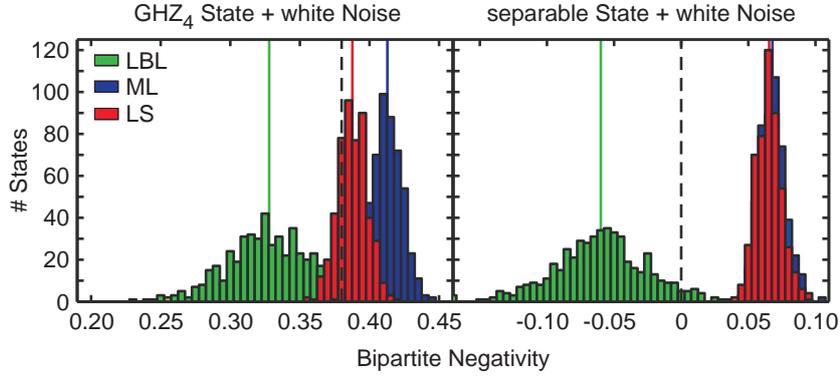


Figure 5.4.: Both maximum likelihood (ML) and least squares (LS) estimation can lead to a systematic overestimation of the negativity which, in consequence, may lead to false entanglement detection when analyzing separable states as shown in b). As discussed in publication P5.1, linearization of the negativity allows to deduce a lower bound (LBL) such that a valid estimate is obtained.

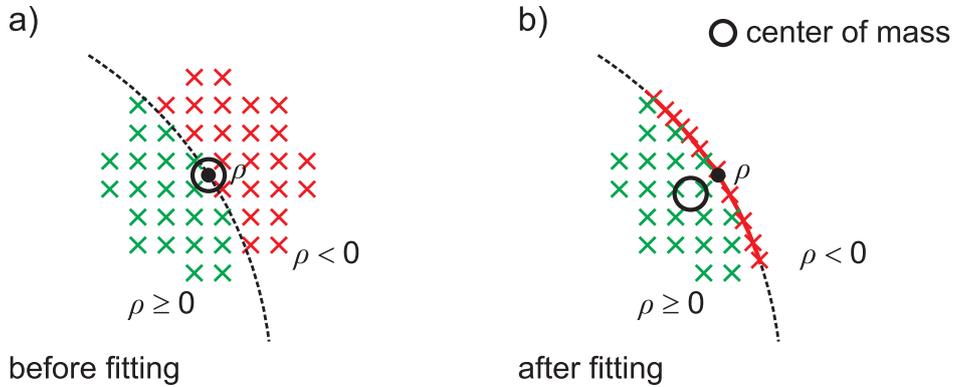


Figure 5.5.: Enforcing physicality in a quantum state estimation scheme leads to biased state estimation. Before fitting, the determined states scatter around the true state ρ with many of them lying outside the physical regime. Fitting algorithms based on maximum likelihood or a least squares minimization map all states with negative eigenvalues on the border of the state space. Effectively, this leads to a shift of the center of mass of the distribution of states which is subsequently observed as a bias.

5.2.2. Biasedness and physicality

As shown in section 5.2.1, for sample sizes typical for multiqubit experiments, both maximum likelihood and least squares fitting algorithms lead to systematic deviations for deduced quantities like, e.g., the fidelity with respect to a target state. Thus, the question arises if such deviations are only a shortcoming of the respective fitting algorithm and one only needs to find a better algorithm or whether the observed effects are of a more general nature. In the following, it will be proven that as soon as the constraint of non-negativity is imposed, systematic deviations are unavoidable. More precisely, the constraint of non-negativity automatically leads to biased state estimation. The proof given below is a more detailed version of the proof presented in publication P5.1.

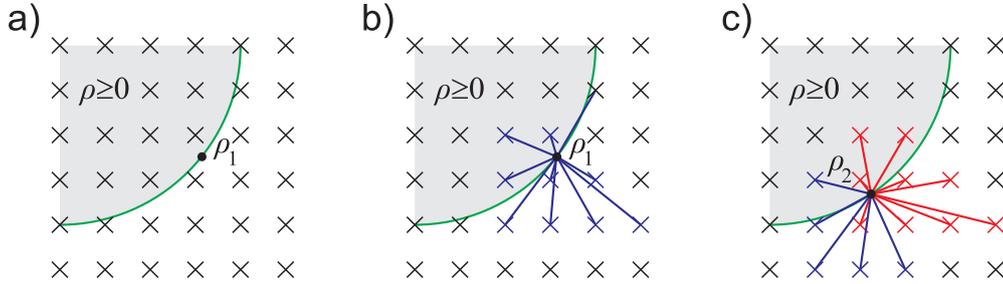


Figure 5.6.: a) A pure quantum state $\varrho_1 = |\psi_1\rangle\langle\psi_1|$ is analyzed by means of quantum state tomography. Because the experimental measurement time is finite, also the set of acquired experimental data has to be finite. Therefore, only a subset of states is accessible, here, marked by crosses. b) Enforcing physicality and unbiasedness of a reconstruction scheme for ϱ_1 , leads to mapping *all* possible results for tomography of ϱ_1 back to ϱ_1 . c) Now a second pure quantum state $\varrho_2 = |\psi_2\rangle\langle\psi_2|$ is to be analyzed which is non-orthogonal to ϱ_1 and, therefore, there are some results which can be obtained for both states. If now the reconstruction scheme shall also be unbiased for ϱ_2 , all possible results would have to be mapped back to ϱ_2 . However, this is not possible because some of the results are already mapped to ϱ_1 , here marked in red. Therefore, a reconstruction scheme which always returns a valid density operator cannot be unbiased.

Theorem. *A reconstruction scheme for quantum state tomography that always yields a valid, i.e., non-negative, density operators is biased.*

Proof. Let us consider experimental quantum state tomography with, due to limited measurement time, a finite sample size. For finite sample sizes, the set of frequencies that can be observed is restricted because all frequencies have to be

multiples of $1/\mathcal{N}$ where \mathcal{N} is a normalization constant that depends on the sample size (for Pauli tomography \mathcal{N} is the number of counts per measurement setting). In Fig. 5.6a), the possible states that can be obtained in a tomography experiment by linear inversion for a certain measurement statistic are marked by crosses. Now consider tomography of a pure quantum state $\varrho_1 = |\psi_1\rangle\langle\psi_1|$ with a set of possible data $\mathcal{S}_1 = \{f_1 | P_{|\psi_1\rangle}(f_1) > 0\}$, where $P_{|\psi_1\rangle}(f_1)$ is the probability to obtain data f_1 . The corresponding states that are obtained from f_1 by linear inversion are marked by the blue crosses in Fig. 5.6b). Let us assume that there exists a reconstruction scheme $\hat{\varrho}$ which returns a valid quantum state $\hat{\varrho}(f) \geq 0$ for all possible outcomes f . Furthermore, the reconstruction scheme shall satisfy Eq. 5.2 for $|\psi_1\rangle$, i.e., $\sum_{\mathcal{S}_1} P_{|\psi_1\rangle}(f_1)\hat{\varrho}(f_1) = |\psi_1\rangle\langle\psi_1|$. This incoherent sum over all $\hat{\varrho}(f_1)$ can only yield the pure state $|\psi_1\rangle\langle\psi_1|$, if and only if, $\hat{\varrho}(f_1) = |\psi_1\rangle\langle\psi_1|$ for all $f_1 \in \mathcal{S}_1$ as shown in Fig. 5.6b). Hence, the outcome of the reconstruction scheme is *identical* for all possible f_1 which is a quite pathological situation already. Now consider a second pure state $\varrho_2 = |\psi_2\rangle\langle\psi_2|$ with $\mathcal{S}_2 = \{f_2 | P_{|\psi_2\rangle}(f_2) > 0\}$ which is non-orthogonal with respect to $|\psi_1\rangle$, i.e., $\langle\psi_1|\psi_2\rangle \neq 0$. For these two states, there exists a non-empty 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 as marked by the red crosses in Fig. 5.6c). Hence, there exist $f_2 \in \mathcal{S}_{12}$ with $\hat{\varrho}(f_2) = |\psi_1\rangle\langle\psi_1| \neq |\psi_2\rangle\langle\psi_2|$. Consequently, the sum in Eq. 5.2 now becomes an incoherent mixture of at least two pure states and thus the equality $\sum_{\mathcal{S}_2} P_{|\psi_2\rangle}(f_2)\hat{\varrho}(f_2) = |\psi_2\rangle\langle\psi_2|$ is violated for $|\psi_2\rangle$. Thus, Eq. 5.2 does not hold for $|\psi_2\rangle$ and, consequently, the reconstruction scheme is biased.

5.3. Publications

5. Systematic errors of standard quantum state estimation

P5.1

PRL 114, 080403 (2015)

PHYSICAL REVIEW LETTERS

week ending
27 FEBRUARY 2015

Systematic Errors in Current Quantum State Tomography Tools

Christian Schwemmer,¹ Lukas Knips,¹ Daniel Richart,¹ Harald Weinfurter,¹ Tobias Moroder,²
Matthias Kleinmann,^{2,3,4} and Otfried Gühne²

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

²Naturwissenschaftlich-Technische Fakultät, Universität Siegen, Walter-Flex-Straße 3, D-57068 Siegen, Germany

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

⁴Department 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 multi-qubit or multiqubit 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].

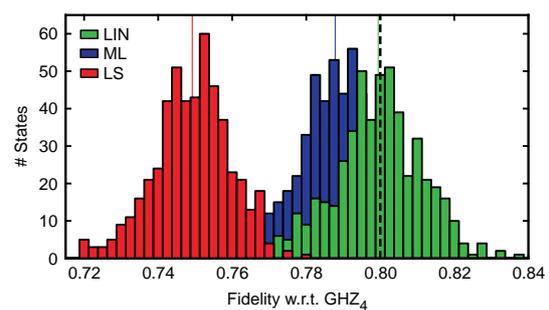


FIG. 1 (color online). Histogram of the fidelity estimates of 500 independent simulations of QST of a noisy four-party GHZ state for three different reconstruction schemes. The values obtained via ML (blue) or LS (red) fluctuate around a value that is lower than the initial fidelity of 80% (dashed line). 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 ρ_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 ρ_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 ρ_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 ρ 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 ρ_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}_{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,

5. Systematic errors of standard quantum state estimation

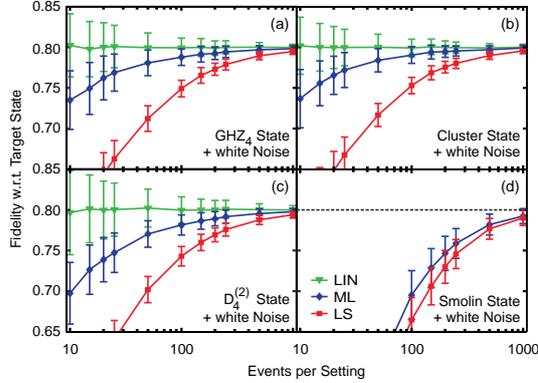


FIG. 2 (color online). The performance of ML, LS, and LIN methods depending on the number of events N_s per setting and for four different noisy initial states ρ_0 . Note that the fidelity can only be calculated linearly if the reference state is pure which is not the case for the Smolin state [25]. 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 ρ_{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

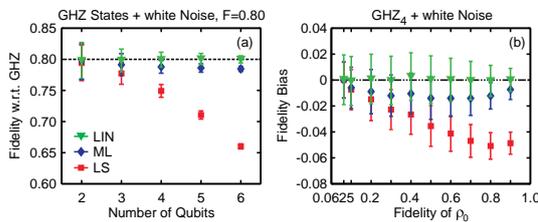


FIG. 3 (color online). The behavior of ML, LS, and LIN depending on the number of qubits n (left) and the fidelity of ρ_0 (right).

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 ρ_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 ρ 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 ρ_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 γ 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 ρ_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 $h^2 = \sum_s (l_{\text{max}}^s - l_{\text{min}}^s)^2$, and $l_{\text{max}/\text{min}}^s$ 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 ρ_{AB} is given by $N(\rho_{AB}) \geq \text{tr}(\rho_{AB} L)$ for any L satisfying $1 \geq L^{T_A} \geq 0$, 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

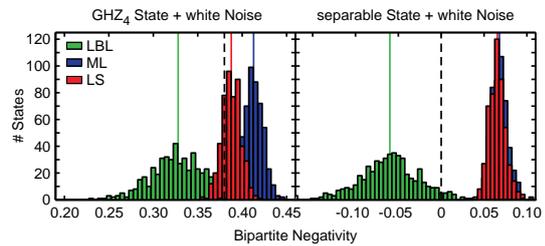


FIG. 4 (color online). Lower bound obtained by linearizing bipartite negativity (LBL) for the GHZ (left) and a four-qubit product state (right) both mixed with white noise resulting in 80% fidelity. The ML and LS reconstruction leads to a systematic overestimation of the negativity, while the lower bound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

5. Systematic errors of standard quantum state estimation

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. Řeháček, (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. Wieczorek, 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. Niggelbaum, 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. Mura, *Phys. Rev. Lett.* **111**, 160406 (2013); R. Blume-Kohout, [arXiv:1202.5270](https://arxiv.org/abs/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. Mura, *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

SUPPLEMENTAL MATERIAL: SYSTEMATIC ERRORS IN CURRENT QUANTUM STATE TOMOGRAPHY TOOLS

Christian Schwemmer,¹ Lukas Knips,¹ Daniel Richart,¹ and Harald Weinfurter,¹ Tobias Moroder,² Matthias Kleinmann,^{2,3,4} and Otfried Gühne,²

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

²*Naturwissenschaftlich-Technische Fakultät, Universität Siegen, Walter-Flex-Straße 3, D-57068 Siegen, Germany*

³*Departamento de Matemática, Universidade Federal de Minas Gerais, Caixa Postal 702, Belo Horizonte, Minas Gerais 31270-901, Brazil*

⁴*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 ρ 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 ν 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_{ν} , 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_{μ} 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_{ν} are replaced by the frequencies f_{ν} 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.

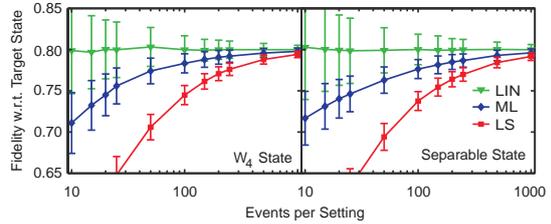


FIG. S1. The behavior of ML, LS and LIN depending on the number of events N_s per setting for different noisy initial states ρ_0 .

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

5. Systematic errors of standard quantum state estimation

2

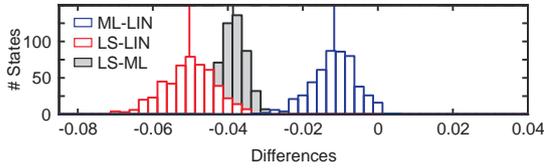


FIG. S2. Here we show the differences of the respective fidelity estimates evaluated for each single simulated tomography experiment as shown in Fig. 1 of the main text. It shows that the respective ML or LS estimate, with one rare exception, is always lower than the LIN estimate. Comparing ML and LS (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 $\hat{P}^s(r) = \text{tr}(\hat{\rho}_0 M_r^s)$ 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_{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 ± 0.011 for ML and 0.700 ± 0.012 for LS) while non-parametric bootstrapping (0.780 ± 0.011 for ML and 0.714 ± 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 $\hat{C}_\varrho(f)$ [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 $Q(\varrho)$ for all $\varrho \in \hat{C}_\varrho(f)$. 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 μ_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.

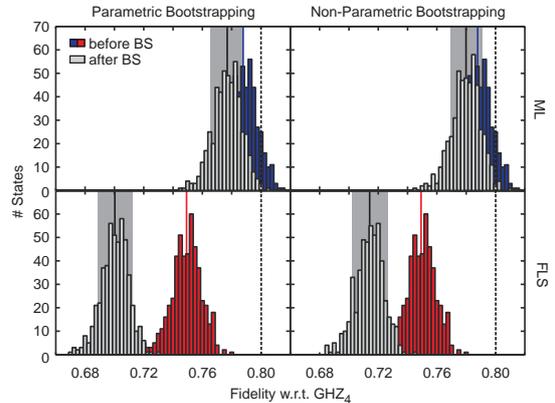


FIG. S3. Error bar computation for the fidelity of the four-qubit GHZ state via Monte-Carlo simulation using either parametric or non-parametric bootstrapping with the data from Fig. 1. For each of these 500 observations f_{obs} , 100 new data sets $f^{(i)}$ were generated and reconstructed in order 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.

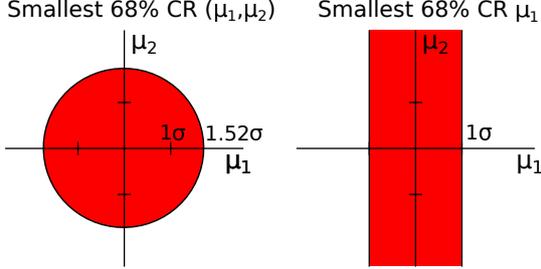


FIG. S4. Which confidence region is the smallest? If one is interested in both mean values $\vec{\mu} = (\mu_1, \mu_2)$ then clearly the left one represents the smallest one, but if $Q(\vec{\mu}) = \mu_1$ is chosen, then the CR obtained by projection onto the μ_1 -axis is much larger for the disc-shaped 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 ϱ_0 . Recall that while the guess ϱ_{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 ϱ to determine the parameter θ in an evolution $U(\theta, H) =$

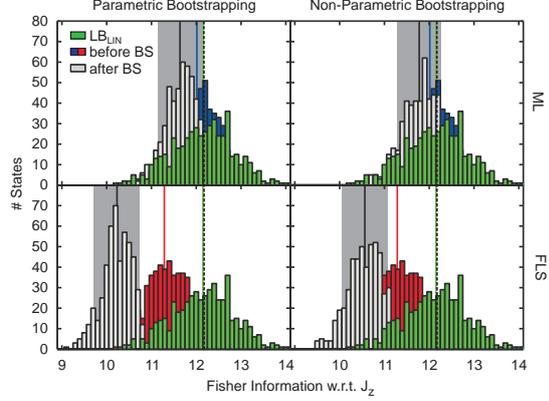


FIG. S5. Full analysis of a Pauli QST scheme with $N_s = 100$ on four qubits in order to deduce the quantum Fisher information with respect to $H = J_z$. As the true underlying state we assume again a noisy four-party GHZ state. We observe that the quantum Fisher information is underestimated from both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

5. Systematic errors of standard quantum state estimation

4

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. Murao, *Phys. Rev. Lett.* **111**, 160406 (2013); R. Blume-Kohout, arXiv:1202.5270; J. Shang, H. K. Ng, A. Schrawat, 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).

6. Conclusions

In the present thesis, highly efficient novel tools for entanglement detection and the tomographic analysis of genuinely multipartite entangled quantum states were presented. These tools were experimentally implemented on the example of various multiqubit states with up to six photons. The focus of this research work was twofold: First, the emphasis was on the development and application of experimentally friendly entanglement indicators which allow for conclusive statements about the presence of entanglement from possibly few measurements. In this context, it was also investigated which types of correlations are actually necessary for genuine N -partite entanglement. Second, new approaches for partial quantum state tomography were studied. For the experimental implementation of these schemes, and in order to compare them against full state tomography, it was necessary to further increase the pump power of the photon source and to improve the efficiency of the whole optical setup.

Many entanglement criteria that are state independent require knowledge of the complete density matrix and are therefore experimentally costly. In contrast, other criteria, like entanglement witnesses, are experimentally far more friendly in terms of the measurement effort but dependent on the state to be investigated. Ideally, an entangled criterion should be, at the same time, state independent and require only few measurements. In this work, two schemes for fast entanglement detection were presented that are both state independent and do not require a full state tomography [52, 53] (see publications P3.1 and P3.2). The first scheme is designed for pure two-qubit states and consists of two steps. In the first step, the local Schmidt bases are determined and the state is (up to a phase) transformed into the Schmidt basis. Then, in the second step, entanglement can easily be verified with a simple criterion [51] which requires at maximum three correlation measurements. The second scheme is based on the principle of correlation complementarity [63] and can also be applied for states with more than two qubits. It can be expressed in a compact form called decision tree which optimally reflects the adaptivity of the scheme, i.e., the result of the previous correlation measurement determines which is the next measurement to be carried out.

The approach to use correlation measurements to detect entanglement was also applied to a particular three-qubit state [54] (see publication P3.3). Generally, entanglement criteria that are used to verify N -partite genuine entanglement depend on N -partite correlations [51, 60, 64–71]. However, interestingly, quantum mechanics allows for states that are genuinely N -partite entangled without

6. Conclusions

N -partite correlations [74]. In this work, a general scheme was presented to construct quantum states with vanishing N -partite correlations for an odd number of qubits. In order to characterize such remarkable states, two experiments aiming at the preparation of symmetric Dicke states with four and six qubits were carried out. The target states could then be obtained by performing appropriate projection measurements [127].

In case one wants to gain more information about a quantum state than just whether it is entangled or not but full tomography is unwanted or not feasible, the application of a partial tomography scheme is the method of choice. In the present work, two novel schemes for partial tomography schemes, compressed sensing and permutationally invariant tomography, were discussed. They were both experimentally implemented to characterize symmetric Dicke states with four and six photons [56, 58] (see publications P4.1 and P4.2). In these experiments, it could be shown that full tomography and permutationally invariant tomography deliver compatible results. For the analysis of six qubits, it could be shown that the experimental effort in comparison to full tomography could be reduced considerably. In detail, the effort could be reduced by a factor of 3 for compressed sensing, by a factor of 25 for permutationally invariant tomography, and by a factor of 50 for a combination of the two methods. In this context, it is also highly important that one has efficient algorithms for data analysis at hand [57] (see P4.3). Therefore, a novel algorithm based on convex optimization was applied.

For every experiment, a careful analysis of all the errors that might influence the final result is required. In this work, it was investigated whether the standard approach for data post processing itself can lead to systematic errors [59] (see publication P5.1). There, it was observed that for the measurement statistics that are typical for today's multiqubit experiments, data analysis itself can lead to a considerably biased estimate. More precisely, it could be shown that this bias can be attributed to restricting the final result to lie within the physical regime.

As a closing remark, a short comment on possible further research based on this work and on the obstacles that have to be surmounted seems appropriate. In general, the preparation of photonic multiqubit states is still a great challenge due to several reasons. First, as deterministic single photon sources are still under development [44–48], one has to resort to the process of spontaneous parametric downconversion. One of the largest drawbacks is that for downconversion sources, there is always a nonvanishing probability to generate more photons than intended which leads to higher order noise and consequently to a decreased quality of the prepared states [58, 296]. Although the higher order noise cannot be completely eliminated, one possibility to at least reduce its strength would be to decrease the energy per pump pulse and, at the same time, to increase the repetition rate. This could, for example, be realized by utilizing a shorter enhancement cavity such that every pulse inside the cavity has to circulate n times before it interferes with the next pulse coming from the laser. A major problem of multiphoton experiments is

that they normally suffer from low count rates. In order to overcome this problem, a whole bundle of measures is required. For example, by Huang et al. [33], it was observed that replacing the lithium-triborate crystal used for second harmonic generation with a bismuth-triborate crystal led to a better collection efficiency of the down conversion photons and thus to higher count rates. Another drawback is the limited photon detection efficiency of state of the art silicon based avalanche photo diodes with a detection efficiency of around 60% in the near infrared. Since most commercially available single photon counter are trimmed towards a low dark count rate, it is, at least in principle, possible to increase the detection efficiency at the price of more dark counts. However, if the dark count rate becomes too high, this can also reduce the quality of the prepared state. Here, a possible counter measure would be to use shorter time windows for the coincidence logic. Alternatively, one could also utilize the recently developed superconducting bolometric detectors which achieve a detection efficiency close to 100% and are also intrinsically free of dark counts [297, 298]. However, the experimental overhead of operating superconducting circuits in comparison to semiconductor devices like avalanche photo diodes has to be taken into account. All in all, a combination of these measures should hopefully pave the way to increase the photon number to at least eight and to prepare states with a high fidelity. Apart from increasing the number of photons, an alternative approach to increase the size of the Hilbert space is to encode several qubits on the same photon using different degrees of freedom like, e.g., polarization and momentum using interferometers [299, 300]. For this purpose, it might also be worth considering to use integrated optical circuits which have been extensively discussed in the context of boson sampling [301–305] to implement interferometers.

At the time of writing, it is not yet clear which one of the different approaches for the realization of a quantum computer will prevail. For both ion traps and superconducting qubits, two-qubit gates with a process fidelity of close to 100% can be implemented now [28, 306]. Hence, there is good reason to assume that for stationary quantum computing, one of these approaches will turn out to be the most favorable one. For communication tasks however, photons are generally considered as the best approach. Thus, most likely, a future quantum network will have a hybrid structure with superconducting qubits or trapped ions for stationary quantum computing and data storage, and photons for communication tasks between the different nodes of the network.

6. Conclusions

A. Appendix

A.1. Additional remarks on convex optimization

The convex optimization algorithm developed for this thesis requires the knowledge of both the first derivatives and the Hessian matrix [57, 250]. Therefore, in the following, these quantities shall be derived for the target functions (see Eq. 4.15 and 4.16)

$$F_T^{MLE} = - \sum_{\lambda} f_{\lambda} \log(\text{Tr}(\varrho |\phi_{\lambda}\rangle \langle \phi_{\lambda}|)) \quad (\text{A.1})$$

and

$$F_T^{LS} = \sum_{\lambda} \frac{1}{w_{\lambda}} [f_{\lambda} - \text{Tr}(\varrho |\phi_{\lambda}\rangle \langle \phi_{\lambda}|)]^2 \quad (\text{A.2})$$

with experimentally observed frequencies f_{λ} when projecting on $|\phi_{\lambda}\rangle$ and weights $\frac{1}{w_{\lambda}}$. For simplicity, the shorthand notations

$$\Phi_{\lambda} = |\phi_{\lambda}\rangle \langle \phi_{\lambda}| \quad (\text{A.3})$$

$$p_{\lambda} = \text{Tr}[\varrho \Phi_{\lambda}] \quad (\text{A.4})$$

will be used. The state ϱ is parametrized as

$$\varrho(\vec{x}) = \frac{1}{2^N} \mathbb{1}^{\otimes N} + \sum_i x_i S_i \quad (\text{A.5})$$

where $x \in \mathbb{R}^{4N-1}$ and $\{S_i\}$ (together with $\frac{1}{2^n} \mathbb{1}^{\otimes N}$) is some appropriate operator basis like, e.g., the Pauli basis. As it will be necessary below, the derivative $\frac{\partial p_{\lambda}}{\partial x_i}$, is stated explicitly

$$\frac{\partial p_{\lambda}}{\partial x_i} = \frac{\partial \text{Tr}[\varrho \Phi_{\lambda}]}{\partial x_i} = \text{Tr}[S_i \Phi_{\lambda}]. \quad (\text{A.6})$$

It can easily be determined from Eq. A.4.

A.1.1. Maximum likelihood

First, the gradient of F_T^{MLE} for the above chosen parametrization of ϱ is determined. For a better readability, Eq. A.1 is rewritten as

$$F_T^{MLE} = - \sum_{\lambda} f_{\lambda} \log(p_{\lambda}). \quad (\text{A.7})$$

Appendix

Direct calculation shows that the gradient $\frac{\partial F_t^{MLE}}{\partial x_i}$ is given by

$$\frac{\partial F_t^{MLE}}{\partial x_i} = -\sum_{\lambda} \frac{f_{\lambda}}{p_{\lambda}} \frac{\partial p_{\lambda}}{\partial x_i} = -\sum_{\lambda} \frac{f_{\lambda}}{p_{\lambda}} \text{Tr}[S_i \Phi_{\lambda}]. \quad (\text{A.8})$$

Using the result from Eq. A.8, the Hessian matrix can directly be calculated as

$$\frac{\partial^2 F_t^{MLE}}{\partial x_i \partial x_j} = \sum_{\lambda} \frac{f_{\lambda}}{p_{\lambda}^2} \text{Tr}[S_i \Phi_{\lambda}] \text{Tr}[S_j \Phi_{\lambda}]. \quad (\text{A.9})$$

A.1.2. Least squares

For the following derivation, the weights $\frac{1}{w_{\lambda}}$ are set to $\frac{1}{p_{\lambda}}$ which is the common choice for photonic experiments [211]. Then Eq. A.2 can be written as

$$F_t^{LS} = \sum_{\lambda} \frac{(p_{\lambda} - f_{\lambda})^2}{p_{\lambda}}. \quad (\text{A.10})$$

The gradient of F_t^{LS} can be inferred as

$$\frac{\partial F_t^{LS}}{\partial x_i} = \sum_{\lambda} \left(\frac{2(p_{\lambda} - f_{\lambda})}{p_{\lambda}} \frac{\partial p_{\lambda}}{\partial x_i} - \frac{(p_{\lambda} - f_{\lambda})^2}{p_{\lambda}^2} \frac{\partial p_{\lambda}}{\partial x_i} \right) \quad (\text{A.11})$$

$$= \sum_{\lambda} \left(2 - \frac{(p_{\lambda} - f_{\lambda})}{p_{\lambda}} \right) \frac{(p_{\lambda} - f_{\lambda})}{p_{\lambda}} \frac{\partial p_{\lambda}}{\partial x_i} \quad (\text{A.12})$$

$$= \sum_{\lambda} \left(2 - 1 + \frac{f_{\lambda}}{p_{\lambda}} \right) \left(1 - \frac{f_{\lambda}}{p_{\lambda}} \right) \frac{\partial p_{\lambda}}{\partial x_i} \quad (\text{A.13})$$

$$= \sum_{\lambda} \left(1 + \frac{f_{\lambda}}{p_{\lambda}} \right) \left(1 - \frac{f_{\lambda}}{p_{\lambda}} \right) \frac{\partial p_{\lambda}}{\partial x_i} \quad (\text{A.14})$$

$$= \sum_{\lambda} \left(1 - \frac{f_{\lambda}^2}{p_{\lambda}^2} \right) \frac{\partial p_{\lambda}}{\partial x_i} \quad (\text{A.15})$$

$$= \sum_{\lambda} \left(1 - \frac{f_{\lambda}^2}{p_{\lambda}^2} \right) \text{Tr}[S_i \Phi_{\lambda}] \quad (\text{A.16})$$

$$(\text{A.17})$$

Accordingly, the Hessian is determined as

$$\frac{\partial^2 F_t^{LS}}{\partial x_i \partial x_j} = \sum_{\lambda} 2 \left(\frac{f_{\lambda}^2}{p_{\lambda}^3} \right) \text{Tr}[S_i \Phi_{\lambda}] \text{Tr}[S_j \Phi_{\lambda}]. \quad (\text{A.18})$$

A.2. Measurement settings for PI tomography

A.1.3. Barrier term

The gradient and the Hessian matrix does not only have to be known for the respective target function but also for the barrier term (see Eq. 4.22)

$$F_B = -\log[\det(\varrho)]. \quad (\text{A.19})$$

From a direct calculation, it follows that [57, 250]

$$\frac{\partial F_B}{\partial x_i} = -\text{Tr}[\varrho^{-1} S_i] \quad (\text{A.20})$$

$$\frac{\partial^2 F_B}{\partial x_i \partial x_j} = \text{Tr}[\varrho^{-1} S_i \varrho^{-1} B_j]. \quad (\text{A.21})$$

In order to obtain the gradient and the Hessian of the modified target function $\tilde{F}_T(\varrho)$, i.e., the target function plus the barrier term, one simply has to add the respective expressions of the original target function and the barrier term.

A.1.4. Implementation

The convex optimization algorithm described in chapter 4.1.3 was coded in C++ because the programming language is known for its high numerical performance. Additionally, the program was parallelized using openMP[®] to benefit from the multicore architecture of today's CPUs. Basic linear algebra operations like, e.g., determining the inverse of a matrix, were carried out using highly optimized numerical libraries like Armadillo or GotoBLAS. Since many operations had to be repeated in an identical way at every step of the algorithm, the corresponding results, were stored in lookuptables, see e.g. Eq. A.6. The program was compiled with the g++ compiler from the gnu compiler collection using the optimization level -O3. On a standard desktop PC with an AMD Phenom[™] II X4 840 quad core processor with 8GB memory, reconstruction of a four-qubit state took only a few seconds and a six-qubit state could be reconstructed in about seven minutes.

A.2. Measurement settings for PI tomography

In this appendix, the measurement settings used for PI tomography of four- and six-qubit Dicke states are given. As explained in chapter 4.2.2, the measurement operators for PI tomography are of the form $A_j^{\otimes N}$ with $A_j = \vec{n}_j \cdot \vec{\sigma} = n_j^x \sigma_x + n_j^y \sigma_y + n_j^z \sigma_z$ and $|\vec{n}_j| = 1$. Experimentally, as described in detail in section 2.3.2, the operators A_j are measured with a polarization analysis consisting of a HWP(Θ_1), a QWP(Θ_2) and a PBS. The angles given in Tab. A.1-A.3 were used in order to measure the corresponding A_j 's (expressed in terms of the respective n_j^x , n_j^y and n_j^z) in publications P4.1 and P4.2. Please note that the angles are not unique, i.e., generally, the same measurement setting can be realized with several different choices of Θ_1 and Θ_2 .

Appendix

j	n_j^x	n_j^y	n_j^z	HWP(Θ_1)	HWP(Θ_2)
1	0.3946	0.9173	-0.0530	7.78°	56.73°
2	0.6843	-0.5344	-0.4961	39.56°	106.20°
3	-0.4307	-0.2082	-0.8782	54.53°	96.01°
4	0.9903	-0.1388	0.0001	24.49°	93.99°
5	-0.0455	0.2979	-0.9535	5.02°	98.67°
6	-0.6880	0.7156	-0.1207	53.59°	67.15°
7	0.1821	-0.9542	-0.2372	53.77°	126.30°
8	0.8563	0.3493	0.3803	11.40°	79.78°
9	0.6330	0.0326	-0.7734	34.71°	89.06°
10	-0.3226	-0.7620	-0.5615	64.88°	114.80°
11	0.4119	0.6373	-0.6513	27.03°	70.21°
12	-0.6787	0.3455	-0.6481	51.53°	79.89°
13	-0.1934	0.8361	-0.5134	35.98°	61.63°
14	0.8639	0.4393	-0.2465	19.97°	76.97°
15	-0.1918	0.4158	0.8890	80.81°	77.71°

Table A.1.: Optimized measurement settings and angle settings of the corresponding wave plates in the polarization analyses for PI tomography of the state $|D_4^{(2)}\rangle$.

j	n_j^x	n_j^y	n_j^z	HWP(Θ_1)	HWP(Θ_2)
1	-0.8096	-0.4635	0.3601	80.40°	103.8°
2	-0.6506	-0.2752	0.7078	83.35°	97.99°
3	-0.6676	0.4786	0.5703	39.78°	104.30°
4	0.3366	0.4390	-0.8330	32.99°	76.98°
5	0.5555	-0.4444	0.7028	16.18°	103.20°
6	0.2141	-0.8069	0.5505	18.76°	116.90°
7	-0.2604	-0.9640	0.0546	89.10°	127.30°
8	-0.6521	-0.5568	0.5145	85.53°	106.90°
9	0.0099	-0.1144	-0.9934	46.50°	93.28°
10	-0.9204	0.2377	0.3104	68.72°	83.12°
11	-0.8738	-0.1626	-0.4583	62.92°	94.68°
12	0.4096	-0.0143	0.9121	6.25°	90.41°
13	0.2048	0.0665	-0.9765	41.09°	88.09°
14	-0.5149	-0.1140	0.8497	83.83°	93.27°
15	-0.4491	0.6434	-0.6200	43.97°	69.98°

Table A.2.: Randomly chosen measurement settings and angle settings of the corresponding wave plates in the polarization analyses for PI tomography of the state $|D_4^{(2)}\rangle$.

A.2. Measurement settings for PI tomography

j	n_j^x	n_j^y	n_j^z	HWP(Θ_1)	HWP(Θ_2)
1	1.0000	0	0	22.50°	0°
2	0	1.0000	0	0°	45.00°
3	0	0	1.0000	0°	0°
4	-0.1840	0.4866	-0.8540	55.32°	14.56°
5	0.2761	0.4409	-0.8540	47.06°	13.08°
6	-0.2420	-0.9238	-0.2967	71.67°	-56.26°
7	0.8242	-0.2374	-0.5141	27.06°	-6.87°
8	-0.5932	0.5688	-0.5698	47.87°	72.67°
9	-0.8922	0.4513	0.0161	74.46°	13.41°
10	0.2405	-0.8479	0.4726	21.24°	118.99°
11	-0.5998	0.1537	-0.7852	52.13°	85.58°
12	-0.8930	0.2191	-0.3931	58.40°	83.67°
13	-0.5902	0.7942	-0.1445	77.21°	26.29°
14	-0.4929	0.1832	0.8506	85.12°	5.28°
15	0.7114	0.1824	-0.6787	30.79°	84.75°
16	0.2199	0.8453	-0.4870	24.50°	61.15°
17	-0.3580	-0.2422	-0.9017	46.91°	-7.01°
18	0.5985	0.6802	0.4233	24.40°	21.43°
19	-0.8941	-0.4051	-0.1909	58.51°	-11.95°
20	0.5607	0.8215	-0.1036	38.92°	27.62°
21	0.3725	-0.6367	-0.6752	47.66°	109.77°
22	0.6012	0.6017	-0.5258	42.04°	18.50°
23	-0.6481	0.6890	0.3244	63.26°	68.22°
24	0.2087	0.6907	0.6924	15.11°	21.84°
25	0.2316	-0.9181	-0.3217	19.40°	-33.32°
26	-0.0757	0.3756	0.9237	83.31°	78.97°
27	0.8968	0.3563	-0.2623	21.36°	79.56°
28	-0.7470	-0.2677	-0.6085	61.59°	97.76°

Table A.3.: Optimized measurement settings and angle settings of the corresponding wave plates in the polarization analyses for PI tomography of the state $|D_6^{(3)}\rangle$.

Publication list

Publications related to this work (ordered by Chapter)

Wiesław Laskowski, Daniel Richart, **Christian Schwemmer**, Tomasz Paterek, and Harald Weinfurter,

Experimental Schmidt Decomposition and State Independent Entanglement Detection,

Phys. Rev. Lett. **108**, 240501 (2012).

Contribution: Carried out the experiment together with D.R., analyzed the data with D.R., and wrote parts of the manuscript.

Wiesław Laskowski, **Christian Schwemmer**, Daniel Richart, Lukas Knips, Tomasz Paterek, and Harald Weinfurter,

Optimized state-independent entanglement detection based on a geometrical threshold criterion,

Phys. Rev. A **88**, 022327 (2013).

Contribution: Performed the experiment together with D.R., analyzed the data with D.R. and L.K., and wrote parts of the manuscript.

Christian Schwemmer, Lukas Knips, Minh Cong Tran, Anna de Rosier, Wiesław Laskowski, Tomasz Paterek, and Harald Weinfurter,

Experimental multipartite entanglement without multipartite correlations,

Phys. Rev. Lett. **114**, 180501 (2015).

Contribution: Conceived the experiment with W.L, T.P., L.K., and H.W., carried out the experiment with L.K., analyzed the data together with L.K., and wrote parts of the manuscript.

Géza Tóth, Witlef Wieczorek, David Gross, Roland Krischek, **Christian Schwemmer** and Harald Weinfurter,

Permutationally Invariant Quantum Tomography,

Phys. Rev. Lett. **105**, 250403 (2010).

Contribution: Performed the experiment together with R.K. and W.W., analyzed the data with G.T., W.W., D.G., and R.K., discussed and edited the manuscript with all other coauthors.

Publication list

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,
Phys. Rev. Lett. **113**, 040503 (2014).

Contribution: Conceived the experiment with G.T. and H.W., carried out the experiment with A.N., analyzed the data together with A.N., T.M., and D.G., wrote the manuscript.

Tobias Moroder, Philipp Hyllus, Géza Tóth, **Christian Schwemmer**, Alexander Niggebaum, Stefanie Gaile, Otfried Gühne, and Harald Weinfurter,
Permutationally invariant state reconstruction,
New J. Phys. **14**, 105001 (2012).

Contribution: Edited the manuscript and discussed the paper with respect to an experimental context.

Christian Schwemmer, Lukas Knips, Daniel Richart, Tobias Moroder, Matthias Kleinmann, Otfried Gühne, and Harald Weinfurter,
Systematic Errors in Current Quantum State Tomography Tools,
Phys. Rev. Lett. **114**, 080403 (2015).

Contribution: Conceived the numerical experiments with L.K., T.M., M.K., D.R., and H.W., developed the optimization algorithm together with L.K. and D.R., carried out the numerical simulations with L.K. and D.R., wrote parts of the manuscript.

Further publications not mentioned in this work (ordered by time with the most recent first)

Witlef Wieczorek, Nikolai Kiesel, Christian Schmid, **Christian Schwemmer**, Géza Tóth, Wiesław Laskowski, and Harald Weinfurter
Experimental characterization of multipartite entanglement,
Planned to be submitted with Review of Modern Physics, *in preparation*.

Contribution: Wrote parts of the manuscript and discussed the paper with all other coauthors.

Lukas Knips, **Christian Schwemmer**, Nico Klein, Marcin Wieśniak, and Harald Weinfurter
Multipartite entanglement detection with minimal effort,
arXiv:1412.5881 (2014).

Contribution: Designed the experiment together with L.K., M.W., and H.W.,

performed the measurements with L.K. and N.K., analyzed the data with L.K. and N.K., wrote parts of the manuscript.

Philipp Hyllus, Wiesław Laskowski, Roland Krischek, **Christian Schwemmer**, Witlef Wieczorek, Harald Weinfurter, Luca Pezzé, and Augusto Smerzi, Fisher information and multiparticle entanglement, *Phys. Rev. A* **85**, 022321 (2012).

Contribution: Edited and discussed the manuscript.

Roland Krischek, **Christian Schwemmer**, Witlef Wieczorek, Harald Weinfurter, Philipp Hyllus, Luca Pezzé, and Augusto Smerzi, Useful Multiparticle Entanglement and Sub-Shot-Noise Sensitivity in Experimental Phase Estimation, *Phys. Rev. Lett.* **107**, 080504 (2011).

Contribution: conceived the experiment together with R.K., W.W., H.W., P.H., L.P., and A.S, performed the experiment with R.K. and W.W., analyzed the data with R.K., W.W., P.H., and L.P., and wrote parts of the manuscript.

Publication list

Bibliography

- [1] Planck, M., Zur Theorie des Gesetzes der Energieverteilung im Normalspectrum, *Verhandlungen der Deutschen physikalischen Gesellschaft*, **2**, 17 (1900).
- [2] Schilpp, P. A. (editor), Philosophen des 20. Jahrhunderts: Albert Einstein als Philosoph und Naturforscher, (Vieweg 1979).
- [3] Wheeler, J. A. and Zurek, W. H. (editors), Quantum Theory and Measurement, (Princeton Legacy Library 2014).
- [4] Einstein, A., Podolsky, B., and Rosen, N., Can Quantum-Mechanical Description of Physical Reality Be Considered Complete?, *Phys. Rev.*, **47**, 777 (1935).
- [5] Schrödinger, E., Die gegenwärtige Situation in der Quantenmechanik, *Naturwissenschaften*, **23**, 807 (1935).
- [6] Bell, J. S., On the Einstein-Podolsky-Rosen paradox, *Physics*, **1**, 195 (1964).
- [7] Freedman, S. J. and Clauser, J. F., Experimental Test of Local Hidden-Variable Theories, *Phys. Rev. Lett.*, **28**, 938 (1972).
- [8] Aspect, A., Grangier, P., and Roger, G., Experimental Tests of Realistic Local Theories via Bell's Theorem, *Phys. Rev. Lett.*, **47**, 460 (1981).
- [9] Aspect, A., Grangier, P., and Roger, G., Experimental Realization of Einstein-Podolsky-Rosen-Bohm *Gedankenexperiment* : A New Violation of Bell's Inequalities, *Phys. Rev. Lett.*, **49**, 91 (1982).
- [10] Weihs, G., Jennewein, T., Simon, C., Weinfurter, H., and Zeilinger, A., Violation of Bell's Inequality under Strict Einstein Locality Conditions, *Phys. Rev. Lett.*, **81**, 5039 (1998).
- [11] Rowe, M. A., Kielpinski, D., Meyer, V., Sackett, C. A., Itano, W. M., Monroe, C., and Wineland, D. J., Experimental violation of a Bell's inequality with efficient detection, *Nature*, **409**, 791 (2001).
- [12] Hofmann, J., Krug, M., Ortegel, N., Gérard, L., Weber, M., Rosenfeld, W., and Weinfurter, H., Heralded Entanglement Between Widely Separated Atoms, *Science*, **337**, 72 (2012).

Bibliography

- [13] Bernien, H., Hensen, B., Pfaff, W., Koolstra, G., Blok, M. S., Robledo, L., Tamini, T. H., Markham, M., Twitchen, D. J., Childress, L., and Hanson, R., Heralded entanglement between solid-state qubits separated by three metres, *Nature*, **497**, 86 (2013).
- [14] Giustina, M., Mech, A., Ramelow, S., Wittmann, B., Kofler, J., Beyer, J., and Brice Calkins, A. L., Gerrits, T., Nam, S. W., Ursin, R., and Anton, Z., Bell violation using entangled photons without the fair-sampling assumption, *Nature*, **497**, 227 (2013).
- [15] Feynman, R. P., Simulating physics with computers, *International journal of theoretical physics*, **21**, 467 (1982).
- [16] Lloyd, S., Universal Quantum Simulators, *Science*, **273**, 1073 (1996).
- [17] Bennett, C. H. and Wiesner, S. J., Communication via one- and two-particle operators on Einstein-Podolsky-Rosen states, *Phys. Rev. Lett.*, **69**, 2881 (1992).
- [18] Ekert, A. K., Quantum cryptography based on Bell's theorem, *Phys. Rev. Lett.*, **67**, 661 (1991).
- [19] Giovannetti, V., Lloyd, S., and Maccone, L., Quantum-Enhanced Measurements: Beating the Standard Quantum Limit, *Science*, **306**, 1330 (2004).
- [20] Bennett, C. H., Brassard, G., Crépeau, C., Jozsa, R., Peres, A., and Wootters, W. K., Teleporting an unknown quantum state via dual classical and Einstein-Podolsky-Rosen channels, *Phys. Rev. Lett.*, **70**, 1895 (1993).
- [21] Grover, L. K., Quantum Mechanics Helps in Searching for a Needle in a Haystack, *Phys. Rev. Lett.*, **79**, 325 (1997).
- [22] Shor, P. W., Polynomial-Time Algorithms for Prime Factorization and Discrete Logarithms on a Quantum Computer, *SIAM Journal on Computing*, **26**, 1484 (1997).
- [23] Schumacher, B., Quantum coding, *Phys. Rev. A*, **51**, 2738 (1995).
- [24] Kimble, H. J., The quantum internet, *Nature*, **453**, 1023 (2008).
- [25] Bloch, I., Quantum coherence and entanglement with ultracold atoms in optical lattices, *Nature*, **453**, 1016 (2008).
- [26] Blatt, R. and Wineland, D., Entangled states of trapped atomic ions, *Nature*, **453**, 1008 (2008).

-
- [27] Clarke, J. and Wilhelm, F. K., Superconducting quantum bits, *Nature*, **453**, 1031 (2008).
- [28] Barends, R., Kelly, J., Megrant, A., Veitia, A., Sank, D., Jeffrey, E., White, T.-C., Mutus, J., Fowler, A.-G., Campbell, B., *et al.*, Superconducting quantum circuits at the surface code threshold for fault tolerance, *Nature*, **508**, 500 (2014).
- [29] Kelly, J., Barends, R., Fowler, A. G., Megrant, A., Jeffrey, E., White, T. C., Sank, D., Mutus, J. Y., Campbell, B., Chen, Y., Chen, Z., Chiaro, B., Dunsworth, A., Hoi, I.-C., Neill, C., O'Malley, P. J. J., C. Quintana, P. R., Vainsencher, A., Wenner, J., Cleland, A. N., and Martinis, J. M., State preservation by repetitive error detection in a superconducting quantum circuit, *Nature*, **519**, 66 (2015).
- [30] Hanson, R. and Awschalom, D. D., Coherent manipulation of single spins in semiconductors, *Nature*, **453**, 1043 (2008).
- [31] Pan, J.-W., Chen, Z.-B., Lu, C.-Y., Weinfurter, H., Zeilinger, A., and Żukowski, M., Multiphoton entanglement and interferometry, *Rev. Mod. Phys.*, **84**, 777 (2012).
- [32] Häffner, H., Hänsel, W., Roos, C. F., Benhelm, J., Chek-al kar, D., Chwalla, M., Korber, T., Rapol, U. D., Riebe, M., Schmidt, P. O., Becher, C., Gühne, O., Dür, W., and Blatt, R., Scalable multiparticle entanglement of trapped ions, *Nature*, **438**, 643 (2005).
- [33] Huang, Y.-F., Liu, B.-H., Peng, L., Li, Y.-H., Li, L., Li, C.-F., and Guo, G.-C., Experimental generation of an eight-photon Greenberger Horne Zeilinger state, *Nature Communications*, **2**, 546 (2011).
- [34] Monz, T., Schindler, P., Barreiro, J. T., Chwalla, M., Nigg, D., Coish, W. A., Harlander, M., Hänsel, W., Hennrich, M., and Blatt, R., 14-Qubit Entanglement: Creation and Coherence, *Phys. Rev. Lett.*, **106**, 130506 (2011).
- [35] Yao, X.-C., Wang, T.-X., Xu, P., Lu, H., Pan, G.-S., Bao, X.-H., Peng, C.-Z., Lu, C.-Y., Chen, Y.-A., and Pan, J.-W., Observation of eight-photon entanglement, *Nature Photonics*, **6**, 225 (2012).
- [36] Demkowicz-Dobrzański, R., Kołodyński, J., and Guţă, M., The elusive Heisenberg limit in quantum-enhanced metrology, *Nature communications*, **3**, 1063 (2012).
- [37] Demkowicz-Dobrzanski, R. and Markiewicz, M., From quantum metrological precision bounds to quantum computation speed-up limits, *arXiv:1412.6111* (2014).

Bibliography

- [38] Knill, E., Laflamme, R., and Milburn, G. J., A scheme for efficient quantum computation with linear optics, *Nature*, **409**, 46 (2001).
- [39] Briegel, H. J. and Raussendorf, R., Persistent Entanglement in Arrays of Interacting Particles, *Phys. Rev. Lett.*, **86**, 910 (2001).
- [40] Raussendorf, R. and Briegel, H. J., A One-Way Quantum Computer, *Phys. Rev. Lett.*, **86**, 5188 (2001).
- [41] Lapaire, G. G., Kok, P., Dowling, J. P., and Sipe, J. E., Conditional linear-optical measurement schemes generate effective photon nonlinearities, *Phys. Rev. A*, **68**, 042314 (2003).
- [42] Nielsen, M. A., Optical Quantum Computation Using Cluster States, *Phys. Rev. Lett.*, **93**, 040503 (2004).
- [43] Kok, P., Munro, W. J., Nemoto, K., Ralph, T. C., Dowling, J. P., and Milburn, G. J., Linear optical quantum computing with photonic qubits, *Rev. Mod. Phys.*, **79**, 135 (2007).
- [44] Nguyen, H. S., Sallen, G., Voisin, C., Roussignol, P., Diederichs, C., and Cassabois, G., Ultra-coherent single photon source, *Applied Physics Letters*, **99**, 261904 (2011).
- [45] Laucht, A., Pütz, S., Günthner, T., Hauke, N., Saive, R., Frédérick, S., Bichler, M., Amann, M.-C., Holleitner, A. W., Kaniber, M., and Finley, J. J., A Waveguide-Coupled On-Chip Single-Photon Source, *Phys. Rev. X*, **2**, 011014 (2012).
- [46] Mizuochi, N., Makino, T., Kato, H., Takeuchi, D., Ogura, M., Okushi, H., Nothhaft, M., Neumann, P., Gali, A., Jelezko, F., *et al.*, Electrically driven single-photon source at room temperature in diamond, *Nature photonics*, **6**, 299 (2012).
- [47] Gazzano, O., Almeida, M. P., Nowak, A. K., Portalupi, S. L., Lemaître, A., Sagnes, I., White, A. G., and Senellart, P., Entangling Quantum-Logic Gate Operated with an Ultrabright Semiconductor Single-Photon Source, *Phys. Rev. Lett.*, **110**, 250501 (2013).
- [48] He, Y.-M., He, Y., Wei, Y.-J., Wu, D., Atatüre, M., Schneider, C., Höfling, S., Kamp, M., Lu, C.-Y., and Pan, J.-W., On-demand semiconductor single-photon source with near-unity indistinguishability, *Nature Nanotechnology*, **8**, 213 (2013).
- [49] Mandel, L. and Wolf, E., *Optical Coherence and Quantum Optics*, (Cambridge University Press 1995).

- [50] Kwiat, P. G., Mattle, K., Weinfurter, H., Zeilinger, A., Sergienko, A. V., and Shih, Y., New High-Intensity Source of Polarization-Entangled Photon Pairs, *Phys. Rev. Lett.*, **75**, 4337 (1995).
- [51] Badziag, P., Brukner, Č., Laskowski, W., Paterek, T., and Żukowski, M., Experimentally Friendly Geometrical Criteria for Entanglement, *Phys. Rev. Lett.*, **100**, 140403 (2008).
- [52] Laskowski, W., Richart, D., Schwemmer, C., Paterek, T., and Weinfurter, H., Experimental Schmidt Decomposition and State Independent Entanglement Detection, *Phys. Rev. Lett.*, **108**, 240501 (2012).
- [53] Laskowski, W., Schwemmer, C., Richart, D., Knips, L., Paterek, T., and Weinfurter, H., Optimized state-independent entanglement detection based on a geometrical threshold criterion, *Phys. Rev. A*, **88**, 022327 (2013).
- [54] Schwemmer, C., Knips, L., Tran, M. C., de Rosier, A., Laskowski, W., Paterek, T., and Weinfurter, H., Genuine Multipartite Entanglement without Multipartite Correlations, *Phys. Rev. Lett.*, **114**, 180501 (2015).
- [55] Gross, D., Liu, Y.-K., Flammia, S. T., Becker, S., and Eisert, J., Quantum State Tomography via Compressed Sensing, *Phys. Rev. Lett.*, **105**, 150401 (2010).
- [56] Tóth, G., Wieczorek, W., Gross, D., Krischek, R., Schwemmer, C., and Weinfurter, H., Permutationally Invariant Quantum Tomography, *Phys. Rev. Lett.*, **105**, 250403 (2010).
- [57] Moroder, T., Hyllus, P., Tóth, G., Schwemmer, C., Niggebaum, A., Gaile, S., Gühne, O., and Weinfurter, H., Permutationally invariant state reconstruction, *New J. Phys.*, **14**, 105001 (2012).
- [58] Schwemmer, C., Tóth, G., Niggebaum, A., Moroder, T., Gross, D., Gühne, O., and Weinfurter, H., Experimental Comparison of Efficient Tomography Schemes for a Six-Qubit State, *Phys. Rev. Lett.*, **113**, 040503 (2014).
- [59] Schwemmer, C., Knips, L., Richart, D., Weinfurter, H., Moroder, T., Kleinmann, M., and Gühne, O., Systematic Errors in Current Quantum State Tomography Tools, *Phys. Rev. Lett.*, **114**, 080403 (2015).
- [60] Tóth, G. and Gühne, O., Entanglement detection in the stabilizer formalism, *Phys. Rev. A*, **72**, 022340 (2005).
- [61] Tóth, G., Detection of multipartite entanglement in the vicinity of symmetric Dicke states, *J. Opt. Soc. Am. B*, **24**, 275 (2007).

Bibliography

- [62] Nielsen, M. A. and Chuang, I. L., Quantum Computation and Quantum Information, (Cambridge University Press 2000).
- [63] Kurzyński, P., Paterek, T., Ramanathan, R., Laskowski, W., and Kaszlikowski, D., Correlation Complementarity Yields Bell Monogamy Relations, *Phys. Rev. Lett.*, **106**, 180402 (2011).
- [64] Lücke, B., Peise, J., Vitagliano, G., Arlt, J., Santos, L., Tóth, G., and Klempt, C., Detecting Multiparticle Entanglement of Dicke States, *Phys. Rev. Lett.*, **112**, 155304 (2014).
- [65] de Vicente, J. I. and Huber, M., Multipartite entanglement detection from correlation tensors, *Phys. Rev. A*, **84**, 062306 (2011).
- [66] Jungnitsch, B., Moroder, T., and Gühne, O., Taming Multiparticle Entanglement, *Phys. Rev. Lett.*, **106**, 190502 (2011).
- [67] Laskowski, W., Markiewicz, M., Paterek, T., and Żukowski, M., Correlation-tensor criteria for genuine multiqubit entanglement, *Phys. Rev. A*, **84**, 062305 (2011).
- [68] Clauser, J. F., Horne, M. A., Shimony, A., and Holt, R. A., Proposed Experiment to Test Local Hidden-Variable Theories, *Phys. Rev. Lett.*, **23**, 15, 880 (1969).
- [69] Mermin, N. D., Quantum mechanics vs local realism near the classical limit: A Bell inequality for spin s , *Phys. Rev. D*, **22**, 356 (1980).
- [70] Ardehali, M., Bell inequalities with a magnitude of violation that grows exponentially with the number of particles, *Phys. Rev. A*, **46**, 5375 (1992).
- [71] Żukowski, M. and Brukner, Č., Bell's Theorem for General N -Qubit States, *Phys. Rev. Lett.*, **88**, 210401 (2002).
- [72] Mendel, J., Tutorial on higher-order statistics (spectra) in signal processing and system theory: theoretical results and some applications, *Proceedings of the IEEE*, **79**, 278 (1991).
- [73] Chua, K. C., Chandran, V., Acharya, U. R., and Lim, C. M., Application of higher order statistics/spectra in biomedical signals—A review, *Medical Engineering & Physics*, **32**, 679 (2010).
- [74] Kaszlikowski, D., Sen(De), A., Sen, U., Vedral, V., and Winter, A., Quantum Correlation without Classical Correlations, *Phys. Rev. Lett.*, **101**, 070502 (2008).

-
- [75] Bennett, C. H., Grudka, A., Horodecki, M., Horodecki, P., and Horodecki, R., Postulates for measures of genuine multipartite correlations, *Phys. Rev. A*, **83**, 012312 (2011).
- [76] Giorgi, G. L., Bellomo, B., Galve, F., and Zambrini, R., Genuine Quantum and Classical Correlations in Multipartite Systems, *Phys. Rev. Lett.*, **107**, 190501 (2011).
- [77] Cramer, M., Plenio, M. B., Flammia, S. T., Somma, R., Gross, D., Bartlett, S. D., Landon-Cardinal, O., Poulin, D., and Liu, Y.-K., Efficient quantum state tomography, *Nature Communications*, **1**, 149 (2010).
- [78] Dicke, R. H., Coherence in Spontaneous Radiation Processes, *Phys. Rev.*, **93**, 99 (1954).
- [79] Stockton, J. K., Geremia, J. M., Doherty, A. C., and Mabuchi, H., Characterizing the entanglement of symmetric many-particle spin- $\frac{1}{2}$ systems, *Phys. Rev. A*, **67**, 022112 (2003).
- [80] Lehmann, E. L. and Casella, G., Theory of Point Estimation, (Springer Berlin Heidelberg 1998).
- [81] van Enk, S. J. and Blume-Kohout, R., When quantum tomography goes wrong: drift of quantum sources and other errors, *New J. Phys.*, **14**, 025024 (2012).
- [82] Werner, R. F., Quantum states with Einstein-Podolsky-Rosen correlations admitting a hidden-variable model, *Phys. Rev. A*, **40**, 4277 (1989).
- [83] Dür, W., Vidal, G., and Cirac, J. I., Three qubits can be entangled in two inequivalent ways, *Phys. Rev. A*, **62**, 062314 (2000).
- [84] Gühne, T. and Tóth, G., Entanglement detection, *Physics Reports*, **474(1-6)**, 1 (2009).
- [85] Horodecki, R., Horodecki, P., Horodecki, M., and Horodecki, K., Quantum entanglement, *Rev. Mod. Phys.*, **81**, 865 (2009).
- [86] Eisert, J. and Plenio, M. B., A comparison of entanglement measures, *J. Mod. Optic.*, **46**, 145 (1999).
- [87] Plenio, M. B. and Virmani, S., An introduction to entanglement measures, *Quantum Information & Computation*, **7**, 1 (2007).
- [88] Wieczorek, W., Kiesel, N., Schmid, C., Schwemmer, C., Tóth, G., Laskowski, W., and Weinfurter, H., Experimental characterization of multipartite entanglement, *to be published*.

Bibliography

- [89] Kiesel, N., Experiments on Multiphoton Entanglement, Ph.D. thesis, Ludwig-Maximilians-Universität München (2007).
- [90] Schmid, C., Multi-photon entanglement and applications in quantum information, Ph.D. thesis, Ludwig-Maximilians-Universität München (2008).
- [91] Wieczorek, W., Multi-Photon Entanglement: Experimental Observation, Characterization, and Application of up to Six-Photon Entangled States, Ph.D. thesis, Ludwig-Maximilians-Universität München (2009).
- [92] Bouwmeester, D., Ekert, A., and Zeilinger, A. (editors), The Physics of Quantum Information, (Springer 2000), 1st edition.
- [93] Bruss, D. and Leuchs, G. (editors), Lectures on Quantum Information, (Wiley VCH 2007).
- [94] Mermin, N. D., Quantum Computer Science: An Introduction (English), (Cambridge University Press 2007).
- [95] Vedral, V., Introduction to Quantum Information Science, (Oxford University Press 2013).
- [96] Uhlmann, A., The “transition” probability in the state space of a $*$ -algebra, *Reports on Mathematical Physics*, **9**, 273 (1976).
- [97] Jozsa, R., Fidelity for Mixed Quantum States, *J. Mod. Opt.*, **41**, 2315 (1994).
- [98] Seevinck, M. and Uffink, J., Partial separability and entanglement criteria for multiqubit quantum states, *Phys. Rev. A*, **78**, 032101 (2008).
- [99] Bennett, C. H., Bernstein, H. J., Popescu, S., and Schumacher, B., Concentrating partial entanglement by local operations, *Phys. Rev. A*, **53**, 2046 (1996).
- [100] Ghiu, I., Bourennane, M., and Karlsson, A., Entanglement-assisted local transformations between inequivalent classes of three-particle entangled states, *Physics Letters A*, **287**, 12 (2001).
- [101] Acín, A., Andrianov, A., Costa, L., Jané, E., Latorre, J. I., and Tarrach, R., Generalized Schmidt Decomposition and Classification of Three-Quantum-Bit States, *Phys. Rev. Lett.*, **85**, 1560 (2000).
- [102] Verstraete, F., Dehaene, J., Moor, B. D., and Verschelde, H., Four qubits can be entangled in nine different ways, *Phys. Rev. A*, **65**, 052112 (2002).

-
- [103] Lamata, L., Leon, J., Salgado, D., and Solano, E., Inductive entanglement classification of four qubits under stochastic local operations and classical communication, *Phys. Rev. A*, **75**, 022318 (2007).
- [104] Lamata, L., Leon, J., Salgado, D., and Solano, E., Inductive classification of multipartite entanglement under stochastic local operations and classical communication, *Phys. Rev. A*, **74**, 052336 (2006).
- [105] Kruszynska, C. and Kraus, B., Local entanglability and multipartite entanglement, *Phys. Rev. A*, **79**, 052304 (2009).
- [106] Miyake, A., Classification of multipartite entangled states by multidimensional determinants, *Phys. Rev. A*, **67**, 012108 (2003).
- [107] Hein, M., Eisert, J., and Briegel, H. J., Multiparty entanglement in graph states, *Phys. Rev. A*, **69**, 062311 (2004).
- [108] Raussendorf, R., Browne, D. E., and Briegel, H. J., Measurement-based quantum computation on cluster states, *Phys. Rev. A*, **68**, 022312 (2003).
- [109] Kempe, J., Bacon, D., Lidar, D. A., and Whaley, K. B., Theory of decoherence-free fault-tolerant universal quantum computation, *Phys. Rev. A*, **63**, 042307 (2001).
- [110] Cabello, A., N -Particle N -Level Singlet States: Some Properties and Applications, *Phys. Rev. Lett.*, **89**, 100402 (2002).
- [111] Lamas-Linares, A., Howell, J. C., and Bouwmeester, D., Stimulated emission of polarization-entangled photons, *Nature*, **412**, 887 (2001).
- [112] Zanardi, P. and Rasetti, M., Noiseless Quantum Codes, *Phys. Rev. Lett.*, **79**, 3306 (1997).
- [113] Horodecki, M., Horodecki, P., and Horodecki, R., Mixed-State Entanglement and Distillation: Is there a “Bound” Entanglement in Nature?, *Phys. Rev. Lett.*, **80**, 5239 (1998).
- [114] Amsalem, E. and Bourennane, M., Experimental four-qubit bound entanglement, *Nat. Phys.*, **5**, 748 (2009).
- [115] Lavoie, J., Kaltenbaek, R., Piani, M., and Resch, K. J., Experimental Bound Entanglement in a Four-Photon State, *Phys. Rev. Lett.*, **105**, 130501 (2010).
- [116] Klumper, A., Schadschneider, A., and Zittartz, J., Equivalence and solution of anisotropic spin-1 models and generalized t-J fermion models in one dimension, *Journal of Physics A: Mathematical and General*, **24**, L955 (1991).

Bibliography

- [117] Fannes, M., Nachtergaele, B., and Werner, R., Finitely correlated states on quantum spin chains, *Communications in Mathematical Physics*, **144**, 443 (1992).
- [118] Perez-Garcia, D., Verstraete, F., Wolf, M. M., and Cirac, J. I., Matrix Product State Representations, *Quantum Inf. Comput.*, **7**, 401 (2007).
- [119] Gottesman, D., Stabilizer Codes and Quantum Error Correction, *arXiv:quant-ph/9705052* (1997).
- [120] Hao, J.-C., Li, C.-F., and Guo, G.-C., Controlled dense coding using the Greenberger-Horne-Zeilinger state, *Phys. Rev. A*, **63**, 054301 (2001).
- [121] Hillery, M., Bužek, V., and Berthiaume, A., Quantum secret sharing, *Phys. Rev. A*, **59**, 1829 (1999).
- [122] Walther, P., Resch, K. J., Rudolph, T., Schenck, E., Weinfurter, H., Vedral, V., Aspelmeyer, M., and Zeilinger, A., Experimental one-way quantum computing, *Nature*, **434**, 169 (2005).
- [123] Gühne, O., Bodoky, F., and Blaauboer, M., Multiparticle entanglement under the influence of decoherence, *Phys. Rev. A*, **78**, 060301 (2008).
- [124] Sen(De), A., Sen, U., Wiesniak, M., Kaszlikowski, D., and Żukowski, M., Multiqubit W states lead to stronger nonclassicality than Greenberger-Horne-Zeilinger states, *Phys. Rev. A*, **68**, 062306 (2003).
- [125] Tashima, T., Ozdemir, S. K., Yamamoto, T., Koashi, M., and Imoto, N., Elementary optical gate for expanding an entanglement web, *Phys. Rev. A*, **77**, 030302 (2008).
- [126] Tóth, G., Knapp, C., Gühne, O., and Briegel, H. J., Optimal Spin Squeezing Inequalities Detect Bound Entanglement in Spin Models, *Phys. Rev. Lett.*, **99**, 250405 (2007).
- [127] Wieczorek, W., Kiesel, N., Schmid, C., and Weinfurter, H., Multiqubit entanglement engineering via projective measurements, *Phys. Rev. A*, **79**, 022311 (2009).
- [128] Eibl, M., Kiesel, N., Bourennane, M., Kurtsiefer, C., and Weinfurter, H., Experimental Realization of a Three-Qubit Entangled W State, *Phys. Rev. Lett.*, **92**, 077901 (2004).
- [129] Kiesel, N., Schmid, C., Tóth, G., Solano, E., and Weinfurter, H., Experimental Observation of Four-Photon Entangled Dicke State with High Fidelity, *Phys. Rev. Lett.*, **98**, 063604 (2007).

-
- [130] Prevedel, R., Cronenberg, G., Tame, M. S., Paternostro, M., Walther, P., Kim, M. S., and Zeilinger, A., Experimental Realization of Dicke States of up to Six Qubits for Multiparty Quantum Networking, *Phys. Rev. Lett.*, **103**, 020503 (2009).
- [131] Wieczorek, W., Krischek, R., Kiesel, N., Michelberger, P., Tóth, G., and Weinfurter, H., Experimental Entanglement of a Six-Photon Symmetric Dicke State, *Phys. Rev. Lett.*, **103**, 020504 (2009).
- [132] Bruß, D., Characterizing Entanglement, *Jour. Math. Phys.*, **43**, 4237 (2002).
- [133] Eisert, J. and Gross, D., Multi-particle entanglement, *arXiv:quant-ph/0505149v2* (2005).
- [134] van Enk, S. J., Lütkenhaus, N., and Kimble, H. J., Experimental procedures for entanglement verification, *Phys. Rev. A*, **75**, 052318 (2007).
- [135] Gisin, N., Bell's inequality holds for all non-product states, *Physics Letters A*, **154**, 5-6, 201 (1991).
- [136] Hyllus, P., Gühne, O., Bruß, D., and Lewenstein, M., Relations between entanglement witnesses and Bell inequalities, *Phys. Rev. A*, **72**, 012321 (2005).
- [137] Zeilinger, A., Experiment and the foundations of quantum physics, *Rev. Mod. Phys.*, **71**, S288 (1999).
- [138] Bell, J. S., *Speakable and Unspeakable in Quantum Mechanics: Collected Papers on Quantum Philosophy*, (Cambridge University Press 2004).
- [139] Genovese, M., Research on hidden variable theories: A review of recent progresses, *Physics Reports*, **413**, 6, 319 (2005).
- [140] Buhrman, H., Cleve, R., Massar, S., and de Wolf, R., Nonlocality and communication complexity, *Rev. Mod. Phys.*, **82**, 665 (2010).
- [141] Brunner, N., Cavalcanti, D., Pironio, S., Scarani, V., and Wehner, S., Bell nonlocality, *Rev. Mod. Phys.*, **86**, 419 (2014).
- [142] Peres, A., Separability Criterion for Density Matrices, *Phys. Rev. Lett.*, **77**, 1413 (1996).
- [143] Horodecki, M., Horodecki, P., and Horodecki, R., Separability of mixed states: necessary and sufficient conditions, *Phys. Lett. A*, **223**, 1 (1996).
- [144] Terhal, B. M., Bell inequalities and the separability criterion, *Phys. Lett. A*, **271**, 319 (2000).

Bibliography

- [145] Helstrom, C. W., Quantum detection and estimation theory, *J. Stat. Phys.*, **1**, 231 (1969).
- [146] Braunstein, S. L. and Caves, C. M., Statistical distance and the geometry of quantum states, *Phys. Rev. Lett.*, **72**, 3439 (1994).
- [147] Dowling, J. P., Quantum optical metrology – the lowdown on high-N00N states, *Contemporary Physics*, **49**, 125 (2008).
- [148] Paris, M. G. A., Quantum Estimation for Quantum Technology, *International Journal of Quantum Information*, **07**, 125 (2009).
- [149] Caves, C. M. and Shaji, A., Quantum-circuit guide to optical and atomic interferometry, *Optics Communications*, **283**, 695 (2010).
- [150] R. Demkowicz-Dobrzański, J. K., M. Jarzyna, Quantum limits in optical interferometry, *arXiv:1405.7703* (2014).
- [151] Pezzé, L. and Smerzi, A., Quantum theory of phase estimation, *arXiv:1411.5164* (2014).
- [152] Tóth, G. and Apellaniz, I., Quantum metrology from a quantum information science perspective, *Journal of Physics A: Mathematical and Theoretical*, **47**, 424006 (2014).
- [153] Pezze, L. and Smerzi, A., Entanglement, Nonlinear Dynamics, and the Heisenberg Limit, *Phys. Rev. Lett.*, **102**, 100401 (2009).
- [154] Hyllus, P., Laskowski, W., Krischek, R., Schwemmer, C., Wieczorek, W., Weinfurter, H., Pezzé, L., and Smerzi, A., Fisher information and multipartite entanglement, *Phys. Rev. A*, **85**, 022321 (2012).
- [155] Tóth, G., Multipartite entanglement and high-precision metrology, *Phys. Rev. A*, **85**, 022322 (2012).
- [156] Kitagawa, M. and Ueda, M., Squeezed spin states, *Phys. Rev. A*, **47**, 5138 (1993).
- [157] Wineland, D. J., Bollinger, J. J., Itano, W. M., and Heinzen, D. J., Squeezed atomic states and projection noise in spectroscopy, *Phys. Rev. A*, **50**, 67 (1994).
- [158] Tóth, G., Knapp, C., Gühne, O., and Briegel, H. J., Spin squeezing and entanglement, *Phys. Rev. A*, **79**, 042334 (2009).
- [159] Vitagliano, G., Apellaniz, I., Egusquiza, I. L., and Tóth, G., Spin squeezing and entanglement for an arbitrary spin, *Phys. Rev. A*, **89**, 032307 (2014).

-
- [160] Gühne, O. and Seevinck, M., Separability criteria for genuine multiparticle entanglement, *New J. Phys.*, **12**, 053002 (2010).
- [161] Uhlmann, A., Entropy and Optimal Decompositions of States Relative to a Maximal Commutative Subalgebra, *Open Sys. Information Dyn.*, **5**, 209 (1998).
- [162] Geza Toth, O. G., Tobias Moroder, Evaluation of convex roof entanglement measures, *arXiv:1409.3806* (2014).
- [163] Bennett, C. H., DiVincenzo, D. P., Smolin, J. A., and Wootters, W. K., Mixed-state entanglement and quantum error correction, *Phys. Rev. A*, **54**, 3824 (1996).
- [164] Vidal, G. and Werner, R. F., Computable measure of entanglement, *Phys. Rev. A*, **65**, 032314 (2002).
- [165] Love, P., van den Brink, A., Smirnov, A., Amin, M., Grajcar, M., Ilichev, E., Izmalkov, A., and Zagoskin, A., A Characterization of Global Entanglement, *Quant. Inf. Proc.*, **6**, 187 (2007).
- [166] Sabín, C. and García-Alcaine, G., A classification of entanglement in three-qubit systems, *The European Physical Journal D*, **48**, 435 (2008).
- [167] Hill, S. and Wootters, W. K., Entanglement of a Pair of Quantum Bits, *Phys. Rev. Lett.*, **78**, 5022 (1997).
- [168] Wootters, W. K., Entanglement of Formation of an Arbitrary State of Two Qubits, *Phys. Rev. Lett.*, **80**, 2245 (1998).
- [169] Aolita, L. and Mintert, F., Measuring Multipartite Concurrence with a Single Factorizable Observable, *Phys. Rev. Lett.*, **97**, 050501 (2006).
- [170] Mintert, F., Kus, M., and Buchleitner, A., Concurrence of mixed multipartite quantum states, *Phys. Rev. Lett.*, **95**, 260502 (2005).
- [171] Walborn, S. P., Ribeiro, P. H. S., Davidovich, L., Mintert, F., and Buchleitner, A., Experimental determination of entanglement by a projective measurement, *Phys. Rev. A*, **75**, 032338 (2007).
- [172] Walborn, S. P., Souto Ribeiro, P. H., Davidovich, L., Mintert, F., and Buchleitner, A., Experimental determination of entanglement with a single measurement, *Nature*, **440**, 1022 (2006).
- [173] Aolita, L., Buchleitner, A., and Mintert, F., Scalable method to estimate experimentally the entanglement of multipartite systems, *Phys. Rev. A*, **78**, 022308 (2008).

Bibliography

- [174] Mintert, F. and Buchleitner, A., Observable Entanglement Measure for Mixed Quantum States, *Phys. Rev. Lett.*, **98**, 140505 (2007).
- [175] Barnum, H. and Linden, N., Monotones and invariants for multi-particle quantum states, *J. Phys. A*, **34**, 6787 (2001).
- [176] Shimony, A., Degree of entanglement, *Ann. N. Y. Acad. Sci.*, **755**, 675 (1995).
- [177] Wei, T.-C. and Goldbart, P. M., Geometric measure of entanglement and applications to bipartite and multipartite quantum states, *Phys. Rev. A*, **68**, 042307 (2003).
- [178] Eisert, J., Brandao, F. G. S. L., and Audenaert, K. M. R., Quantitative entanglement witnesses, *New J. Phys.*, **9**, 46 (2007).
- [179] Gühne, O., Reimpell, M., and Werner, R. F., Estimating Entanglement Measures in Experiments, *Phys. Rev. Lett.*, **98**, 110502 (2007).
- [180] Gühne, O., Reimpell, M., and Werner, R. F., Lower bounds on entanglement measures from incomplete information, *Phys. Rev. A*, **77**, 052317 (2008).
- [181] Vidal, G. and Tarrach, R., Robustness of entanglement, *Phys. Rev. A*, **59**, 141 (1999).
- [182] Coffman, V., Kundu, J., and Wootters, W. K., Distributed entanglement, *Phys. Rev. A*, **61**, 052306 (2000).
- [183] Lohmayer, R., Osterloh, A., Siewert, J., and Uhlmann, A., Entangled Three-Qubit States without Concurrence and Three-Tangle, *Phys. Rev. Lett.*, **97**, 260502 (2006).
- [184] Horodecki, M., Entanglement measures, *Quantum Inform. Compu.*, **1**, 3 (2001).
- [185] Chruściński, D. and Sarbicki, G., Entanglement witnesses: construction, analysis and classification, *Journal of Physics A: Mathematical and Theoretical*, **47**, 483001 (2014).
- [186] Louisell, W. H., Yariv, A., and Siegman, A. E., Quantum Fluctuations and Noise in Parametric Processes. I., *Phys. Rev.*, **124**, 1646 (1961).
- [187] Kleinman, D. A., Theory of Optical Parametric Noise, *Phys. Rev.*, **174**, 1027 (1968).
- [188] Hong, C. K. and Mandel, L., Theory of parametric frequency down conversion of light, *Phys. Rev. A*, **31**, 2409 (1985).

- [189] Keller, T. E. and Rubin, M. H., Theory of two-photon entanglement for spontaneous parametric down-conversion driven by a narrow pump pulse, *Phys. Rev. A*, **56**, 1534 (1997).
- [190] Chwedeńczuk, J. and Wasilewski, W., Intensity of parametric fluorescence pumped by ultrashort pulses, *Phys. Rev. A*, **78**, 063823 (2008).
- [191] Trojek, P., Efficient Generation of Photonic Entanglement and Multiparty Quantum Communication, Ph.D. thesis, Ludwig-Maximilians-Universität München (2007).
- [192] Krischek, R., Experimental multi-photon entanglement and quantum enhanced metrology, Ph.D. thesis, Ludwig-Maximilians-Universität München (2011).
- [193] Saleh, B. and Teich, M., Fundamentals of Photonics, (John Wiley & Sons, Inc. 1991).
- [194] Dmitriev, V. G., Gurzadyan, G. G., and Nikogosyan, D. N., Handbook of Nonlinear Optical Crystals, (Springer-Verlag Berlin Heidelberg 1999).
- [195] Michelberger, P., Femtosecond pulsed enhancement cavity for multi-photon entanglement studies, Master's thesis, Ludwig-Maximilians-Universität München (2009).
- [196] Krischek, R., Wieczorek, W., Ozawa, A., Kiesel, N., Michelberger, P., Udem, T., and Weinfurter, H., Ultraviolet enhancement cavity for ultrafast nonlinear optics and high-rate multiphoton entanglement experiments, *Nat. Photon.*, **4**, 170 (2010).
- [197] Hänsch, T. W. and Couillaud, B., Laser frequency stabilization by polarization spectroscopy of a reflecting reference cavity, *Opt. Comm.*, **35**, 441 (1980).
- [198] Bennink, R. S., Liu, Y., Earl, D. D., and Grice, W. P., Spatial distinguishability of photons produced by spontaneous parametric down-conversion with a focused pump, *Phys. Rev. A*, **74**, 023802 (2006).
- [199] Oberparleiter, M., Effiziente Erzeugung verschränkter Photonenpaare, Ph.D. thesis, Ludwig-Maximilians-Universität München (2002).
- [200] Kurtsiefer, C., Oberparleiter, M., and Weinfurter, H., High-efficiency entangled photon pair collection in type-II parametric fluorescence, *Phys. Rev. A*, **64**, 023802 (2001).

Bibliography

- [201] Di Giuseppe, G., Haiberger, L., De Martini, F., and Sergienko, A. V., Quantum interference and indistinguishability with femtosecond pulses, *Phys. Rev. A*, **56**, R21 (1997).
- [202] Grice, W. P., Erdmann, R., Walmsley, I. A., and Branning, D., Spectral distinguishability in ultrafast parametric down-conversion, *Phys. Rev. A*, **57**, R2289 (1998).
- [203] Atatüre, M., Sergienko, A. V., Jost, B. M., Saleh, B. E. A., and Teich, M. C., Partial Distinguishability in Femtosecond Optical Spontaneous Parametric Down-Conversion, *Phys. Rev. Lett.*, **83**, 1323 (1999).
- [204] Kliger, D. S., Lewis, J. W., and Randall, C. E., Polarized Light in Optics and Spectroscopy, (Academic Press Limited 1990).
- [205] Goos, F. and Hänchen, H., Ein neuer und fundamentaler Versuch zur Totalreflexion, *Annalen der Physik*, **436**, 333 (1947).
- [206] Trojek, P. and Weinfurter, H., Collinear source of polarization-entangled photon pairs at nondegenerate wavelengths, *Appl. Phys. Lett.*, **92**, 211103 (2008).
- [207] Wieczorek, W., Schmid, C., Kiesel, N., Pohlner, R., Gühne, O., and Weinfurter, H., Experimental Observation of an Entire Family of Four-Photon Entangled States, *Phys. Rev. Lett.*, **101**, 010503 (2008).
- [208] Zeilinger, A., General properties of lossless beam splitters in interferometry, *Am. J. Phys.*, **49**, 882 (1981).
- [209] Peres, A., Quantum Theory: Concepts and Methods, (Springer Netherlands 2002).
- [210] Bužek, V., Hillery, M., and Werner, F., Universal-NOT gate, *Journal of Modern Optics*, **47**, 211 (2000).
- [211] James, D. F. V., Kwiat, P. G., Munro, W. J., and White, A. G., Measurement of qubits, *Phys. Rev. A*, **64**, 052312 (2001).
- [212] Resch, K. J., Walther, P., and Zeilinger, A., Full Characterization of a Three-Photon Greenberger-Horne-Zeilinger State Using Quantum State Tomography, *Phys. Rev. Lett.*, **94**, 070402 (2005).
- [213] Steffen, M., Ansmann, M., McDermott, R., Katz, N., Bialczak, R. C., Lucero, E., Neeley, M., Weig, E. M., Cleland, A. N., and Martinis, J. M., State Tomography of Capacitively Shunted Phase Qubits with High Fidelity, *Phys. Rev. Lett.*, **97**, 050502 (2006).

-
- [214] DiCarlo, L., Chow, J. M., Gambetta, J. M., Bishop, L. S., Johnson, B. R., Schuster, D. I., Majer, J., Blais, A., Frunzio, L., Girvin, S. M., and Schoelkopf, R. J., Demonstration of two-qubit algorithms with a superconducting quantum processor, *Nature*, **460**, 240 (2009).
- [215] Rippe, L., Julsgaard, B., Walther, A., Ying, Y., and Kröll, S., Experimental quantum-state tomography of a solid-state qubit, *Phys. Rev. A*, **77**, 022307 (2008).
- [216] Thew, R. T., Nemoto, K., White, A. G., and Munro, W. J., Qudit quantum-state tomography, *Phys. Rev. A*, **66**, 012303 (2002).
- [217] Agnew, M., Leach, J., McLaren, M., Roux, F. S., and Boyd, R. W., Tomography of the quantum state of photons entangled in high dimensions, *Phys. Rev. A*, **84**, 062101 (2011).
- [218] Lvovsky, A. I. and Raymer, M. G., Continuous-variable optical quantum-state tomography, *Rev. Mod. Phys.*, **81**, 299 (2009).
- [219] Hradil, Z., Quantum-state estimation, *Phys. Rev. A*, **55**, R1561 (1997).
- [220] Perez-Garcia, D., Verstraete, F., Cirac, J., and Wolf, M., PEPS as unique ground states of local Hamiltonians, *Quant. Inf. Comp.*, **8**, 650 (2008).
- [221] Flammia, S. T., Gross, D., Bartlett, S. D., and Somma, R., Heralded Polynomial-Time Quantum State Tomography, *arXiv:1002.3839* (2010).
- [222] Landon-Cardinal, O. and Liu, Y.-K. and Poulin, D., Efficient Direct Tomography for Matrix Product States, *arXiv:1002.4632* (2010).
- [223] Baumgratz, T., Nüßeler, A., Cramer, M., and Plenio, M. B., A scalable maximum likelihood method for quantum state tomography, *New J. Phys.*, **14**, 125004 (2013).
- [224] Baumgratz, T., Gross, D., Cramer, M., and Plenio, M. B., Scalable Reconstruction of Density Matrices, *Phys. Rev. Lett.*, **111**, 020401 (2013).
- [225] Haack, G., Steffens, A., Eisert, J., and Hübener, R., Continuous matrix product state tomography of quantum transport experiments, *arXiv:1504.04194* (2015).
- [226] Steffens, A., Riofrío, C. A., Hübener, R., and Eisert, J., Quantum field tomography, *arXiv:1406.363* (2014).
- [227] Steffens, A., Friesdorf, M., Langen, T., Rauer, B., Schweigler, T., Hübener, R., Schmiedmayer, J., Riofrío, C. A., and Eisert, J., Towards experimental quantum field tomography with ultracold atoms, *arXiv:1406.3632* (2014).

Bibliography

- [228] Flammia, S. T., Gross, D., Liu, Y.-K., and Eisert, J., Quantum tomography via compressed sensing: error bounds, sample complexity and efficient estimators, *New J. Phys.*, **14**, 095022 (2012).
- [229] Liu, W.-T., Zhang, T., Liu, J.-Y., Chen, P.-X., and Yuan, J.-M., Experimental Quantum State Tomography via Compressed Sampling, *Phys. Rev. Lett.*, **108**, 170403 (2012).
- [230] Ohliger, N., Nesme, V., and Eisert, J., Efficient and feasible state tomography of quantum many-body systems, *New J. Phys.*, **14**, 015024 (2012).
- [231] Smith, A., Riofrío, C. A., Anderson, B. E., Sosa-Martinez, H., Deutsch, I. H., and Jessen, P. S., Quantum state tomography by continuous measurement and compressed sensing, *Phys. Rev. A*, **87**, 030102 (2013).
- [232] Carpentier, A., Eisert, J., and andsu Richard Nickl, D. G., Uncertainty Quantification for Matrix Compressed Sensing and Quantum Tomography Problems, *arXiv:1504.03234* (2015).
- [233] Yin, J. O. S. and van Enk, S. J., Information criteria for efficient quantum state estimation, *Phys. Rev. A*, **83**, 062110 (2011).
- [234] Banaszek, K., Cramer, M., and Gross, D., Focus on quantum tomography, *New J. Phys.*, **15**, 125020 (2013).
- [235] Fischer, D. G., Kienle, S. H., and Freyberger, M., Quantum-state estimation by self-learning measurements, *Phys. Rev. A*, **61**, 032306 (2000).
- [236] Hannemann, T., Reiss, D., Balzer, C., Neuhauser, W., Toschek, P. E., and Wunderlich, C., Self-learning estimation of quantum states, *Phys. Rev. A*, **65**, 050303 (2002).
- [237] Sugiyama, T., Finite Sample Analysis in Quantum Estimation, in Finite Sample Analysis in Quantum Estimation, Springer Theses, 89–112, (Springer Japan 2014).
- [238] Sugiyama, Takanori and Turner, Peter S. and Muraio, Mio, Adaptive experimental design for one-qubit state estimation with finite data based on a statistical update criterion, *Phys. Rev. A*, **85**, 052107 (2012).
- [239] Řeháček, J., Englert, B.-G., and Kaszlikowski, D., Minimal qubit tomography, *Phys. Rev. A*, **70**, 052321 (2004).
- [240] Huszár, F. and Houlby, N. M. T., Adaptive Bayesian quantum tomography, *Phys. Rev. A*, **85**, 052120 (2012).

-
- [241] Kravtsov, K. S., Straupe, S. S., Radchenko, I. V., Houlby, N. M. T., Huszár, F., and Kulik, S. P., Experimental adaptive Bayesian tomography, *Phys. Rev. A*, **87**, 062122 (2013).
- [242] Okamoto, R., Iefuji, M., Oyama, S., Yamagata, K., Imai, H., Fujiwara, A., and Takeuchi, S., Experimental Demonstration of Adaptive Quantum State Estimation, *Phys. Rev. Lett.*, **109**, 130404 (2012).
- [243] D'Ariano, G. M. and Paris, M. G. A., Adaptive quantum homodyne tomography, *Phys. Rev. A*, **60**, 518 (1999).
- [244] Paris, M. and Řeháček, J. (editors), Quantum State Estimation, (Springer Berlin Heidelberg 2004).
- [245] Penrose, A., A generalized inverse for matrices, *Proc. Cambridge Phil. Soc.*, **51**, 406 (1955).
- [246] Langford, N. K., Errors in quantum tomography: diagnosing systematic versus statistical errors, *New J. Phys.*, **15**, 035003 (2013).
- [247] Řeháček, J., Hradil, Z., Knill, E., and Lvovsky, A. I., Diluted maximum-likelihood algorithm for quantum tomography, *Phys. Rev. A*, **75**, 042108 (2007).
- [248] Fiurášek, J., Maximum-likelihood estimation of quantum measurement, *Phys. Rev. A*, **64**, 024102 (2001).
- [249] Vardi, Y. and Lee, D., From Image Deblurring to Optimal Investments: Maximum Likelihood Solutions for Positive Linear Inverse Problems, *Journal of the Royal Statistical Society. Series B (Methodological)*, **55**, pp. 569 (1993).
- [250] Boyd, S. and Vandenberghe, S., Convex optimization, (Cambridge University Press 2004).
- [251] Reimpell, M., Quantum Information and Convex Optimization, Ph.D. thesis, Technische Universität Carolo-Wilhelmina zu Braunschweig (2007).
- [252] Nocedal, J. and Wright, S. J., Numerical Optimization, (Springer New York 1999).
- [253] Wolfe, P., Convergence Conditions for Ascent Methods, *SIAM Review*, **11**(2), 226 (1969).
- [254] Broyden, C. G., A class of methods for solving nonlinear simultaneous equations, *Math. Comp.*, **19**, 577 (1965).

Bibliography

- [255] Davidon, W. C., Variable Metric Method for Minimization, *Technical Report ANL-5990 (revised)*, Argonne National Laboratory, Argonnem Il (1959).
- [256] Davidon, W. C., Variable Metric Method for Minimization, *SIAM J. Optim.*, **1**, 1 (1991).
- [257] Broyden, C. G., The Convergence of a Class of Double-rank Minimization Algorithms 1. General Considerations, *IMA J. Appl. Math.*, **6**, 76 (1970).
- [258] Fletcher, R., A new approach to variable metric algorithms, *The Computer Journal*, **13**, 317 (1970).
- [259] Goldfarb, D., A family of variable-metric methods derived by variational means, *Math. Comp.*, **24**, 23 (1970).
- [260] Shanno, D. F., Conditioning of quasi-Newton methods for function minimization, *Math. Comp.*, **24**, 647 (1970).
- [261] Hestenes, M. R. and Stiefel, E., Methods of conjugate gradients for solving linear systems, *Journal of Research of the National Bureau of Standards*, **49**, 409 (1952).
- [262] Bogdanov, Y. I., Brida, G., Bukeev, I. D., Genovese, M., Kravtsov, K. S., Kulik, S. P., Moreva, E. V., Soloviev, A. A., and Shurupov, A. P., Statistical estimation of the quality of quantum-tomography protocols, *Phys. Rev. A*, **84**, 042108 (2011).
- [263] Christandl, M. and Renner, R., Reliable Quantum State Tomography, *Phys. Rev. Lett.*, **109**, 120403 (2012).
- [264] Cramer, M. and Plenio, M. B., Reconstructing quantum states efficiently, *arXiv:1002.3780* (2010).
- [265] Flammia, S. T. and Liu, Y.-K., Direct Fidelity Estimation from Few Pauli Measurements, *Phys. Rev. Lett.*, **106**, 230501 (2011).
- [266] Heinosaari, T., Mazzarella, L., and Wolf, M., Quantum Tomography under Prior Information, *Communications in Mathematical Physics*, **318**, 355 (2013).
- [267] O. Landon-Cardinal, and D. Poulin, Practical learning method for multi-scale entangled states, *New J. Phys.*, **14**, 085004 (2012).
- [268] Taylor, H. L., Banks, S. C., and McCoy, J. F., Deconvolution with the ℓ_1 norm, *GEOPHYSICS*, **44**, 39 (1979).

-
- [269] Santosa, F. and Symes, W., Linear Inversion of Band-Limited Reflection Seismograms, *SIAM Journal on Scientific and Statistical Computing*, **7**, 1307 (1986).
- [270] Candes, E. and Tao, T., Near-Optimal Signal Recovery From Random Projections: Universal Encoding Strategies?, *Information Theory, IEEE Transactions on*, **52**, 5406 (2006).
- [271] Donoho, D., Compressed sensing, *Information Theory, IEEE Transactions on*, **52**, 1289 (2006).
- [272] Guta, M., Kypraios, T., and Dryden, I., Rank-based model selection for multiple ions quantum tomography, *New J. Phys.*, **14**, 105002 (2012).
- [273] Chase, B. A. and Geremia, J. M., Collective processes of an ensemble of spin-1/2 particles, *Phys. Rev. A*, **78**, 052101 (2008).
- [274] D’Ariano, G. M., Maccone, L., and Painsi, M., Spin tomography, *Journal of Optics B: Quantum and Semiclassical Optics*, **5**, 77 (2003).
- [275] Adamson, R. B. A., Shalm, L. K., Mitchell, M. W., and Steinberg, A. M., Multiparticle State Tomography: Hidden Differences, *Physical Review Letters*, **98**, 043601 (2007).
- [276] Adamson, R. B. A., Turner, P. S., Mitchell, M. W., and Steinberg, A. M., Detecting hidden differences via permutation symmetries, *Phys. Rev. A*, **78**, 033832 (2008).
- [277] Shalm, L. K., Adamson, R. B. A., and Steinberg, A. M., Squeezing and over-squeezing of triphotons, *Nature*, **457**, 67 (2009).
- [278] Tóth, G., Wieczorek, W., Krischek, R., Kiesel, N., Michelberger, P., and Weinfurter, H., Practical methods for witnessing genuine multi-qubit entanglement in the vicinity of symmetric states, *New J. Phys*, **11**, 083002 (2009).
- [279] Tóth, G. and Gühne, O., Detecting Genuine Multipartite Entanglement with Two Local Measurements, *Phys. Rev. Lett.*, **94**, 060501 (2005).
- [280] Cirac, J. I., Ekert, A. K., and Macchiavello, C., Optimal Purification of Single Qubits, *Phys. Rev. Lett.*, **82**, 4344 (1999).
- [281] Blume-Kohout, R., Hedged Maximum Likelihood Quantum State Estimation, *Phys. Rev. Lett.*, **105**, 200504 (2010).
- [282] Ferrie, C. and Blume-Kohout, R., Minimax quantum tomography: the ultimate bounds on accuracy, *arXiv:1503.03100* (2015).

Bibliography

- [283] Blume-Kohout, R., Robust error bars for quantum tomography, [*arXiv:1202.5270*](#) (2012).
- [284] Shang, J., Ng, H. K., Sehrawat, A., Li, X., and Englert, B.-G., Optimal error regions for quantum state estimation, *New J. Phys.*, **15**, 123026 (2013).
- [285] Sugiyama, Takanori and Turner, Peter S. and Murao, Mio, Precision-Guaranteed Quantum Tomography, *Phys. Rev. Lett.*, **111**, 160406 (2013).
- [286] Shang, J., Seah, Y.-L., Ng, H. K., Nott, D. J., and Englert, B.-G., Monte Carlo sampling from the quantum state space. I, *New J. Phys.*, **17**, 043017 (2015).
- [287] Seah, Y.-L., Shang, J., Ng, H. K., Nott, D. J., and Englert, B.-G., Monte Carlo sampling from the quantum state space. II, *New J. Phys.*, **17**, 043018 (2015).
- [288] Moroder, T., Kleinmann, M., Schindler, P., Monz, T., Gühne, O., and Blatt, R., Certifying Systematic Errors in Quantum Experiments, *Phys. Rev. Lett.*, **110**, 180401 (2013).
- [289] Rosset, D., Ferretti-Schöbitz, R., Bancal, J.-D., Gisin, N., and Liang, Y.-C., Imperfect measurement settings: Implications for quantum state tomography and entanglement witnesses, *Phys. Rev. A*, **86**, 062325 (2012).
- [290] Stark, C., Self-consistent tomography of the state-measurement Gram matrix, *Phys. Rev. A*, **89**, 052109 (2014).
- [291] van Enk, S. J. and Blume-Kohout, R., When quantum tomography goes wrong: drift of quantum sources and other errors, *New J. Phys.*, **15**, 025024 (2013).
- [292] Merkel, S. T., Gambetta, J. M., Smolin, J. A., Poletto, S., Córcoles, A. D., Johnson, B. R., Ryan, C. A., and Steffen, M., Self-consistent quantum process tomography, *Phys. Rev. A*, **87**, 062119 (2013).
- [293] Jaynes, E. T., Probability Theory, (Cambridge University Press 2003).
- [294] Hoeffding, W., Probability Inequalities for Sums of Bounded Random Variables, *Journal of the American Statistical Association*, **58**, pp. 13 (1963).
- [295] Mood, A. F., Introduction to the Theory of Statistics, (McGraw-Hill Inc., New York 1974).
- [296] Laskowski, W., Wieśniak, M., Żukowski, M., Bourennane, M., and Weinfurter, H., Interference contrast in multisource few-photon optics, *Journal of Physics B: Atomic, Molecular and Optical Physics*, **42**, 114004 (2009).

-
- [297] Lita, A. E., Miller, A. J., and Nam, S. W., Counting near-infrared single-photons with 95% efficiency, *Opt. Express*, **16**, 3032 (2008).
- [298] Ramelow, S., Mech, A., Giustina, M., Gröblacher, S., Wieczorek, W., Beyer, J., Lita, A., Calkins, B., Gerrits, T., Nam, S. W., Zeilinger, A., and Ursin, R., Highly efficient heralding of entangled single photons, *Opt. Express*, **21**, 6707 (2013).
- [299] Gao, W.-B., Lu, C.-Y., Yao, X.-C., Xu, P., Gühne, O., Goebel, A., Chen, Y.-A., Peng, C.-Z., Chen, Z.-B., and Pan, J.-W., Experimental demonstration of a hyper-entangled ten-qubit Schrödinger cat state, *Nature Physics*, **6**, 5, 331 (2010).
- [300] Knips, L., Schwemmer, C., Klein, N., Wieśniak, M., and Weinfurter, H., Multipartite entanglement detection with minimal effort, *arXiv:1412.5881* (2014).
- [301] Broome, M. A., Fedrizzi, A., Rahimi-Keshari, S., Dove, J., Aaronson, S., Ralph, T. C., and White, A. G., Photonic Boson Sampling in a Tunable Circuit, *Science*, **339**, 794 (2013).
- [302] Spring, J. B., Metcalf, B. J., Humphreys, P. C., Kolthammer, W. S., Jin, X.-M., Barbieri, M., Datta, A., Thomas-Peter, N., Langford, N. K., Kundys, D., Gates, J. C., Smith, B. J., Smith, P. G. R., and Walmsley, I. A., Boson Sampling on a Photonic Chip, *Science*, **339**, 798 (2013).
- [303] Tillmann, M., Dakić, B., Heilmann, R., Nolte, S., Szameit, A., and Walther, P., Experimental boson sampling, *Nature Photonics*, **7**, 540 (2013).
- [304] Crespi, A., Osellame, R., Ramponi, R., Brod, D. J., Galvão, E. F., Spagnolo, N., Vitelli, C., Maiorino, E., Mataloni, P., and Sciarrino, F., Integrated multimode interferometers with arbitrary designs for photonic boson sampling, *Nature Photonics*, **7**, 545 (2013).
- [305] Ralph, T., Quantum computation: Boson sampling on a chip, *Nature Photonics*, **7**, 514 (2013).
- [306] Benhelm, J., Kirchmair, G., Roos, C. F., and Blatt, R., Towards fault-tolerant quantum computing with trapped ions, *Nat. Phys.*, **4**, 463 (2008).

Bibliography

Acknowledgments

A doctoral thesis cannot be written without the support of many others! Therefore, I would like to thank all those who made this work possible. *Cui honorem, honorem!*

At first, I want to express my sincere gratitude to my advisor Prof. Dr. Harald Weinfurter who gave me the possibility to carry out this research work in the fascinating field of quantum optics. He always had an open ear for the needs and problems of a young experimenter. In the past five years, we had countless fascinating and pleasurable discussions about our experiments and the physics behind them. Thank you for that!

I also want to thank my former colleagues Roland Krischek and Witlef Wiczorek for introducing me into the multiphoton experiment and all the fascinating discussion about physics and the world in general. Furthermore, I want to thank Lukas Knips and Daniel Richart with whom I carried out most of the experiments described in this thesis. Writing and optimizing Matlab and C++ programs together was really great fun! In the course of the last years, they became far more than just colleagues. Without their support, especially in mentally tough times, this work might not have been finished! Special thanks goes to Thomas Schulte-Herbrüggen for organizing and letting me participate in all those wonderful QCCC meetings and workshops.

A special thanks goes to all collaborators I had the honor to work with, namely, Philipp Hyllus, David Gross, Otfried Gühne, Matthias Kleinmann, Wiesław Laskowski, Tobias Moroder, Tomasz Paterek, Luca Pezzé, Augusto Smerzi, Géza Tóth, and Marcin Wieśniak. Furthermore, I would like to express my gratitude to our former students Alex Niggebaum, Yvo Fischer and Nico Klein for the good collaboration and for pushing the experiments forward. I also want to thank Thomas Reimann for all the highly entertaining discussions about physics, photography and life in general. I hope, one day, we find the time to continue our discussions! I am also thankful to my former lab neighbor Caro Hahn for launching our journal club, the good conversations we had and our hiking tours.

Most people of our working group are not located in Garching but at the city center of Munich. Although the way from Garching to Munich can be quite long, particularly when the subway is not working, there has always been a strong feeling of togetherness and a great cooperativeness. Especially when there were electronic problems or new circuits had to be designed, Sebastian Nauwerth, Markus Rau and Wenjamin Rosenfeld turned out as real wizards! Whenever there were

Acknowledgments

problems with one of the down conversion sources, Daniel Schlenk's expertise was invaluable.

I want to thank Daniel Burchardt, Florian Henkel, Michael Krug, Lars Liebermeister, Gwenaëlle Melen, Norbert Ortegel, Kai Redeker, Daniel Schlenk, Almut Tröller, Markus Weber, and Martin Zeitlmair for the weekly breakfasts and all the fun we had enjoying the nightlife of Munich. A special thanks goes to Michael Krug and Daniel Schlenk for pushing me to start swimming on a regular basis.

I also want to thank my old friends from Erlangen, namely, Mareike Bugla, Elisa Gärtner, Martin Hupfer, Hermann Kalb, Mathias Mühlbacher, Tobias Röthlingshöfer, Florian Seperant, and Tobias Wirth, for not forgetting me. I should have spent more time with you, I know! Johannes Mittmann and Thomas Pohl, there are no better friends than you! Andrea und Bernhard Lohrey, I hope to see you soon and on a more regular basis!

Last but not least, I want to thank my parents for giving me the opportunity to go to university and my brother for all his warmth and for coloring up my life! My aunt Elfriede was a great help for my family and me in the last two years which have been really difficult. Thank you for that!

Curriculum vitae

Personal data

Name	Christian Michael Schwemmer
Date of birth	23rd September 1982
Place of birth	Sulzbach-Rosenberg
Nationality	German

Education

2009 to present	PhD thesis in the Experimental Quantum Physics group of Prof. Dr. Harald Weinfurter at the Max-Planck-Institute of Quantum Optics in Garching on “Efficient Tomography and Entanglement Detection of Multiphoton States” as member of the international PhD program “Quantum Computing, Control and Communication” which is part of the Elitenetwork of Bavaria.
08/2008-04/2009	Diploma thesis in the group of Prof. Dr. Alexey Ustinov at the Karlsruher Institut für Technologie, Karlsruhe, Germany on “Noise Measurements of Cryogenic Amplifiers for Qubit Experiments”.
10/2003-08/2008	Diploma studies at the Friedrich-Alexander-Universität Erlangen-Nürnberg, Erlangen, Germany.
10/2005-06/2006	Semester abroad at Imperial College, London, UK, within the Erasmus exchange program.
07/2002-03/2003	Military service.
2000-2002	Grammar school, Christoph-Jacob-Treu-Gymnasium in Lauf a. d. Pegnitz, Germany.
1994-2000	Grammar school, Paul-Pfinzing-Gymnasium in Hersbruck, Germany.
1989-1994	Elementary school, Grund- und Mittelschule Happurg, Germany.

Scholarships

2003-2009	e-fellows.net – Das Online-Stipendium
2008	Scholarship of the Eva Schleip-Stiftung

Ich sage euch: man muss noch Chaos in sich haben,
um einen tanzenden Stern gebären zu können.

Friedrich Nietzsche, *Also sprach Zarathustra*

