

# **Mathematical Structures in Quantum Information Theory: Tensors, Correlations and State Estimation**

DISSERTATION

zur Erlangung des Grades eines Doktors  
der Naturwissenschaften

vorgelegt von

Jonathan Steinberg

eingereicht bei der Naturwissenschaftlich-Technischen Fakultät  
der Universität Siegen,  
Siegen 2023



Gutachter:

Prof. Dr. Otfried Gühne

Prof. Dr. Karol Życzkowski

Prüfer:

Prof. Dr. Otfried Gühne

Prof. Dr. Karol Życzkowski

Prof. Dr. Tommy Hofmann

PD Dr. Michael Johanning

Prof. Dr. Gregor Nickel

Tag der mündlichen Prüfung:

03.08.2023

## Abstract

This thesis is devoted to different aspects of quantum information science as well as the closely related topic of multilinear algebra. We present several results in the fields of quantum measurement theory, the theory of bipartite Bell inequalities, the activation of nonlocal quantum correlations, the identification of resourceful multipartite quantum states as well as the characterization of the eigenstructure of highly symmetric real-valued tensors. Gearing towards applications of quantum technology, we develop scalable methods that allow for the simultaneous prediction of many observables of a multi-qubit system.

First, we investigate how quantum measurements with many outcomes can be simulated by measurements with fewer outcomes. In particular, we present a minimal scheme that certifies this simulability based on correlations. Further, we analyze this minimal scheme with respect to noise robustness. Afterwards, we pick up the quantum measurement problem and discuss whether the realization of a partially observed measurement is compatible with the universality of the unitary time evolution.

Second, we introduce a family of bipartite Bell inequalities and associated quantum correlations that allow for an extremely low detection efficiency as well as high robustness to noise. Further, we discuss how these inequalities can be optimized by means of symmetry considerations. Subsequently, we examine the phenomenon of activation of quantum correlations. We develop methods that allow for rigorous statements on the statistical significance of such an experimental demonstration. These methods include the construction of a suitable confidence polytope as well as an algorithm to determine the correlation class of a quantum state.

Third, we present an algorithm that allows for finding maximally resourceful multipartite quantum states. We provide a rigorous proof of convergence and apply it to multiple quantifiers of quantum resources, e.g., the geometric measure of entanglement. This reveals an interesting connection to so-called absolutely maximally entangled states. Then, we discuss the eigenstructure of certain highly symmetric tensors, whose construction is based on simplex frames. We provide a full characterization of the eigenvectors for an arbitrary number of parties and local dimension two. Further, we discuss whether the eigenvectors can be obtained by the power iteration method.

The last part of this thesis is concerned with scalable methods that allow for simultaneously predicting many expectation values of a multi-qubit system with high accuracy. For this purpose, we extend the technique of classical shadows, originally based on projective measurements, to generalized measurements. This yields a simple formulation, allowing for the incorporation of symmetries and the possibility of optimizing the measurement directions towards a set of targeted observables. Moreover, we combine classical shadows with error mitigation techniques, rendering the incorporation of preparation errors in the estimation of many expectation values possible.

## Zusammenfassung

In dieser Dissertation werden verschiedene Fragestellungen aus den Bereichen der Quanteninformationstheorie und dem damit verknüpften Gebiet der multilinearen Algebra untersucht. Es wird eine Vielzahl von Resultaten auf den Gebieten der quantenmechanischen Messtheorie, der bipartiten Bellschen Ungleichungen und deren Detektionseffizienz, der Aktivierung von quantenmechanischen Korrelationen, der Charakterisierung von Verschränkung multipartiter Systeme, als auch der Eigenstruktur von sogenannten supersymmetrischen reellen Tensoren erbracht. Ferner werden skalierbare Methoden zur Berechnung von Erwartungswerten von Observablen diskutiert und deren Relevanz für Quantencomputer erörtert.

Im ersten Abschnitt dieser Arbeit wird untersucht, inwiefern sich quantenmechanische Messungen mit vielen Messausgängen vermöge Messungen mit weniger Ausgängen simulieren lassen. Dabei stellt sich insbesondere die Frage, wie sich diese Simulierbarkeit auf Grundlage experimenteller Daten zertifizieren lässt. Wir präsentieren ein minimales Szenario für diese Zertifizierung und diskutieren die Robustheit bezüglich experimenteller Fehler. Im Anschluss greifen wir das Messproblem der Quantenmechanik auf und diskutieren, inwiefern die Realisierung einer nur partiell beobachteten Messung kompatibel mit der unitären Zeitentwicklung der Quantenmechanik ist.

Im zweiten Abschnitt präsentieren wir eine Familie von bipartiten bellschen Ungleichungen und dazugehörigen Quantenkorrelationen, welche eine hohe Toleranz gegenüber ineffizienten Detektoren und experimentellem Rauschen besitzt. Ferner wird diskutiert, wie die erhaltenen bellschen Ungleichungen auf Grundlage von Symmetriebetrachtungen weiter optimiert werden können. Im Anschluss betrachten wir das Phänomen der Aktivierung von Quantenkorrelationen. In diesem Kontext werden Methoden entwickelt, welche einen experimentellen Nachweis mit hoher statistischer Sicherheit ermöglichen. Diese umfassen die Konstruktion von einem neuartigen Konfidenz-Polytop als auch einem Algorithmus, der die Korrelationsklasse eines Quantenzustands bestimmen kann.

Im dritten Abschnitt widmen wir uns dem Auffinden von multipartiten Zuständen, welche für gewisse Zwecke und Protokolle besonders hilfreich sind. Dazu entwickeln wir einen Algorithmus, präsentieren einen entsprechenden Konvergenzbeweis und illustrieren die Flexibilität durch eine Vielzahl von Anwendungen. Als Beispiel diskutieren wir im Detail das geometrische Maß der Verschränkung und beobachten eine Verbindung zu sogenannten absolut-maximal verschränkten Zuständen. Im Anschluss untersuchen wir die Eigenstruktur von reellen Tensoren, welche ausgehend von einem Simplex-Frame konstruiert wurden. Hier liefern wir eine vollständige Klassifizierung der Eigenstruktur für beliebige Modenzahlen und lokale Dimension zwei. Zusätzlich erörtern wir, inwiefern die gefundenen Eigenwerte mittels der Potenzmethode berechnet

werden können.

Ein weiterer großer Teil dieser Arbeit beschäftigt sich mit skalierbaren Methoden zur Bestimmung von Erwartungswerten von Observablen großer Quantensysteme. Dazu erweitern wir das Konzept der auf projektiven Messungen basierenden klassischen Schatten auf verallgemeinerte Messungen. Dies liefert eine einfachere Formulierung, erlaubt die Einbeziehung von Symmetrien, als auch die Optimierung der Messrichtungen. Außerdem kombinieren wir klassische Schatten mit Techniken der Fehlerminimierung. Diese Kombination erlaubt es mithilfe klassischer Computer den Effekt experimenteller Ungenauigkeiten auf Erwartungswerte von Observablen zu minimieren und somit den Bereich der Anwendbarkeit von Quantencomputern zu erweitern.



# Contents

<b>Introduction</b>	<b>1</b>
<b>1 Preliminaries</b>	<b>5</b>
1.1 Mathematical framework of quantum theory . . . . .	5
1.1.1 Quantum states . . . . .	5
1.1.2 Quantum measurements . . . . .	12
1.1.3 Quantum channels . . . . .	18
1.1.4 The quantum measurement problem . . . . .	22
1.2 Correlations and quantum theory . . . . .	27
1.2.1 The EPR argument . . . . .	27
1.2.2 Quantum entanglement . . . . .	28
1.2.3 Bell nonlocality . . . . .	42
1.2.4 Quantum steering . . . . .	54
1.2.5 The measurement problem revisited . . . . .	58
1.3 Further concepts and applications . . . . .	61
1.3.1 Graphs and graph states . . . . .	61
1.3.2 Quantum computing . . . . .	68
1.4 Computational and algorithmic aspects . . . . .	73
1.4.1 General form and terminology . . . . .	74
1.4.2 Lagrange duality . . . . .	75
1.4.3 Convex optimization . . . . .	76
<b>2 Certifying irreducible measurements in a prepare-and-measure scenario</b>	<b>79</b>
2.1 Motivation . . . . .	79
2.2 Concepts and notation . . . . .	81
2.2.1 The prepare-and-measure scenario . . . . .	81
2.2.2 The structure of measurements . . . . .	81
2.2.3 Unambiguous state discrimination . . . . .	82
2.3 Certification from correlations . . . . .	83
2.3.1 Certifying trichotomic measurements . . . . .	85
2.3.2 The minimal scenario . . . . .	85
2.3.3 The geometry of trichotomic correlations . . . . .	87



2.4	Robustness of trichotomic correlations . . . . .	89
2.4.1	The computation of upper bounds . . . . .	89
2.4.2	The computation of lower bounds . . . . .	90
2.5	Estimating the state space dimension . . . . .	94
2.6	Experimental feasibility . . . . .	95
2.7	Outlook and discussion . . . . .	97
<b>3</b>	<b>Partially observed measurements in the Wigner's friend scenario</b>	<b>99</b>
3.1	Motivation . . . . .	99
3.2	A minimal protocol and the assumptions . . . . .	101
3.2.1	The scenario . . . . .	101
3.2.2	Relative event by incomplete information . . . . .	102
3.2.3	Freedom of choice and locality . . . . .	103
3.3	Combining the assumptions . . . . .	104
3.3.1	The correlation polytope . . . . .	105
3.3.2	The quantum violation . . . . .	106
3.4	Subtleties and an extended protocol . . . . .	108
3.5	Further protocols . . . . .	109
3.5.1	Alternative protocol I . . . . .	109
3.5.2	Alternative protocol II . . . . .	111
3.5.3	Alternative protocol III . . . . .	112
3.6	Discussion of the results . . . . .	113
<b>4</b>	<b>Bipartite Bell inequalities with low detection efficiency</b>	<b>115</b>
4.1	Motivation and previous works . . . . .	115
4.2	Strategies to compute the detection efficiency . . . . .	117
4.3	Constructing Bell inequalities from graphs . . . . .	118
4.4	Examples of nonlocal correlations with low detection efficiency . . . . .	121
4.4.1	The graph Pauli-4320 . . . . .	122
4.4.2	The graph Pauli-36720 . . . . .	122
4.5	Independence number and quantum value of Newman graphs . . . . .	124
4.6	Incorporating noise in graph-based Bell inequalities . . . . .	129
4.7	Bell inequalities with arbitrary low detection efficiency . . . . .	131
4.8	Optimizing Bell inequalities using symmetries . . . . .	135
4.8.1	Gilbert's algorithm . . . . .	136
4.8.2	Gilbert's algorithm with symmetry . . . . .	138
4.8.3	Generating the set of symmetrized vertices . . . . .	140
4.8.4	Application to $\mathcal{P}_2(\mathbb{R})$ . . . . .	140
4.9	Conclusion and discussion . . . . .	142

<b>5</b>	<b>Certifying activation of quantum correlations with finite data</b>	<b>145</b>
5.1	Motivation . . . . .	145
5.2	Construction of a simple confidence polytope . . . . .	147
5.3	Steerability with dichotomic measurements . . . . .	150
5.4	Activation of nonlocality by local filtering . . . . .	154
5.4.1	Locality of the targeted state family . . . . .	155
5.4.2	Application to Bell-nonlocality . . . . .	156
5.4.3	Application to quantum steering . . . . .	158
5.5	Conclusion and discussion . . . . .	159
<b>6</b>	<b>Finding resourceful multipartite quantum states</b>	<b>161</b>
6.1	Motivation . . . . .	161
6.2	Concepts and notation . . . . .	163
6.3	A simple algorithm for maximizing the geometric measure . . . . .	164
6.4	The proof of monotonicity . . . . .	165
6.5	Maximally entangled states of the geometric measure . . . . .	168
6.5.1	Finding a concise representation of quantum states . . . . .	169
6.5.2	Results for qubit systems . . . . .	170
6.5.3	Higher-dimensional systems . . . . .	171
6.6	The algorithm for typical states . . . . .	174
6.6.1	The entanglement of typical states . . . . .	174
6.6.2	Performance for random starting points . . . . .	177
6.7	The algorithm for entangled subspaces . . . . .	178
6.8	Application to states with a fixed stabilizer rank . . . . .	179
6.8.1	The stabilizer rank . . . . .	179
6.8.2	The modified algorithm and results . . . . .	180
6.9	Application to states with a fixed Schmidt rank . . . . .	182
6.9.1	Schmidt rank, tensor rank and border rank . . . . .	182
6.9.2	Computing rank- $k$ approximations . . . . .	183
6.9.3	The modified algorithm and results . . . . .	184
6.10	Application to independent triangle preparable states . . . . .	187
6.10.1	Network state approximations . . . . .	187
6.10.2	The modified algorithm and results . . . . .	188
6.11	Relation to upper bounds on the geometric measure . . . . .	189
6.11.1	Bounds on the maximal entanglement . . . . .	190
6.11.2	Asymptotic scaling . . . . .	191
6.11.3	The maximal entangled state and optimal norm constants . . . . .	191
6.12	Conclusion and discussion . . . . .	192

<b>7</b>	<b>Real eigenstructure of regular simplex tensors</b>	<b>193</b>
7.1	Motivation . . . . .	193
7.2	Mathematical concepts and notation . . . . .	195
7.2.1	Tensor eigenvalues and the tensor power method . . . . .	195
7.2.2	Frames and simplex tensors . . . . .	197
7.3	Characterizing eigenpairs of regular simplex tensors . . . . .	197
7.4	Eigenstructure for local dimension $n = 2$ . . . . .	211
7.5	Eigenstructure for local dimension $n = 3$ . . . . .	214
7.6	Robustness analysis . . . . .	216
7.6.1	Local dimension $n = 2$ . . . . .	217
7.6.2	Local dimension $n = 3$ . . . . .	220
7.7	Discussion and Conclusion . . . . .	220
<b>8</b>	<b>Shadow tomography with generalized measurements</b>	<b>223</b>
8.1	Motivation . . . . .	223
8.2	Formulating shadow tomography with POVMs . . . . .	225
8.2.1	Shadow tomography with generalized measurements . . . . .	225
8.2.2	Shadows from the least-squares estimator . . . . .	225
8.2.3	Estimation of observables and sample complexities . . . . .	227
8.2.4	Relation to randomized ideal measurements . . . . .	228
8.3	Symmetries and generalized measurements . . . . .	229
8.4	Optimization of measurements . . . . .	231
8.4.1	The single qubit case . . . . .	232
8.4.2	Tensoring construction and the multi-qubit case . . . . .	233
8.5	Optimality of measurements . . . . .	234
8.6	Mitigation of measurement noise . . . . .	236
8.7	Conclusion and discussion . . . . .	238
<b>9</b>	<b>Error mitigated classical shadows</b>	<b>241</b>
9.1	Motivation . . . . .	241
9.2	Introduction and notation . . . . .	244
9.2.1	Classical shadows . . . . .	244
9.2.2	Probabilistic error cancellation . . . . .	245
9.3	Probabilistic error canceled shadows . . . . .	246
9.3.1	The protocol . . . . .	247
9.3.2	Rigorous performance guarantees . . . . .	249
9.3.3	Classical post-processing algorithms . . . . .	256
9.4	Further error mitigation techniques . . . . .	259
9.4.1	Error extrapolated shadows . . . . .	259
9.4.2	Symmetry verified shadows . . . . .	260

9.5 Applications . . . . .	261
9.5.1 Ground-state preparation . . . . .	261
9.5.2 Error mitigated estimation of entanglement entropies . . . . .	264
9.5.3 Further applications . . . . .	266
9.6 Discussion and conclusion . . . . .	266
<b>Summary and Outlook</b>	<b>269</b>
<b>Acknowledgments</b>	<b>271</b>
<b>List of publications</b>	<b>273</b>
<b>Bibliography</b>	<b>275</b>



# Introduction

Quantum information science is a synthesis of two major advances of the 20<sup>th</sup> century: quantum mechanics and information technology. Quantum mechanics describes Nature at the scale of atoms and subatomic particles and has been appreciated as one of the most accurate theories in science [1]. Despite its success, the meaning and the origin of its counterintuitive Hilbert space formalism are difficult to grasp and many of its implications challenge our perception of reality. On the other hand, information technology is concerned with storing, retrieving, manipulating, and communicating data and information [2]. It gave rise to modern computers, digital communication and devices that are ubiquitous in our daily life. Interestingly, it turns out that conceptual difficulties and apparent paradoxes in the realm of quantum mechanics translate to a new generation of protocols and devices in information technology that provide significant advantages over their classical counterparts.

One of the most striking and apparently paradoxical features, lying at the heart of quantum mechanics, is certainly entanglement. It entails that there exist multipartite quantum states that do not admit a local description in terms of its constituents, existing independently of the state of the others. Consequently, an entangled state must be seen as a single entity, regardless of the distance between the parties. That entangled states can lead to situations that are in conflict with our classical intuition was first pointed out in a seminal paper by Einstein, Podolsky and Rosen (EPR) [3]. According to their view, the quantum-mechanical description of the physical reality by means of the wave function was incomplete. However, they believed that quantum mechanics could be supplemented by additional variables "restoring to the theory causality and locality" [4], which are today known as local hidden variables (LHV). In 1964, Bell formalized the idea of EPR and showed that the statistical predictions of quantum mechanics are incompatible with the correlations allowed by *any* physical theory assuming locality, realism and free will [5]. Moreover, the resulting Bell inequalities allow for experimentally testable deviations of quantum mechanics from any theory based on LHV models. Many experiments aimed at the demonstration of quantum correlations by violating the Clauser-Horne-Shimony-Holt [6] inequality or one of its equivalent versions [7]. However, early attempts were prone to loopholes, as the experimental setups suffered from inefficient detectors [8,9] or an insufficient separation of the involved parties [10, 11]. Finally, in 2015, the first loophole-free experimental violations of a Bell inequality were announced [12, 13].

Apart from being fundamental to our understanding of Nature, the existence of nonlocal quantum correlations is indispensable for the security of certain quantum key distribution protocols. While the security of the first protocols, conceived by Wiesner [14], and Bennett and Brassard [15], relies on the no-cloning theorem [16], i.e., the impossibility of creating an independent and identical copy of an unknown quantum state, Ekert [17] realized that interventions of an eavesdropper would introduce "elements of physical reality" into the correlations shared between the involved parties [18]. Consequently, if the observed correlations violate a Bell inequality, one can bound an adversary's information by using Bell's theorem. This allows in principle for the creation of a secure key without putting strong restrictions on the eavesdropper's power. These observations triggered intensive research and can be seen as the origin of the fields of device-independent quantum key distribution [18] and self-testing [19]. Along with the rapid development of experimental tools, quantum key distribution is nowadays routinely being demonstrated in laboratories [20] and has been used to distribute secure keys over long distances [21]. However, at its current stage, losses in the transmission and imperfect detectors prevent applications outside laboratories with well controlled losses.

Another difficulty related with the formulation of quantum mechanics is that the Hilbert space of a quantum system scales exponentially with the number of parties. This poses severe computational problems in the field of condensed matter physics and quantum chemistry [22] as the simulation of quantum systems is hard for classical computing machines. However, by realizing that computation is a physical process [23], the idea of using the exponentially large Hilbert space as a resource emerged. It was first suggested by Feynman [24] and Manin [25] that controllable quantum systems could be used for simulating other quantum systems or even for solving general computational problems. These ideas set the stage for quantum simulators [26, 27] and quantum computers [28], and it is believed that those devices can solve problems that are intractable for classical computers [29]. Indeed, Shor's algorithm [30] for prime factorization provides a scheme to factor large integers in polynomial time, while no efficient classical analogue is known. However, this algorithm relies on the ability to implement deep quantum circuits, which is still out of reach for current noisy intermediate-scale quantum technology [31]. Nevertheless, it is expected that some form of early practical quantum advantage just beyond the reach of classical computing could be achieved even with noisy quantum computers [32–34].

In this thesis we are concerned with different aspects of quantum information science. More precisely, we aim at obtaining a better understanding of quantum theory from the perspective of structural and foundational questions as well as the perspective of possible applications and quantum technology.

In the first chapter, we introduce the mathematical terminology and the physical concepts that are needed in this thesis. This includes the axiomatic framework of quan-

tum theory, the notion of Bell nonlocality, quantum steering and local-friendliness. Further, we recapitulate basic concepts from quantum computing and techniques from convex optimization theory.

In the second chapter, we consider the simulability of quantum measurements with many outcomes by means of randomizing quantum measurements with fewer outcomes. We analyze how the irreducibility of a measurement can be certified in a prepare-and-measure scenario with the characteristic dimension constraint. We present a minimal scheme and a family of correlations that allow to probe the irreducibility of a three-outcome measurement on a qubit system. For this family of correlations, we analyze the robustness with respect to noise and present two numerical methods to upper and lower bound this robustness.

In the third chapter, we investigate the incompatibility of the universality of the unitary time evolution and the irreversibility of a measurement event. Motivated by a surge of revival interests in the quantum measurement problem [35–37], here we probe the assumption that the measurement event is realized relatively to *one* observer, who only partially observed the outcome of a measurement with multiple outcomes. We propose a protocol that shows that this assumption is incompatible with the universality of the unitary time evolution given that locality and free will holds.

In the fourth chapter, we introduce a family of bipartite Bell inequalities that are constructed from graphs and whose classical and quantum bounds can be expressed in terms of invariants of the underlying graph. Based on a connection to state-independent contextuality sets, we assign to each Bell inequality quantum correlations, such that their combination allows for an extremely low detection efficiency as well as high noise robustness. In addition, we discuss how these inequalities can be further optimized by means of symmetry considerations.

In the fifth chapter, we develop theoretical methods that can be used to obtain rigorous statements on the statistical significance of experiments demonstrating the activation of nonlocal quantum correlations. These contain techniques to construct a suitable confidence region in form of a polytope from the measured data and an efficient algorithm to classify the correlation class of a quantum state. We illustrate how our methods can be used to analyze the activation of quantum correlations by local filtering, specifically, for Bell-nonlocality and quantum steerability.

In the sixth chapter, we propose an iterative algorithm to find maximally resourceful quantum states of several particles for various applications and quantifiers. We present a rigorous proof of convergence, discuss in detail the case of the geometric measure, identifying physically interesting states and also deliver insights to the problem of absolutely maximally entangled states. Moreover, we demonstrate the universality of our approach by applying it to maximally entangled subspaces, the Schmidt-rank, the stabilizer rank as well as the preparability in a triangle network.

In the seventh chapter, we discuss the full real eigenstructure of regular simplex



tensors, including the robustness analysis of all normalized eigenvectors for the case of an arbitrary number of parties and local dimension 2. We show that regular simplex tensors have robust as well as non-robust eigenvectors with respect to the tensor power iteration map. Moreover, we find that the normalized eigenvectors only partly coincide with the generators from the symmetric tensor decomposition.

Finally, in chapter eight and nine, we discuss scalable methods that allow for simultaneously predicting many expectation values of a multi-qubit system. We present a variant of shadow tomography which is based on generalized measurements and provide a detailed study of the implication of symmetries. In addition, we demonstrate how the measurement directions can be optimized towards a set of targeted observables and how measurement errors can be mitigated. Concerning errors in the preparation phase, we generalize error mitigation techniques, such that they can be applied directly to classical shadows. We discuss the technique of probabilistic error cancellation in detail and provide rigorous theoretical sample complexities.

# 1 Preliminaries

In this Chapter we give a concise introduction into the basic notions of quantum information theory that are needed throughout this thesis. First, we will recapitulate the mathematical framework of quantum theory in Section 1.1. We proceed in Section 1.2 by introducing different types of bipartite correlations such as entanglement, quantum steering and Bell nonlocality. In Section 1.3 we outline concepts from graph theory, the notion of graph states as well as basics of the field of quantum computing. In addition, we mention necessary computational tools from optimization theory in Section 1.4.

Of course, the aim of this Chapter is not to give a comprehensive exposition of quantum theory. It is intended to provide a sufficient amount of basic knowledge, such that the subsequent chapters can be easily understood without the need of consulting other literature. The contents presented here can be found in any textbook about quantum mechanics with a focus on quantum information theory, see for instance the books by Peres [1], Holevo [38], Heinosaari and Ziman [39], or Nielsen and Chuang [29].

## 1.1 Mathematical framework of quantum theory

The framework of quantum theory offers the possibility to calculate probabilities for the outcomes of measurements, given that a system was prepared in a particular state and subjected to a particular evolution. In Section 1.1.1 we will discuss the notion of a quantum state, which comprises all the information that is available about the system under investigation. We then proceed in Section 1.1.2 by introducing the concept of quantum measurements by which means we extract classical information about the quantum system. Finally, we discuss quantum channels in Section 1.1.2 describing the dynamics the system is subjected to.

### 1.1.1 Quantum states

Quantum theory starts by associating a Hilbert space  $\mathcal{H}$  to a physical system. A Hilbert space is a complex inner product space that is complete with respect to the metric induced by this inner product. Here complex inner product means linearity in the first argument, conjugate symmetry under exchange of the arguments as well as positive definiteness, i.e., for any  $\psi \in \mathcal{H}$  one has  $\langle \psi | \psi \rangle > 0$ , unless  $\psi = 0$  in which case the inner product equals zero. An inner product on a space  $\mathcal{H}$  gives rise to a hierarchy of

structures. By virtue of  $\|\psi\| = \sqrt{\langle\psi|\psi\rangle}$ , the inner product yields a norm on  $\mathcal{H}$ , which itself induces a metric via  $d(\psi, \phi) = \|\psi - \phi\|$ .

The associated Hilbert space comprises all different configurations in which the system can possibly be. To each configuration, there exists a vector of unit length  $|\psi\rangle \in \mathcal{H}$ , called *ket*, describing the state of the system. One assumes that  $|\psi\rangle$  contains all information about the system that one can have, that is, knowing  $|\psi\rangle$  means to have a full description of the actual state of the system. For this reason, a state vector  $|\psi\rangle$  is also called a pure state.

More precisely, a quantum system having  $1 \leq d < \infty$  degrees of freedom (also called *d-level system* or *qudit system*) is associated with the Hilbert space  $\mathbb{C}^d$ . Any such Hilbert space admits an orthonormal basis  $\{|\alpha\rangle\}$  for  $0 \leq \alpha \leq d-1$  and any element of  $\mathcal{H}$  can be uniquely expanded with respect to such a basis, that is, for  $|\psi\rangle \in \mathbb{C}^d$  there are coefficients  $c_\alpha \in \mathbb{C}$ , such that

$$|\psi\rangle = \sum_{\alpha=0}^{d-1} c_\alpha |\alpha\rangle. \quad (1.1)$$

The number of elements that appear in a basis is called the dimension of the system. From the normalization we directly obtain  $1 = \langle\psi|\psi\rangle = \sum_{\alpha=0}^{d-1} |c_\alpha|^2$ . Although there is no distinguished basis for a Hilbert space, it is often convenient to work in the so-called *computational basis*. The computational basis ket  $|j\rangle$  has entry 1 at the  $j$ th position and 0 elsewhere. Similar, any state  $|\psi\rangle$  can be decomposed with respect to that basis

$$|\psi\rangle = \sum_{j=0}^{d-1} c_j |j\rangle, \quad (1.2)$$

where  $c_j \in \mathbb{C}$  for  $0 \leq j \leq d-1$ .

So far, we have considered the description of a single physical system. However, it is often the case that the system under investigation is of composed form, that is, the whole system consists of multiple individual constituents.

Given  $n$  particles, each associated with a Hilbert space  $\mathbb{C}^{d_j}$ , the joint system is described by the Hilbert space

$$\mathcal{H} = \mathbb{C}^{d_1} \otimes \dots \otimes \mathbb{C}^{d_n}. \quad (1.3)$$

Because  $\mathcal{H}$  in Eq. (1.3) shares an additional tensor structure, it is also called a tensor product space. Assuming that the independent pure quantum states of the individual systems are  $\{|\psi_j\rangle\}_{j=1}^n$ , the state of the joint system is given by  $|\psi\rangle = \otimes_{j=1}^n |\psi_j\rangle \in \mathbb{C}^{d_1} \otimes \dots \otimes \mathbb{C}^{d_n}$ . As a tensor product space is also a vector space, any multipartite quantum state can be expanded in any orthonormal basis. A common choice is to expand the state with respect to the tensor product of the local computational bases. It is often convenient to introduce the following short hand notation

$$|j_1 j_2 \dots j_n\rangle := |j_1\rangle |j_2\rangle \dots |j_n\rangle := |j_1\rangle \otimes |j_2\rangle \otimes \dots \otimes |j_n\rangle. \quad (1.4)$$

Then, any quantum state  $|\psi\rangle$  can be written as

$$|\psi\rangle = \sum_{j_1=0}^{d_1} \cdots \sum_{j_n=0}^{d_n} \psi_{j_1 \cdots j_n} |j_1 \cdots j_n\rangle, \quad (1.5)$$

where  $\psi_{j_1 \cdots j_n} \in \mathbb{C}$  is the coefficient tensor of  $|\psi\rangle$ . However, it can also be useful to decompose a multipartite state into a basis which cannot be written as tensor products of the local bases. For example, consider the Hilbert space  $\mathcal{H} = \mathbb{C}^2 \otimes \mathbb{C}^2$ . The so-called Bell states

$$\begin{aligned} |\phi^+\rangle &= \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle), \\ |\phi^-\rangle &= \frac{1}{\sqrt{2}}(|00\rangle - |11\rangle), \\ |\psi^+\rangle &= \frac{1}{\sqrt{2}}(|01\rangle + |10\rangle), \\ |\psi^-\rangle &= \frac{1}{\sqrt{2}}(|01\rangle - |10\rangle), \end{aligned} \quad (1.6)$$

form an orthonormal basis for the space  $\mathbb{C}^2 \otimes \mathbb{C}^2$ , while not being tensor products of local bases. Furthermore, for reasons that will become clear later (cf. Section 1.2.2) we call the bipartite state

$$|\Phi_d\rangle := \frac{1}{\sqrt{d}} \sum_{j=0}^{d-1} |j\rangle|j\rangle \quad (1.7)$$

the maximally entangled quantum state of the system  $\mathbb{C}^d \otimes \mathbb{C}^d$ .

If one wants to prepare a particular quantum state in the laboratory, one typically has to deal with imperfections in the experimental settings. This often has the consequence that the actually prepared state is not the targeted but a different state. As the particular misalignment of the setup may vary for each preparation, one obtains a statistical mixture of pure quantum states. Such a probability distribution over the set of pure quantum states in  $\mathcal{H}$  is called a mixed state and is represented by a self-adjoint positive semidefinite operator in the space  $\mathcal{B}(\mathcal{H})$  of bounded operators on  $\mathcal{H}$  being of unit trace. More formally, a density operator is given by

$$\varrho = \sum_j p_j |\psi_j\rangle\langle\psi_j|, \quad (1.8)$$

where  $|\psi_j\rangle \in \mathcal{H}$  are pure quantum states. In order to ensure positive semidefiniteness of  $\varrho$ , one requires that  $p_j \geq 0$  and the unit trace property demands that  $\sum_j p_j = 1$ . Clearly, any pure quantum state corresponds to a density operator  $\varrho$  and is given by a rank one projector  $\varrho = |\psi\rangle\langle\psi|$ . Notice that this correspondence is not unique as the associated density operator  $\varrho$  is invariant under global phases of the pure state  $|\psi\rangle$ . However, we will see that quantum theory does not allow to distinguish states

that only differ by a global phase. Consequently, identifying those states is a natural procedure. We denote by  $\mathcal{S}(\mathcal{H}) \subset \mathcal{B}(\mathcal{H})$  the set of all density operators. The state  $\mathcal{S}(\mathbb{C}^d) \ni \varrho = \frac{1}{d}\mathbb{1}$  is called the maximally mixed state. Whether a quantum state  $\varrho$  is pure or mixed can be directly decided by computing the so-called purity of the state, which is defined as  $\mathcal{P}(\varrho) := \text{Tr}[\varrho^2]$ . A state  $\varrho$  is pure if and only if  $\mathcal{P}(\varrho) = 1$  and  $\mathcal{P}(\varrho) = \frac{1}{d}$  for  $\varrho \in \mathcal{B}(\mathbb{C}^d)$  if and only if  $\varrho$  is maximally mixed. Because  $\varrho$  is self-adjoint, the spectral theorem for self-adjoint operators guarantees that  $\varrho$  can be diagonalized with non negative real eigenvalues  $\{\lambda_j\}_j$ . Further, one has  $1 = \text{Tr}[\varrho] = \sum_j \lambda_j$ . This also allows for an interpretation of mixed states in the other direction. Any mixed state can be seen as a statistical mixture of pure states. Again, this correspondence does not need to be unique. For instance, consider two sources  $S_1, S_2$  with  $S_1$  producing the states  $|0\rangle$  and  $|1\rangle$  each with probability  $\frac{1}{2}$  and  $S_2$  producing  $|x^+\rangle$  and  $|x^-\rangle$  each with probability  $\frac{1}{2}$ , where  $|x^+\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$  and  $|x^-\rangle = \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)$ . Then  $\varrho_{S_1} = \varrho_{S_2}$ , while the particular physical realization of the ensemble was different.

Given a physical system composed of two parts, lets say  $AB$ , one is often interested in the effective quantum state that one party, lets say  $A$ , holds. This forgetting or marginalization operation is mathematically described by the partial trace. If the state of the joint system is an element of  $\mathcal{S}(\mathcal{H}_A \otimes \mathcal{H}_B)$ , the partial trace over system  $A$  is a linear mapping

$$\text{Tr}_A : \mathcal{S}(\mathcal{H}_A \otimes \mathcal{H}_B) \rightarrow \mathcal{S}(\mathcal{H}_B), \quad (1.9)$$

which satisfies the relation

$$\text{Tr}[\text{Tr}_A[\varrho]W] = \text{Tr}[\varrho(\mathbb{1} \otimes W)] \quad (1.10)$$

for all  $\varrho \in \mathcal{S}(\mathcal{H}_A \otimes \mathcal{H}_B)$  and all operators  $W \in \mathcal{B}(\mathcal{H}_B)$ . The partial trace  $\text{Tr}_B$  over subsystem  $B$  is defined in a similar way and the generalization to the multipartite case is straightforward. It can be shown that a map defined by Eq. (1.9) and Eq. (1.10) exists and is unique [39]. Similarly to the trace of an operator, also the partial trace can be obtained by summing over an orthonormal basis. Indeed, if  $\varrho = \sum_{ijkl} \varrho_{kl}^{ij} |ij\rangle\langle kl|$  one has

$$\varrho_B = \text{Tr}_A[\varrho] = \sum_{mn} \sum_{ijkl} \varrho_{kl}^{ij} \langle m|i\rangle\langle k|n\rangle |j\rangle\langle l| = \sum_{ijkl} \varrho_{kl}^{ij} |j\rangle\langle l|. \quad (1.11)$$

The partial trace is particularly easy to compute if the multipartite state  $|\psi\rangle$  is pure, i.e., we have  $|\psi\rangle = \otimes_{j=1}^n |\psi_j\rangle$ . In this case, the corresponding density operator is given by  $\varrho = \otimes_{j=1}^n |\psi_j\rangle\langle\psi_j|$  and the partial trace over subsystems  $S \subset \{1, \dots, n\}$  is given by

$$\text{Tr}_S[|\psi\rangle\langle\psi|] = \otimes_{j \in S^c} |\psi_j\rangle\langle\psi_j|. \quad (1.12)$$

However, it should be noticed that in general taking the partial trace of a given state does not preserve its purity. For instance, the Bell states defined in Eq. (1.6) are bipartite

pure states, while the one-party reductions are all *maximally mixed*. We will discuss this phenomenon in more detail in Section 1.2.2.

By construction, a mixed state can be seen as a statistical mixture of pure states. As we will see now, there exists at least partially also a reverse statement, which is known as state purification. The idea is that any mixed quantum state  $\rho \in \mathcal{B}(\mathcal{H})$  can be seen as a pure state in a larger Hilbert space  $\mathcal{Q}$ . Assume that  $\rho$  has a decomposition of the form  $\rho = \sum_j p_j |\psi_j\rangle\langle\psi_j|$ . Now consider the state

$$\mathcal{Q} = \mathcal{H} \otimes \mathcal{K} \ni |\eta\rangle := \sum_j \sqrt{p_j} |\psi_j\rangle \otimes |j\rangle, \quad (1.13)$$

where  $\{|j\rangle\}_j$  is some orthonormal basis for the auxiliary Hilbert space  $\mathcal{K}$ . The mixed state  $\rho$  can then be recovered from  $|\eta\rangle$  by tracing out the system  $\mathcal{K}$ , that is,  $\rho = \text{Tr}_{\mathcal{K}}[|\eta\rangle\langle\eta|]$ . The state  $|\eta\rangle$  is called a purification of  $\rho$ . Obviously, as the space  $\mathcal{K}$  as well as the orthonormal basis  $\{|j\rangle\}_j$  can be chosen arbitrarily, the purification is not unique.

### The Bloch sphere and operator bases

As already seen, we can expand every vector in a vector space with respect to some basis. The set of bounded operators acting on a Hilbert space  $\mathcal{H}$ , denoted by  $\mathcal{B}(\mathcal{H})$ , is also a vector space and consequently a similar decomposition exists for operators. Within this space, the set of density matrices arises as the intersection of the convex cone of positive semidefinite matrices and the subspace of operators of unit trace. The subset of  $\mathcal{B}(\mathcal{H})$  containing only self-adjoint operators is denoted by  $\mathcal{B}_H(\mathcal{H})$  and forms an  $\mathbb{R}$ -linear subspace. For the case  $\mathcal{B}_H(\mathbb{C}^2)$ , that is, for qubits, there is a very popular and common choice of basis. The Pauli operators are given by

$$\sigma_0 = \mathbb{1} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \quad (1.14)$$

The operators  $\sigma_0, \sigma_1, \sigma_2, \sigma_3$  are often denoted by  $I, X, Y, Z$ . The Pauli operators fulfill a list of important properties. First, they are self-adjoint  $\sigma_j = \sigma_j^\dagger$ , they are traceless  $\text{Tr}[\sigma_j] = 0$  for  $1 \leq j \leq 3$ , they are self inverse  $\sigma_j^2 = \sigma_0$ , they fulfill the  $\text{SU}(2)$  commutation relation

$$[\sigma_i, \sigma_j] = 2i \sum_k \epsilon_{ijk} \sigma_k, \quad (1.15)$$

as well as the Dirac algebra

$$\{\sigma_i, \sigma_j\} = 2\delta_{ij}\sigma_0. \quad (1.16)$$

This already implies that they are orthogonal with respect to the Hilbert-Schmidt inner product given by  $\langle A, B \rangle = \text{Tr}[A^\dagger B]$ ,

$$\text{Tr}[\sigma_j^\dagger \sigma_k] = \text{Tr}[\sigma_j \sigma_k] = 2\delta_{jk}. \quad (1.17)$$

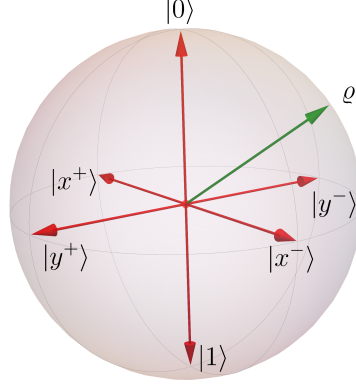


Figure 1.1: Illustration of the Bloch sphere. The red vectors denote the eigenstates of the Pauli operators, while the green vector corresponds to an arbitrary pure qubit state lying at the surface of the ball.

In particular, Eq. (1.16) together with  $\sigma_j^2 = \sigma_0$  means that Pauli operators either commute or anticommute. Sometimes it is more convenient to consider a normalized variant of the Pauli operators, that is,  $P_j := (1/\sqrt{2})\sigma_j$  such that  $\text{Tr}[P_j P_k] = \delta_{jk}$ . In other words, the set  $\{P_0, P_1, P_2, P_3\}$  is an *orthonormal* basis for  $\mathcal{B}_H(\mathbb{C}^2)$  with respect to the Hilbert-Schmidt inner product. Any qubit state can be expressed in the Pauli basis as

$$\rho = \frac{1}{2}(\sigma_0 + r_1\sigma_1 + r_2\sigma_2 + r_3\sigma_3), \quad (1.18)$$

where  $r_j \in \mathbb{R}$  for  $1 \leq j \leq 3$ . In fact, this allows us to associate to each qubit density operator  $\rho$  faithfully a point in  $\mathbb{R}^3$ , i.e.,  $\vec{r} = (r_1, r_2, r_3)$ , which is called Bloch vector. Indeed, from the positive semidefiniteness of  $\rho$  and the normalization  $\text{Tr}[\rho] = 1$  it follows that  $\|\vec{r}\|_2 \leq 1$ . Further, one directly obtains that the purity  $\mathcal{P}(\rho)$  relates to the length of the Bloch vector  $\vec{r}$  via  $\mathcal{P}(\rho) = \frac{1}{2}(1 + \|\vec{r}\|_2^2)$ . Consequently, there is a bijective correspondence between qubit states and vectors in the unit ball of  $\mathbb{R}^3$ , with the maximally mixed state at the origin and the pure states at the surface, see also Fig. 1.1. Due to the orthogonality of the Pauli operators, one can directly compute the coefficients  $r_j$  for a given density operator  $\rho$  via  $r_j = \text{Tr}[\rho\sigma_j]$ .

As we have already seen for pure quantum states, it is often convenient to express the global state as a combination of tensor products of the local bases. This works similar for multipartite mixed states and is simply given by tensor products of Pauli operators. In this way, one obtains the  $n$ -qubit basis of self-adjoint operators acting on the tensor space  $(\mathbb{C}^2)^{\otimes n}$  as

$$\mathcal{P}_n := \{\sigma_\alpha \mid \sigma_\alpha = \sigma_{\alpha_1} \otimes \cdots \otimes \sigma_{\alpha_n}\}. \quad (1.19)$$

By using the orthogonality of single qubit Pauli operators and the fact that the trace is multiplicative with respect to the tensor product, one directly obtains that  $\text{Tr}[\sigma_\alpha \sigma_\beta] =$

$2^n \delta_{\alpha_1 \beta_1} \cdots \delta_{\alpha_n \beta_n}$ . For a Pauli operator only acting on subsystem  $j$  we also write  $X_j$  and similar for the other Pauli operators. For instance, for a four-qubit system one has  $X_2 Y_3 = I X Y I = I \otimes X \otimes Y \otimes I$ . A density operator  $\varrho$  of an  $n$ -qubit system can be expanded with respect to the Pauli basis as

$$\varrho = \frac{1}{2^n} \sum_{\sigma_\alpha \in \mathcal{P}_n} \text{Tr}[\varrho \sigma_\alpha] \sigma_\alpha. \quad (1.20)$$

The support of a Pauli string  $\sigma_\alpha \in \mathcal{P}_n$  is defined as the set of those positions on which the operator  $\sigma_\alpha$  acts nontrivial, that is,  $\text{supp}(\sigma_\alpha) := \{j | \sigma_{\alpha_j} \neq I\}$ . The number of positions on which  $\sigma_\alpha$  acts nontrivial is called the weight,  $\text{wt}(\sigma_\alpha) = |\text{supp}(\sigma_\alpha)|$ .

Although qubit and multi-qubit systems are very important in many applications also systems where the local dimension of the constituents is larger play an important role. In order to obtain a similar description of mixed states as in Eq. (1.20), one needs Bloch representations for higher dimensions. Here the name Bloch representation refers to any choice of orthogonal basis of the operator space. However, in practice one typically uses basis elements that also possess an additional algebraic structure. One option for generalizing the Pauli matrices to higher spins is by using the observation that  $\sigma_1, \sigma_2, \sigma_3$  are generators of a representation of the Lie algebra  $\mathfrak{su}(2)$  associated with the Lie group  $\text{SU}(2)$ . The idea is to take as basis elements the  $d^2 - 1$  matrices generating the Lie algebra associated to the special unitary group  $\text{SU}(d)$ . There is an easy prescription how such matrices, also called generalized Gell-Mann matrices, can be constructed. It turns out that the set of generators decomposes into the subclass of symmetric

$$\lambda_s^{jk} = |j\rangle\langle k| + |k\rangle\langle j|, \quad 0 \leq j < k \leq d-1, \quad (1.21)$$

of antisymmetric

$$\lambda_{\text{as}}^{jk} = -i|j\rangle\langle k| + i|k\rangle\langle j|, \quad 0 \leq j < k \leq d-1, \quad (1.22)$$

and of diagonal matrices

$$\lambda^\ell = \sqrt{\frac{2}{\ell(\ell+1)}} \left( \sum_{j=1}^{\ell} |j\rangle\langle j| - \ell|\ell+1\rangle\langle\ell+1| \right), \quad 0 \leq j < k \leq d-2. \quad (1.23)$$

It follows directly from the definition, that the generalized Gell-Mann operators are self-adjoint and traceless. Further, they satisfy the orthogonality relations

$$\text{Tr}[\lambda_s^{jk} \lambda_s^{\alpha\beta}] = \text{Tr}[\lambda_{\text{as}}^{jk} \lambda_{\text{as}}^{\alpha\beta}] = 2\delta_{j\alpha} \delta_{k\beta}, \quad \text{Tr}[\lambda^j \lambda^k] = 2\delta_{jk}, \quad (1.24)$$

while all other inner products vanish. Following the construction of the multi-qubit Bloch representation, one can also obtain a similar variant for systems of higher dimension. Let  $\mathcal{G}_d$  denote the generalized Gell-Mann operators of dimension  $d$ . Then, a



tensor product basis can be obtained via

$$\Lambda_\alpha = \Lambda_{\alpha_1} \otimes \cdots \otimes \Lambda_{\alpha_n} \quad (1.25)$$

for  $\Lambda_{\alpha_j} \in \mathcal{G}_d$ . Consequently, any many-particle density operator  $\rho$  can then be expanded as

$$\rho = \frac{1}{2^n} \sum_{\alpha} \text{Tr}[\Lambda_\alpha \rho] \Lambda_\alpha. \quad (1.26)$$

A different representation relies on the Heisenberg-Weyl or displacement basis. Albeit these operators share some convenient properties, e.g., they are unitary, they are not self-adjoint. Define the phase operator  $X$  and the shift operator  $Z$  via their action on a basis element

$$X|j\rangle = |(j+1) \bmod d\rangle, \quad Z|j\rangle = \omega^j |j\rangle, \quad (1.27)$$

where the complex phase  $\omega$  is given by  $\omega = e^{\frac{2\pi i}{d}}$ . The operators  $X$  and  $Z$  do in general not commute and obey the relation

$$Z^l X^m = \omega^{lm} X^m Z^l. \quad (1.28)$$

The unitaries corresponding to discrete phase-space displacements for  $d$ -level systems are defined as

$$\mathcal{D}(l, m) = \bar{\omega}^{lm/2} Z^l X^m. \quad (1.29)$$

Displacement operators have many convenient properties which make them particularly useful as a basis set. First, they are a complete set in the sense that they are a basis for  $\mathcal{B}(\mathbb{C}^d)$ . Second, they satisfy the following orthogonality condition

$$\text{Tr}[\mathcal{D}(j, k) \mathcal{D}(l, m)^\dagger] = d \delta_{j,l} \delta_{k,m}. \quad (1.30)$$

Therefore, any quantum state  $\rho \in \mathcal{S}(\mathbb{C}^d)$  can be decomposed into

$$\rho = \frac{1}{d} \sum_{j,k=0}^{d-1} \text{Tr}[\rho \mathcal{D}(j, k)] \mathcal{D}(j, k)^\dagger. \quad (1.31)$$

It should be noticed that the Bloch vector components with respect to the Heisenberg-Weyl basis  $r_{jk} = \text{Tr}[\rho \mathcal{D}(j, k)]$  are generally complex, as the displacement operators are not hermitian. Consequently, in order to fully characterize the density operator  $\rho$ , one has to determine  $d^2 - 1$  complex parameters.

### 1.1.2 Quantum measurements

Up to this point, we have only discussed how the state of a quantum system can be described by Hilbert spaces and density operators. However, no concept introduced so

far has been related to a quantity that could actually be *observed*. Quantum measurements are the means by which we obtain information about a physical system. They allow us to make predictions about observable quantities. However, in difference to classical mechanics, the quantum-theoretical formalism only allows us in general to predict outcome probabilities, but not the particular outcome itself. We will now introduce the notion of observables and projective measurements, which are also called projector-valued measures.

Given a quantum state  $\varrho \in \mathcal{B}(\mathcal{H})$ , an observable is described by a self-adjoint operator  $A \in \mathcal{B}(\mathcal{H})$ . In a run of an experiment, the possible observable outcomes are labeled by the eigenvalues  $\{a_j\}_j$  of  $A$ . The probability to observe a certain outcome  $a_j$  is given by the Born rule

$$p_j := \text{Prob}[a_j|\varrho] := \text{Tr}[\Pi_{a_j}\varrho], \quad (1.32)$$

where  $\Pi_{a_j}$  is the projector corresponding to the eigenspace of  $A$  associated to eigenvalue  $a_j$ . More precisely, as  $A$  is a self-adjoint operator, the spectral theorem allows for a decomposition of the form  $A = \sum_j a_j \Pi_{a_j}$ , where the projector  $\Pi_{a_j}$  takes the case into account, where the eigenspace can be degenerated, i.e., of dimension larger than one. Because the outcomes of a measurement are related to projectors  $\Pi_{a_j}$ , an observable is also called a projective measurement. Further, the expectation value of  $A$  with respect to the state  $\varrho$  is

$$\langle A \rangle_\varrho = \sum_j a_j \text{Prob}[a_j|\varrho] = \sum_j a_j \text{Tr}[\varrho \Pi_{a_j}] = \text{Tr}\left[\varrho \sum_j a_j \Pi_{a_j}\right] = \text{Tr}[\varrho A]. \quad (1.33)$$

In a similar manner as mixed states generalize the concept of pure states, positive-operator valued measures (POVMs) generalize projective measurements. POVMs do not only allow for incorporating the effect of experimental noise into the description of a measurement, e.g., if the measurement apparatus is not coupling to the system perfectly, but can also yield advantages in communication-theoretic tasks.

From an abstract perspective, any quantum measurement yields at its end a certain outcome  $\omega \in \Omega$ , where  $\Omega$  is the outcome space, i.e., the collection of all possible outcomes that can be observed. However, it is natural to be not only interested in the single outcomes but also in certain subsets of them. Therefore, one introduces on the outcome space  $\Omega$  an additional structure, called a  $\sigma$ -algebra, containing all the sets one is typically interested in. Formally, a  $\sigma$ -algebra on  $\Omega$  is a nonempty collection  $\mathcal{A}$  of subsets of  $\Omega$  that is closed under complement, countable unions, and countable intersections. The pair  $(\Omega, \mathcal{A})$  is called a measurable space. If  $\Omega$  is finite, it is convenient to choose the power set of  $\Omega$  as  $\sigma$ -algebra  $\mathcal{A}$ , hence  $\mathcal{A}$  contains  $2^{|\Omega|}$  elements<sup>1</sup>. A probability measure on a measurable space  $(\Omega, \mathcal{A})$  is a mapping  $\mu : \mathcal{A} \rightarrow [0, 1]$  that satisfies

<sup>1</sup>For the case of  $\Omega$  at most countable, one can show that the condition that  $\mathcal{A}$  contains all singleton sets  $\{\omega\}$  for  $\omega \in \Omega$  already implies that  $\mathcal{A}$  is given by the power set.

the conditions  $\mu(\emptyset) = 0$ ,  $\mu(\Omega) = 1$  and for any countable collections  $\{X_j\}_j \subset \mathcal{A}$  of disjoint sets of  $\mathcal{A}$  one has  $\mu(\cup_j X_j) = \sum_j \mu(X_j)$ . For a given event  $X \in \mathcal{A}$ , the number  $\mu(X)$  is then interpreted as the probability that the event  $X$  occurs. In quantum theory, given a Hilbert space  $\mathcal{H}$ , one considers the associated effect space which contains all self-adjoint positive semidefinite operators, whose spectrum is bounded by 1, that is,

$$\mathcal{E}(\mathcal{H}) := \{E \in \mathcal{B}(\mathcal{H}) \mid 0 \leq E \leq \mathbb{1}\}. \quad (1.34)$$

An element  $E \in \mathcal{E}(\mathcal{H})$  is called effect. The effect space offers a geometrical interpretation as a double cone, that is, it can be seen as the intersection of two cones.

A positive operator-valued measure (POVM) is a mapping  $P : (\Omega, \mathcal{A}) \rightarrow \mathcal{E}(\mathcal{H})$  such that the conditions

$$P(\emptyset) = 0 \in \mathcal{B}(\mathcal{H}), \quad (1.35)$$

$$P(\Omega) = \mathbb{1}, \quad (1.36)$$

$$P(\cup_j X_j) = \sum_j P(X_j) \quad (1.37)$$

hold, where Eq. (1.37) is with respect to all countable collections  $\{X_j\}_j \subset \mathcal{A}$  of disjoint sets of  $\mathcal{A}$ . By construction, a given POVM  $P$  induces a family of probability measures  $\mu_\varrho$  on the measurable space  $(\Omega, \mathcal{A})$  by virtue of the Born rule, that is,

$$\mathcal{A} \ni X \mapsto \mu_\varrho(X) := \text{Tr}[\varrho P(X)]. \quad (1.38)$$

In other words, a POVM  $P$  together with the map  $X \mapsto \text{Tr}[\varrho P(X)]$  yields a probability measure for every state  $\varrho \in \mathcal{B}(\mathcal{H})$ . For the case where  $\Omega$  is finite,  $|\Omega| < \infty$ , it is convenient to identify a POVM with its image in the effect space. More precisely, one regards a POVM as a collection of operators  $\{E_\omega\}_{\omega \in \Omega}$ . Some comments are in order. First, the notion of a POVM contains projective measurements as a special case. Indeed, if one chooses the effects to be projectors, i.e.,  $E_\omega = \Pi_\omega$ , one exactly recovers a projective measurement. Second, POVMs do not necessarily arise from a countable outcome space  $\Omega$ . A common example of a continuous outcome POVM is given by spin measurements in random directions [39]. Consider the projection operators  $\sigma_\pm(\vec{n}) := \frac{1}{2}(\mathbb{1} \pm \vec{\sigma} \vec{n})$  for a spin direction  $\vec{n} \in \mathbb{R}^3$  with  $\|\vec{n}\|_2 = 1$ . If the input state  $\varrho \in \mathcal{B}(\mathbb{C}^2)$  has the Bloch decomposition  $\varrho = \frac{1}{2}(\mathbb{1} + \vec{r} \vec{\sigma})$ , the probability to observe outcome  $\pm 1$  given that a spin measurement in direction  $\vec{n}$  was implemented is

$$\text{Tr}[\sigma_+(\vec{n})\varrho] = \text{Tr}[\sigma_-(\vec{n})\varrho] = \frac{1}{2}(1 + \vec{r} \vec{n}). \quad (1.39)$$

As already pointed out, a POVM together with a quantum state  $\varrho = \frac{1}{2}(\mathbb{1} + \vec{r} \vec{\sigma})$  induces a probability measure on the underlying measurable space  $(\Omega, \mathcal{A})$ . Here,  $\Omega$  is given by the 2-sphere  $\mathbb{S}^2 \subset \mathbb{R}^3$ . Further, the associated  $\sigma$ -algebra  $\mathcal{A}$  is simply given by the subspace  $\sigma$ -algebra, i.e.,  $\mathcal{A} = \{A \cap \mathbb{S}^2 \mid A \in \mathfrak{B}(\mathbb{R}^3)\}$ , where  $\mathfrak{B}$  denotes the corresponding

Borel algebra. Therefore, the corresponding probability measure on  $S^2$  is

$$\mathfrak{B}(S^2) \ni X \mapsto \mu_\varrho(X) = \frac{1}{4\pi} \int_X (1 + \vec{r} \cdot \vec{n}) \, d\vec{n}, \quad (1.40)$$

where  $\frac{1}{4\pi} d\vec{n}$  refers to the surface element in spherical coordinates. Consequently, the associated POVM is given by

$$\mathfrak{B}(S^2) \ni X \mapsto P(X) = \frac{1}{4\pi} \int_X (\mathbb{1} + \vec{n} \cdot \vec{\sigma}) \, d\vec{n}. \quad (1.41)$$

We have already seen that any mixed quantum state can be regarded as a pure state in an enlarged Hilbert space. A similar statement is true for POVMs and is known under the name Naimark dilation [39, 40]. This theorem offers a prescription how POVMs can be implemented in an experiment. We will state it here for the case of a finite outcome space  $\Omega$ . Let  $\{E_\omega\}_{\omega \in \Omega} \subset \mathcal{B}(\mathcal{H})$  be a POVM. Then, there exists a Hilbert space  $\mathcal{K}$ , a bounded operator  $V : \mathcal{K} \rightarrow \mathcal{H}$  and a projector-valued measure  $\{\Pi_\omega\}_{\omega \in \Omega}$  on  $\mathcal{B}(\mathcal{K})$  such that

$$E_\omega = V \Pi_\omega V^\dagger. \quad (1.42)$$

Therefore, Naimark's theorem allows for implementing a POVM on a system  $\mathcal{H}$  by first coupling to another quantum system and then performing a projective measurement  $\{\Pi_\omega\}_{\omega \in \Omega}$  on the enlarged system  $\mathcal{K}$ .

### State-update rules

A further important question is how the measurement process affects the state of the system, that is, what the state after the measurement is. This state assignment depends on what kind of measurement has been conducted and also on the particular update rule. Here we will only comment on the case of projective measurements and leave the discussion for POVMs to Section 1.1.3.

In a seminal book [41], von Neumann formulated a rule for how to obtain the state of an ensemble of physical systems after a projective measurement  $A$ . If  $A$  is nondegenerate, i.e.,  $A = \{|a_j\rangle\langle a_j|\}$  with  $\langle a_i | a_j \rangle = \delta_{ij}$ , the measurement of  $A$  with respect to a state  $\varrho$  leads to a collapse of the state into one of the eigenstates  $|a_j\rangle$ , depending on the particular observed outcome  $a_j$ . Notice that any kind of superposition of the state is destroyed and the measurement leads to full decoherence in the entire eigenbasis of the observable. However, if the observable is degenerated, there are in fact two different rules, as well as certain hybrids thereof, for assigning a post-measured state to the system.

According to von Neumann, the measurement device refines the observable  $A$  into another commuting observable  $\tilde{A}$ , which is the actual implemented observable, having a nondegenerate spectrum. The measurement process of  $A$  collapses the state into an

eigenstate of  $\tilde{A}$ , removing any coherence of the original state as the degeneracy has been lifted. It often happens that  $\tilde{A}$  arises from consecutive measurements. In this case,  $A$  is a function of  $\tilde{A}$ , that is,  $A = f(\tilde{A})$  for some function  $f$  and from a measurement result  $\tilde{a}$  of  $\tilde{A}$  for an individual run of the experiment one obtains the corresponding result  $a = f(\tilde{a})$  for  $A$ . There is a second way how post-measurement states can be assigned, which is due to Lüders [42]. Here, a system existing in a superposition of degenerate eigenstates is unaffected by the measurement, such that coherence within these subspaces is preserved. Given an observable  $A = \{\Pi_{a_j}\}$ , the state  $\rho$  transforms to

$$\rho \mapsto \frac{\Pi_{a_j} \rho \Pi_{a_j}}{\text{Tr}[\Pi_{a_j} \rho]}, \quad (1.43)$$

given that the measurement outcome  $a_j$  was observed. In principle, one can also think about hybrid models of the von Neumann and Lüders rule, where one only lifts the degeneracy of an observable  $A$  partially. The question is now which rule is the appropriate choice. As there is no a priori choice for such an update rule that could be derived from the postulates of quantum theory introduced so far, this question should be answered with the help of experiments and not only theoretical considerations. The validity of Lüders's rule has recently been observed experimentally [43].

### State discrimination tasks

In general, state discrimination tasks deal with the question how well two or multiple quantum states can be distinguished in an experiment. However, in difference to quantum state estimation procedures like quantum state tomography or parameter estimation, here one deals with a single copy of the input state. More formally, a source prepares one quantum state among a set of  $n$  possible quantum states  $\rho_1, \dots, \rho_n$  which are known in advance.

The particular preparation in each run follows a probability distribution, i.e., the state  $\rho_k$  occurs with probability  $p_k$  and  $\sum_k p_k = 1$ . The task of state discrimination is to design a (generalized) measurement which can discriminate between these states. More precisely, one seeks a measurement  $N = (N_1, \dots, N_m)$  with  $m \geq n$ , such that the observation of the outcome  $k$  allows for an identification of the state  $\rho_k$ .

In general, there are two approaches to solve the problem. In minimum error discrimination one is forced to declare one of the states to be the prepared one, but one allows for a wrong identification. This means that the POVM  $N$  has exactly as many outcomes as state preparations. Among those POVMs, one aims to find an instance which maximizes the success probability

$$S_{\text{med}} = \sum_k p_k \text{Tr}[N_k \rho_k]. \quad (1.44)$$

For the case of  $n = 2$  pure states  $|\psi_1\rangle, |\psi_2\rangle$  and equal probabilities  $p_1 = p_2 = \frac{1}{2}$ , the optimal measurement can be derived analytically and therefore  $S_{\text{med}}$  can be computed exactly. This bound is known as the Helstrom bound [44] and reads

$$S_{\text{med}} = \frac{1}{2}(1 + \sqrt{1 - |\langle\psi_1|\psi_2\rangle|^2}). \quad (1.45)$$

Clearly, if the states are orthogonal, the success probability equals one and the states can be discriminated perfectly. In case that  $|\psi_1\rangle = |\psi_2\rangle$ , the discrimination between them cannot succeed and we have  $S_{\text{med}} = 0$ . Note that the minimum error discrimination problem for  $n = 2$  can also be solved if the two states are mixed and appear with different probabilities [45].

On the other hand, in unambiguous state discrimination, no wrong identification is allowed, that is,  $\text{Tr}[N_k \rho_l] = 0$  whenever  $k \neq l \leq n$ . However, one allows for an inconclusive result corresponding to the case where the identification was not successful. Clearly, if the quantum states  $\rho_j$  are not mutually orthogonal, no perfect discrimination is possible and hence one must allow for an additional measurement outcome, corresponding to this inconclusive result. Consequently, the measurements of interest have now  $m = n + 1$  outcomes. Similar to the previous, the task is to find an optimal measurement maximizing the success probability

$$S_{\text{usd}} = \sum_k p_k \text{Tr}[N_k \rho_k]. \quad (1.46)$$

### Gleason's Theorem

In the axiomatic formalization of quantum theory, which is mostly due to von Neumann [41], the expectation value of an observable  $A$  with respect to a state  $\rho$  is given by  $\langle A \rangle_\rho = \text{Tr}[A\rho]$ . However, one could ask the question whether one can envisage new axioms for quantum theory that are weaker than those introduced so far and would yield statistical predictions that differ from the rule  $\langle A \rangle_\rho = \text{Tr}[A\rho]$ . In particular, this points to the question whether density operators are an appropriate description of quantum states or if there could be a more sophisticated construction. The theorem of Gleason [46] effectively states that there is no such alternative if the dimension of the Hilbert space is larger than 2. More formally, we are interested in the set of functions  $\omega : \mathcal{P}(\mathcal{H}) \rightarrow \mathbb{R}$  with  $\mathcal{P}(\mathcal{H})$  the set of all projection operators of  $\mathcal{H}$  such that

$$0 \leq \omega(\Pi) \leq 1 \quad \forall \Pi \in \mathcal{P}(\mathcal{H}), \quad (1.47)$$

$$\omega(\mathbb{1}) = 1, \quad (1.48)$$

$$\omega\left(\sum_j \Pi_j\right) = \sum_j \omega(\Pi_j), \quad (1.49)$$

where the property in Eq. (1.49) holds for any set of mutually orthogonal projections, that is, subsets  $\{\Pi_j\}_j \subset \mathcal{P}(\mathcal{H})$  with  $\sum_j \Pi_j \leq \mathbb{1}$ . Gleason's theorem then states that if

$\dim(\mathcal{H}) > 2$  there will exist a density operator  $\varrho \in \mathcal{B}(\mathcal{H})$  such that  $\omega(\Pi) = \text{Tr}[\varrho\Pi]$  for all  $\Pi \in \mathcal{P}(\mathcal{H})$ .

### 1.1.3 Quantum channels

Up to this point, the introduced formalism is static and only allows for state changes due to a measurement process. However, after the state preparation and before the measurement, a non-trivial time evolution might occur. This evolution changes the state of the system, for example, via an interaction of the system with the environment or an ambient field.

Let us start by considering the state evolution of an isolated system whose initial state is given by  $|\psi_0\rangle \in \mathcal{H}$ . If the system's Hamiltonian is  $H$ , the evolution is governed by the Schrödinger equation

$$i\hbar \frac{d}{dt} |\psi_t\rangle = H |\psi_t\rangle. \quad (1.50)$$

Note that in general also the Hamiltonian can be time-dependent,  $H = H(t) = H_0 + V(t)$  for some time-dependent potential  $V(t)$ , which often takes the form of an external field. If the Hamiltonian itself is time-independent  $H = H_0$ , the solution of Eq. (1.50) is given by a matrix exponential and will read

$$|\psi_t\rangle = U_t |\psi_0\rangle \quad \text{with} \quad U_t = e^{-\frac{i}{\hbar} H t}, \quad (1.51)$$

where  $|\psi_0\rangle$  is the initial state of the system. The operators  $(U_t)_{t \in \mathbb{R}}$  form a strongly continuous semigroup of unitary operators and  $-i\frac{1}{\hbar}H$  is also called the generator of time shifts [39]. This particularly implies that the evolution of an isolated system is reversible as it is described by a unitary operation. However, there is a complementary approach for justifying why reversible dynamics on pure quantum states is connected to a unitary evolution. From an abstract perspective, we are interested in invertible mappings that transform the pure states of  $\mathcal{H}$  into pure states. Let  $\mathcal{P}_1$  denote the set of all rank-1 projectors associated with  $\mathcal{H}$ , i.e.,  $\mathcal{P}_1 = \{|\psi\rangle\langle\psi| : |\psi\rangle \in \mathcal{H}\}$ . Wigner's theorem [47] then assures that any bijective function  $f : \mathcal{P}_1 \rightarrow \mathcal{P}_1$  fulfilling  $\text{Tr}[f(|\psi\rangle\langle\psi|)f(|\phi\rangle\langle\phi|)] = |\langle\psi|\phi\rangle|^2$  for all  $|\psi\rangle, |\phi\rangle \in \mathcal{H}$  is either unitary, i.e.,  $f(|\psi\rangle\langle\psi|) = U|\psi\rangle\langle\psi|U^\dagger$  for some unitary  $U$  or anti-unitary, i.e.,  $f(|\psi\rangle\langle\psi|) = U(|\psi\rangle\langle\psi|)^\top U^\dagger$  for some unitary  $U$ .

There is an analogue of the Schrödinger equation for mixed states, which is called von Neumann equation. The von Neumann equation dictates that under a Hamiltonian  $H$  the system evolves according to

$$i\hbar \frac{d\varrho}{dt} = [H, \varrho]. \quad (1.52)$$

Similarly to Eq. (1.51), if the Hamiltonian is time-independent, the von Neumann equation can be easily solved and will yield

$$\varrho_t = U_t \varrho_0 U_t^\dagger \quad \text{with} \quad U_t = e^{-\frac{i}{\hbar} H t}, \quad (1.53)$$

where  $\varrho_0$  is the initial density operator of the system. The von Neumann equation has the appealing property that it can be seen as a quantum analogue of the Liouville-equation from classical mechanics<sup>2</sup>. However, it is often the case that the system under consideration is not isolated and interacts with its environment. Such a system is also called an open quantum system. For instance, assume that the system is in the state  $\varrho \in \mathcal{B}(\mathcal{H}_S)$  and couples to an environment which is initially in a pure state  $|\xi\rangle \in \mathcal{H}_E$ .

As the dynamics of the global system, that is, the system together with its environment, is unitary the effective dynamics of the system can be described by

$$\varrho \mapsto \sigma = \text{Tr}_E[U \varrho \otimes |\xi\rangle \langle \xi| U^\dagger] \quad (1.54)$$

for some unitary operator  $U \in \mathcal{B}(\mathcal{H}_S \otimes \mathcal{H}_E)$ . Time evolutions that are of the form given by Eq. (1.54) are obviously trace preserving, that is,  $\text{Tr}[\varrho] = \text{Tr}[\sigma]$ . What is less obvious is that they are in addition completely positive. Mappings being completely positive and trace preserving (CPTP) are also called quantum channels.

A quantum channel  $\Lambda$  is a linear CPTP map

$$\Lambda : \mathcal{B}(\mathcal{H}_A) \rightarrow \mathcal{B}(\mathcal{H}_B), \quad (1.55)$$

where  $\mathcal{H}_A, \mathcal{H}_B$  are finite dimensional vector spaces. A linear map  $\Lambda$  is called positive if it maps positive operators to positive operators. Further,  $\Lambda$  is called completely positive if also the map  $\Lambda \otimes \text{id}_{\mathcal{K}} : \mathcal{B}(\mathcal{H}_A \otimes \mathcal{K}) \rightarrow \mathcal{B}(\mathcal{H}_B \otimes \mathcal{K})$  is positive for *any* finite dimensional Hilbert space  $\mathcal{K}$ .

Given a quantum channel  $\Lambda$  the question is now how it relates to the state evolution in Eq. (1.54). This connection is established by the so-called Stinespring dilation theorem [39, 48]. Suppose that  $\Lambda : \mathcal{S}(\mathcal{H}) \rightarrow \mathcal{S}(\mathcal{H})$  is a quantum channel acting on the system of interest which is in the state  $\varrho$ . Then, there exists an environment, represented by the Hilbert space  $\mathcal{K}$ , a pure state  $|\xi\rangle \in \mathcal{S}(\mathcal{K})$  and a unitary operator  $U : \mathcal{H} \otimes \mathcal{K} \rightarrow \mathcal{H} \otimes \mathcal{K}$  such that  $\Lambda(\varrho) = \text{Tr}_{\mathcal{K}}[U(\varrho \otimes |\xi\rangle \langle \xi|)U^\dagger]$ . Here it is important to notice that the dilation, that is, the triple  $(\mathcal{K}, |\xi\rangle, U)$  is not uniquely determined by the channel  $\Lambda$ . Indeed, an obvious reason for the non-uniqueness of the dilation is that the dimension of the Hilbert space  $\mathcal{K}$  is not limited.

So far, we have always regarded the states as the dynamical objects of the theory while observables are static. This viewpoint is known as the Schrödinger picture. However, there is no reason why not the observables should evolve in time while keeping

---

<sup>2</sup>In statistical mechanics the Liouville equation describes the time evolution of the phase space distribution function.



the quantum state fixed. The description of the evolution of the system from this perspective is referred to as Heisenberg picture. A third complementary viewpoint is the so-called interaction picture where both, states and observables, are dynamical objects.

### Representations of quantum channels

As quantum channels are linear maps, it is in principle sufficient to describe them via their action on some set of basis vectors. This is the approach of the so-called Pauli-transfer matrix (PTM) representation, which offers an elegant description of multi-qubit channels. We have already seen that for an  $n$ -qubit system the set of tensorized normalized Pauli operators  $\{P_a\}_a$  forms an orthonormal basis for the space of observables. With respect to this set of operators one can assign to an arbitrary operator  $A \in \mathcal{B}(\mathcal{H})$  its vector of coordinates via  $|A\rangle\rangle = \{\text{Tr}[\sigma_a A]\}_a$ . A quantum channel  $\Lambda : \mathcal{B}(\mathcal{H}) \rightarrow \mathcal{B}(\mathcal{H})$  is then just a linear map assigning to a state vector  $|\varrho\rangle\rangle$  a new state vector  $|\Lambda(\varrho)\rangle\rangle$ . Consequently, the whole channel can be represented as a matrix whose entries are given by

$$\Lambda_{ab} = \text{Tr}[P_a \Lambda(P_b)] = \langle\langle P_a | \Lambda(P_b) \rangle\rangle. \quad (1.56)$$

A different representation of a quantum channel is the so-called Kraus or operator-sum form [39]. The Kraus decomposition asserts that any quantum channel  $\Lambda : \mathcal{S}(\mathcal{H}) \rightarrow \mathcal{S}(\mathcal{H})$  can be written as

$$\Lambda(\varrho) = \sum_j K_j \varrho K_j^\dagger \quad \text{with} \quad \sum_j K_j^\dagger K_j = \mathbb{1}, \quad (1.57)$$

where  $K_j : \mathcal{H} \rightarrow \mathcal{H}$ . The operators  $\{K_j\}_j$  are called Kraus operators and the minimal number of operators needed in order to represent the channel  $\Lambda$  is called the Kraus rank of  $\Lambda$ . Clearly, a unitary channel  $\Lambda_U(\varrho) := U\varrho U^\dagger$  for a unitary operator  $U$  has Kraus rank one. In general, if  $\dim(\mathcal{H}) = d < \infty$ , then at most  $d^2$  Kraus operators  $K_j$  are needed in order to represent a channel.

Further, there is a bijective correspondence between quantum channels and bipartite quantum states. This relation is known as the Choi-Jamiołkowski isomorphism [39]. For a given quantum channel  $\Lambda : \mathcal{B}(\mathcal{H}) \rightarrow \mathcal{B}(\mathcal{K})$  the correspondence

$$J_\Lambda := (\Lambda \otimes \text{id}_d)(|\Phi_d\rangle\rangle\langle\langle\Phi_d|), \quad \text{Tr}[A\Lambda(B)] = d \text{Tr}[J_\Lambda A \otimes B^\top] \quad (1.58)$$

is bijective for all  $A \in \mathcal{B}(\mathcal{K})$ ,  $B \in \mathcal{B}(\mathcal{H})$  and  $d = \dim(\mathcal{H})$ . The bipartite operator  $J_\Lambda \in \mathcal{B}(\mathcal{K} \otimes \mathcal{H})$  is called the Choi state of the map  $\Lambda$ . The Choi-Jamiołkowski isomorphism enjoys many useful properties, for instance (i)  $J_\Lambda \geq 0$  if and only if  $\Lambda$  is completely positive, (ii)  $\Lambda$  is trace-preserving if and only if  $\text{Tr}_A[J_\Lambda] = d^{-1}\mathbb{1}$ , (iii)  $\Lambda$  is unital if and only if  $\text{Tr}_A[J_\Lambda] = d^{-1}\mathbb{1}$ . Notice that in principle the state  $|\Phi_d\rangle\rangle$  in the construction of the Choi state could be replaced by any other pure quantum state, while obtaining a similar channel-state duality. Such a correspondence will be one-to-one as long as the respective reduced state is of full rank.

### Examples of important quantum channels

In the following we introduce some important quantum channels that usually appear in the context of quantum information processing, where they are used to model different types of noise present in the system.

A very simple model of a quantum channel is the depolarizing noise channel of a qubit system which is given by

$$\mathcal{D}_p(\rho) := (1 - p)\rho + p\frac{\mathbb{1}}{2}, \quad (1.59)$$

where  $p \in [0, 1]$  denotes the rate of error. Physically, this corresponds to an apparatus which outputs with probability  $p$  the depolarized qubit state  $\frac{1}{2}\mathbb{1}$  and with probability  $1 - p$  the state  $\rho$ . Geometrically, this channel yields a contraction of the Bloch sphere, that is, the image of the Bloch sphere under the map  $\mathcal{D}_p$  is a Bloch sphere of radius  $p$ . In order to transform the channel  $\mathcal{D}_p$  into operator-sum form, observe that for an arbitrary  $\rho$  we have the very useful identity

$$\mathbb{1} = \frac{1}{2}(\rho + X\rho X + Y\rho Y + Z\rho Z). \quad (1.60)$$

Now substituting Eq. (1.60) into Eq. (1.59) we obtain

$$\mathcal{D}_p(\rho) = (1 - \frac{3p}{4})\rho + \frac{p}{4}(X\rho X + Y\rho Y + Z\rho Z). \quad (1.61)$$

Amplitude damping is an approximation to a noisy evolution that occurs in many physical systems due to spontaneous emission of a photon from an atom and arises from the inevitable coupling to the vacuum field. As a consequence, the atom tends to decay from its excited state  $|1\rangle$  to its ground state  $|0\rangle$ , even if the atom is in a superposition of the ground and the excited state. The Kraus operators of the amplitude damping channel  $\mathcal{A}_\gamma$  of a two-level system are given by

$$K_1 = \sqrt{\gamma}|0\rangle\langle 1| \quad \text{and} \quad K_2 = |0\rangle\langle 0| + \sqrt{1 - \gamma}|1\rangle\langle 1|, \quad (1.62)$$

where  $\gamma \in [0, 1]$  is the probability for the event that the system decays from the excited state  $|1\rangle$  to the ground state  $|0\rangle$ . It is obvious from the Kraus operators  $K_1, K_2$  that  $\mathcal{A}_\gamma$  leaves the ground state  $|0\rangle$  invariant. On the other hand, the state  $|1\rangle$  maps to  $\gamma|0\rangle\langle 0| + (1 - \gamma)|1\rangle\langle 1|$ , thus damping the amplitude of the excited state  $|1\rangle$ . In the more general setting of an  $n$ -qubit system, one typically assumes that in first approximation spontaneous emission acts independently on each of the qubits [49]. Therefore, the  $n$ -qubit amplitude damping channel is simply the  $n$ -fold tensor product of the single qubit channel  $\mathcal{A}_\gamma$ .

### Quantum instruments

In general, if a quantum measurement was not destructive, the system will still exist after the measurement and one may aim to extract more information about the system by performing a subsequent measurement. A different way of interpreting this is

by viewing a quantum measurement as a state preparator. For instance, we will see in Chapter 5 that local filtering is a process where upon a measurement and post-selecting on a particular outcome a state with a higher purity is distilled. Both scenarios motivate a concept which is known as quantum instrument [39]. An instrument  $\mathcal{I}$  is a collection of completely positive trace-nonincreasing maps  $\{\mathcal{I}_\omega\}_{\omega \in \Omega}$  where  $\Omega$  is the outcome space such that  $\sum_{\omega \in \Omega} \mathcal{I}_\omega = \Lambda$  with  $\Lambda$  a quantum channel. If an instrument  $\{\mathcal{I}_\omega\}_{\omega \in \Omega}$  acts on a quantum state  $\rho$  it yields in dependence of the observed outcome  $\omega \in \Omega$  the non-normalized state  $\mathcal{I}_\omega(\rho)$ . The probability to observe the outcome  $\omega$  is given by  $\text{Tr}[\mathcal{I}_\omega(\rho)]$  and the state of the system after the measurement is given by  $\rho_\omega = \mathcal{I}_\omega(\rho) / \text{Tr}[\mathcal{I}_\omega(\rho)]$ .

Our main aim for introducing quantum channels is to complete the list of update rules for post-measurement quantum states, see Section 1.1.2. Until now, we only have defined a POVM on a purely statistical level without actually describing how it can be implemented or how the state of the system changes after the measurement. A typical instance of such an implementation is the so-called Lüders instrument. Given a POVM  $\{E_\omega\}_{\omega \in \Omega}$  one defines the maps

$$\mathcal{I}_\omega(\rho) := \sqrt{E_\omega} \rho \sqrt{E_\omega}. \quad (1.63)$$

Using the cyclic property of the trace one can directly verify that for a given state  $\rho$  the instrument recovers the correct output statistics, i.e.,

$$\text{Tr}[\mathcal{I}_\omega(\rho)] = \text{Tr}[\sqrt{E_\omega} \rho \sqrt{E_\omega}] = \text{Tr}[E_\omega \rho]. \quad (1.64)$$

Accordingly, the state of the system upon observing  $\omega \in \Omega$  is given by

$$\rho \mapsto \rho_\omega = \frac{\mathcal{I}_\omega(\rho)}{\text{Tr}[\mathcal{I}_\omega(\rho)]} = \frac{\sqrt{E_\omega} \rho \sqrt{E_\omega}}{\text{Tr}[\rho E_\omega]}. \quad (1.65)$$

### 1.1.4 The quantum measurement problem

Although the name suggests it, the quantum measurement problem is not solely concerned with the role of measurements in quantum theory but points to a much more general problem, the quantum-to-classical transition. The core question is how classical systems and classical properties that we experience in the macroscopic world can emerge from the underlying quantum domain. Indeed, we do not observe superpositions of macroscopic distinguishable positions and this paradox manifests itself in the context of measurements in quantum theory. The results and definitions presented in this Section are covered in Ref. [50].

#### The ideal von Neumann scheme

Even though a measurement apparatus appears classical, e.g., takes definite and well distinguishable values, it consists of atoms which should admit a purely quantum

mechanical description. It is therefore a natural question whether one can describe the process of a measurement entirely by quantum theory, i.e., model the physical interaction between system and apparatus using the formalism of quantum mechanics. However, the von Neumann scheme of a quantum measurement applies to much more situations, as it gives a simple explanation to the question how entanglement between systems arises.

A quantum measurement typically involves a microscopic system  $S$  and a measurement apparatus  $A$ . Both of them are treated as quantum systems. This means that one associates to  $S$  the Hilbert space  $\mathcal{H}_S$  with basis  $\{|s_j\rangle\}$  and to  $A$  the Hilbert space  $\mathcal{H}_A$  with basis  $\{|a_j\rangle\}$ . The states of the apparatus should correspond to the different positions that a pointer can take, indicating the result of the measurement. As those positions are typically distinguishable one assumes that the pointer states  $|a_j\rangle$  are mutually orthonormal. The measurement process is now a dynamical interaction between the system  $S$  and the apparatus. As we aim to infer the state of  $S$  by means of  $A$ , the interaction should be of the form

$$|s_j\rangle_S |R\rangle_A \mapsto |s_j\rangle_S |a_j\rangle_A \quad (1.66)$$

for all possible values of  $j$  where  $|R\rangle$  denotes the initial state of the apparatus. The appearance of  $|a_j\rangle$  implies that the system was in the state  $|s_j\rangle$ . Further, if the system  $S$  was in state  $|s_j\rangle$  the joint system will be found after the interaction in the state  $|s_j\rangle|a_j\rangle$ . Consequently, Eq. (1.66) yields a bijective correspondence between system states and pointer states. This relation also motivates the name *ideal* measurement, as the interaction does not disturb the initial state of the system, i.e., after the interaction with the apparatus the system  $S$  is still in the state  $|s_j\rangle$ .

Applying the evolution defined in Eq. (1.66) to an arbitrary superposition of basis states  $|\psi_0\rangle = \sum_j c_j |s_j\rangle$  with  $c_j \in \mathbb{C}$  yields

$$|\psi_0\rangle |R\rangle \mapsto |\psi_t\rangle = \sum_j c_j |s_j\rangle |a_j\rangle \in \mathcal{H}_S \otimes \mathcal{H}_A. \quad (1.67)$$

The evolution of the factorized quantum state  $|\psi_0\rangle |R\rangle$  into a superposition of system-apparatus states  $|\psi_t\rangle$  represents the von Neumann quantum measurement scheme. Clearly, the superposition has been broadcasted from the system  $S$  to the apparatus. Consequently, it is not possible to ascribe an individual state to the system  $S$ . This raises a serious problem. If the evolution in Eq. (1.67) describes the whole measurement, how can the state  $|\psi_t\rangle$  reflect our experience of definite outcomes? This points to the fact that  $|\psi_t\rangle$  does not yield a complete description of the measurement. In this context, the evolution in Eq. (1.67) is also called a *pre-measurement* [50].

### Unfolding the measurement problem

The notion of a pre-measurement offers a precise formulation of the quantum-to-classical transition and its application to quantum measurements. In the following we will discuss how the measurement problem can be decomposed into three sub-problems, where two of them can be resolved via the so-called decoherence program [50–53].

- (1) **The preferred basis problem:** If one considers the von Neumann scheme in Eq. (1.67), it turns out that the final state  $|\psi_t\rangle$  does not uniquely determine the observable that was intended to be measured. For instance, suppose that we intend to measure  $\sigma_3$  and that the initial state of the system is the +1 eigenstate of  $\sigma_1$ , i.e.,  $\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$ . If the pointer states are denoted by  $|0\rangle$  and  $|1\rangle$ , then the state after the evolution is given by

$$|\psi_t\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle) = \frac{1}{\sqrt{2}}(|\vec{n}_-\rangle|\vec{n}_-\rangle + |\vec{n}_+\rangle|\vec{n}_+\rangle), \quad (1.68)$$

for *any* spin direction  $\vec{n}$  (cf. Schmidt decomposition in Section 1.2.2). However, this implies that the device  $A$  has formed correlations with both  $\sigma_3$  and  $\sigma_{\vec{n}}$ . Consequently, it seems that the device  $A$  has measured simultaneously two noncommuting observables in contradiction with the laws of quantum mechanics.

- (2) **The nonobservability of interference:** From Eq. (1.67) it follows that superpositions of macroscopic objects such as measurement devices should appear in nature. However, for systems of macroscopic size interference effects are typically observed to vanish. This raises the question why we do not observe macroscopic interference while being predicted by the von Neumann scheme.
- (3) **The problem of outcomes:** On the one hand, if one conducts a measurement on a system, one would expect to obtain a definite result, e.g., a distinguishable pointer position. On the other hand, the von Neumann scheme in Eq. (1.67) predicts that the final state will be in a superposition of system-apparatus states. This raises the problem of how these different scenarios can be reconciled. However, even if one can explain the appearance of definite pointer states the question remains how one arrives at a particular outcome. These two problems are termed the problem of outcomes [50].

### The thought experiment of Wigner

The problem of outcomes raises the question how one can explain the experience of an observer to obtain a definite, classical outcome. This seems to point to the fact that a division of the world into a part that is treated with quantum theory and a part that is not, i.e., the classical apparatus, is necessary. However, if the location of

the quantum-classical boundary marks the distinction between the observer and what is being observed, what about situations that involve multiple observers? More precisely, is each observer permitted to treat the other observer as a quantum system? Questions of this form can be seen as reformulations of the problem when and where the collapse-inducing measurement takes place.

Consideration of scenarios involving multiple observers was initiated by Wigner [54] and has the name thought experiment for two reasons. First, it is a purely hypothetical experiment. Second, Wigner designed the experiment in order to support his view that consciousness, i.e., a thoughtful observer, is necessary to complete the quantum measurement process.

Suppose that a friend, named Alice, is located in a laboratory that is perfectly isolated from its environment. Within the lab she has access to a microscopic physical two-level system  $S$  on which she can perform a dichotomic measurement, which can yield outcomes  $a = 0$  or  $a = 1$ . Further, outside of the lab, there is Wigner who aims to describe the experiment which his friend Alice has implemented inside the lab. For him, the whole lab appears as a physical system whose evolution is governed by the Schrödinger equation. For simplicity the friend Alice is also described by a two-level system as this abstraction captures all the relevant information, even though Alice can be of arbitrary complexity. We assume that before Alice implements the measurement she is in some state  $|R\rangle_A \in \mathcal{H}_A$ . The von Neumann scheme then implies that

$$|0\rangle_S |R\rangle_A \mapsto |0\rangle_S |0\rangle_A \quad \text{and} \quad |1\rangle_S |R\rangle_A \mapsto |1\rangle_S |1\rangle_A. \quad (1.69)$$

Suppose that the system  $S$  was initially prepared in the state  $|+\rangle$ , a fact known to Wigner. As he treats the lab as an isolated system, after the interaction, he assigns the state

$$|+\rangle |R\rangle \mapsto |\psi\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle). \quad (1.70)$$

Wigner then enters the room to inquire about the outcome of the experiment. Alice, describing her measurement according to the measurement postulate will answer that she has observed either 0 or 1. Then Wigner has to conclude that from his perspective, the superposed state in Eq. (1.70) has collapsed onto either one of its two components  $|00\rangle$  or  $|11\rangle$ . However, what would Alice have said about the outcome of the experiment before Wigner had entered the lab and asked the first question?

At the time Wigner proposed the experiment, he concluded that the superposition in Eq. (1.70) must be regarded as absurd as it contains two "distinct states of consciousness" [54]. For him, consciousness must break the unitary evolution and induces a collapse of the wave function onto a definite state of the conscious observer [54].

### Deutsch's version of Wigner's friend

The thought experiment developed by Deutsch [55] is formally similar to the previous one, but introduces two novel ideas to the setup. These two ideas will become relevant in Section 3 when discussing the measurement problem from the viewpoint of so-called local friendliness correlations. First, there can be a physical record of an experiment being performed in the past, which can exist independently from its outcome. More precisely, it allows Wigner to obtain information on whether the friend has observed a definite outcome upon her measurement or not without revealing which particular outcome has been observed. Second, Deutsch added a further step at the end of the protocol where Alice's measurement of the system  $S$  is undone. This has the surprising consequence that Wigner can experimentally distinguish whether or not Alice's lab was in a superposition before he entered the laboratory.

Formally the scenario is as follows: The friend inside the lab does not only record what she observed but also whether or not she has observed a definite outcome. For this purpose, one allows the system where Alice stores this information, i.e., the information of outcome and of the definiteness of the result, to be of the form  $\mathcal{H}_{A_1} \otimes \mathcal{H}_{A_2}$ . Assume that she is initially in the state  $|0\rangle_{A_1}|0\rangle_{A_2}$ . The correspondence between the measurement outcome and the internal state is given by

$$|0\rangle_S|0\rangle_{A_1}|0\rangle_{A_2} \mapsto |0\rangle_S|0\rangle_{A_1}|1\rangle_{A_2} \quad \text{and} \quad |1\rangle_S|0\rangle_{A_1}|0\rangle_{A_2} \mapsto |1\rangle_S|1\rangle_{A_1}|1\rangle_{A_2}. \quad (1.71)$$

If the system is initially in the state  $|+\rangle$ , then the interaction according to Eq. (1.71) yields the state

$$|\psi\rangle = \frac{1}{\sqrt{2}}(|0\rangle_S|0\rangle_{A_1}|1\rangle_{A_2} + |1\rangle_S|1\rangle_{A_1}|1\rangle_{A_2}) = \frac{1}{\sqrt{2}}(|0\rangle_S|0\rangle_{A_1} + |1\rangle_S|1\rangle_{A_1})|1\rangle_{A_2}. \quad (1.72)$$

The important point is now that we can reveal the information whether Alice has observed a definite outcome or not without inferring any information about the particular realized outcome. For this, introduce two measurement operators

$$\mathcal{B}(\mathcal{H}_{A_1} \otimes \mathcal{H}_{A_2}) \ni P = |01\rangle\langle 01| + |11\rangle\langle 11| \quad (1.73)$$

and

$$\mathcal{B}(\mathcal{H}_{A_1} \otimes \mathcal{H}_{A_2}) \ni Q = |10\rangle\langle 10| + |11\rangle\langle 11| - |00\rangle\langle 00| - |01\rangle\langle 01|. \quad (1.74)$$

Clearly  $P$  is a projector and has the two eigenvalues 0, 1, where the corresponding subspaces are degenerate and of dimension two. The  $\lambda = 0$  subspace is spanned by  $|00\rangle$  and  $|10\rangle$  while the  $\lambda = 1$  subspace is spanned by  $|01\rangle$  and  $|11\rangle$ . The eigenvalues of  $Q$  are given by  $\pm 1$ , each with degeneracy two. The  $\lambda = 1$  subspace is spanned by  $|10\rangle$  and  $|11\rangle$  while the  $\lambda = -1$  subspace is spanned by  $|00\rangle$  and  $|01\rangle$ . The form of the observables  $P$  and  $Q$  has been chosen in a thoughtful way. Indeed, a measurement of

$P$  reveals whether the friend sees a definite outcome or not while a measurement of  $Q$  yields the outcome observed by Alice. Here it is crucial that the measurement of  $P$  does not contain any information about the actually observed value  $a = 0$  or  $a = 1$ . If it would, then the state in Eq. (1.72) would not factorize with respect to the bipartition  $SA_1|A_2$ . In addition, if Wigner would have access to the information of the realized outcome, lets say  $a = 1$ , he would assign to Alice's lab the state  $|1\rangle_S|1\rangle_{A_1}|1\rangle_{A_2}$  rather than a superposition.

In the final step, after Alice has completed the measurement, that is, she has updated her knowledge from  $|00\rangle$  to either  $|01\rangle$  or  $|11\rangle$  depending on the state of the system, some external control is applied to Alice's entire lab  $SA_1A_2$ . This control restores the information about the observable  $Q$  and the system but not about  $P$ . More precisely, the action on basis kets is given by

$$|001\rangle \mapsto |001\rangle, \quad |111\rangle \mapsto |100\rangle. \quad (1.75)$$

Even though this evolution may be difficult to realize in practice, e.g., the system may be of macroscopic size, it is not forbidden if one assumes quantum theory to be universally valid. Thus assuming that it can be implemented, the resulting state of Alice's lab is

$$|\psi\rangle \mapsto |\eta\rangle = \frac{1}{\sqrt{2}}(|001\rangle + |101\rangle) = |+\rangle|0\rangle|1\rangle. \quad (1.76)$$

The state  $|\eta\rangle$  does not contain any information about the outcome of the measurement and the system  $S$  has returned to its initial state. However, there still exists a record that the experiment had been performed in the past. Now Wigner can test this superposition by applying a direct measurement of  $\sigma_1$  to  $S$ . As  $|+\rangle$  is an eigenstate of  $\sigma_1$ , quantum theory predicts that the outcome is with certainty  $+1$ . On the other hand, according to the measurement postulate, after Alice completes her measurement the state would be in one of the states  $|001\rangle$  or  $|111\rangle$ . Applying the external control would yield the state  $|000\rangle$  or  $|100\rangle$ . Wigner's final measurement with respect to  $\sigma_1$  would thus yield uniformly random outcomes. Therefore, this scheme distinguishes whether or not Alice's lab was in a superposition state.

## 1.2 Correlations and quantum theory

### 1.2.1 The EPR argument

Already in 1935 quantum theory had established itself as a very successful and accurate theory and seemed *fine for all practical purposes* [56]. However, due to its abstract formulation involving Hilbert spaces, it was not clear to what extent the quantum formalism relates to the physical reality. In particular, one question was whether the



quantum state  $|\psi\rangle$  mirrors the physical reality perfectly, i.e., whether the wave function corresponds to something physical real.

The argument of Einstein, Podolsky and Rosen (EPR) from 1935 picks up this question and asks for the completeness of quantum theory [3]. Even though they do not attempt to fully characterize what a complete theory is, they present a necessary condition. They demand that every element of the physical reality must have a counterpart in the physical theory. More precisely, if one can extinguish an element of reality for which the physical theory cannot provide a corresponding counterpart, the physical theory cannot be regarded as complete [50]. Further, they regard as a sufficient condition for a physical quantity to be an element of reality, that it can be predicted with certainty without disturbing the system. This resembles our intuition that a measurement of a physical quantity reveals a property that has already existed before the measurement was performed. In their work, EPR apply the criterion for elements of physical reality to a composite system of distant particles, i.e., any action on the first system cannot affect the physical situation of the second. In the version of Bohm [57], one considers a pair of qubits in the state  $|\psi^-\rangle$ . We have already seen in Eq. (1.68) that

$$|\psi^-\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle) = \frac{1}{\sqrt{2}}(|--\rangle + |++\rangle), \quad (1.77)$$

where  $|\pm\rangle$  are the eigenstates of  $\sigma_1$ . If now Alice decides to measure  $\sigma_3$ , the wave function on Bob's side changes to  $|0\rangle$  or  $|1\rangle$  depending on Alice's outcome. In particular, as Bob's state after the measurement is in an eigenstate of  $\sigma_3$ , the measurement of  $\sigma_3$  on the second system can be predicted with certainty. However, if  $\sigma_1$  is measured on Alice's side, the state of Bob will change to  $|+\rangle$  or  $|-\rangle$ , and thus also can be predicted with certainty. If one now assumes that both parties are sufficiently separated, the measurement on the first subsystem does not cause a change in the second subsystem. Therefore both,  $\sigma_1$  and  $\sigma_3$ , simultaneously correspond to elements of reality. But quantum theory precludes the simultaneous assignment of definite values to noncommuting operators as  $\sigma_1$  and  $\sigma_3$ . From this EPR concluded that quantum theory must be incomplete.

### 1.2.2 Quantum entanglement

We have seen in Section 1.1.1 that a pure multipartite quantum state is described by a state vector of the form

$$|\psi\rangle = \sum_{j_1, \dots, j_n} \psi_{j_1 \dots j_n} |j_1\rangle \cdots |j_n\rangle \quad (1.78)$$

for certain coefficients  $\psi_{j_1 \dots j_n} \in \mathbb{C}$ . In general, those states cannot be written as a tensor product of local tensor factors, i.e.,  $|\psi\rangle \neq |\psi_1\rangle \cdots |\psi_n\rangle$ . If  $|\psi\rangle$  factorizes, it is called a product state. If  $|\psi\rangle$  is not a product state, this has the consequence that it is not

possible to assign a single state vector to each of the  $n$  subsystems. The fact that  $|\psi\rangle$  is not factorizable expresses on a formal level that  $|\psi\rangle$  is *entangled*. Easy examples of entangled states are the Bell states introduced in Eq. (1.6). We will refer to entanglement in pure states as pure state entanglement. In the case that  $|\psi\rangle$  is a product state,  $|\psi\rangle = |\psi_1\rangle \cdots |\psi_n\rangle$ , it is also called fully separable. Further, we call  $|\psi\rangle$   $m$ -separable if there exists a partition  $P = \{P_1, \dots, P_m\}$  of the set  $\{1, \dots, n\}$  such that

$$|\psi\rangle = \otimes_{j=1}^m |\psi_{P_j}\rangle, \quad (1.79)$$

where  $|\psi_{P_j}\rangle \in \mathcal{H}_{P_j} := \otimes_{k \in P_j} \mathcal{H}_k$ . This means that while the state factorizes with respect to the given partition  $P$ , it could still be entangled within the subspaces  $\mathcal{H}_{P_j}$ . If each party independently measures an observable, that is, party  $j$  measures  $A_j$ , on a product state, the expectation value of  $A = A_1 \otimes \cdots \otimes A_n$  factorizes as

$$\langle A \rangle_{|\psi\rangle} = \text{Tr}[A_1 |\psi_1\rangle \langle \psi_1|] \cdots \text{Tr}[A_n |\psi_n\rangle \langle \psi_n|]. \quad (1.80)$$

However, for a realistic description of an experiment, mixed quantum states are the appropriate notion. In order to extend the concept of entanglement to mixed states one characterizes those states that should *not* fall into that class. This yields the set of so-called classical correlated quantum states and entangled states are the complement thereof. Consider a scenario where two parties Alice (A) and Bob (B) receive one part of a quantum system  $\varrho \in \mathcal{B}(\mathcal{H}_A \otimes \mathcal{H}_B)$ , each originating from a different preparation devices. Suppose that each of the devices  $P \in \{A, B\}$  has  $j = 0, \dots, n-1$  different settings and produces upon choice a quantum state  $\varrho_P^j$ . Further suppose that in addition to the two preparation devices one has access to a random number generator which yields random numbers among the set  $\{0, \dots, n-1\}$  with probability  $p_j$ . Now, in each round of the experiment, one can first draw a random number  $j$  and accordingly to the output, one prepares the states  $\varrho_P^j$ . If Alice measures observable  $O_A$  and Bob measures observable  $O_B$ , the observable statistics would be

$$\langle O_A \otimes O_B \rangle_\varrho = \sum_{j=1}^{n-1} p_j \text{Tr}[\varrho_A^j O_A] \text{Tr}[\varrho_B^j O_B] = \text{Tr}[O_A \otimes O_B \varrho], \quad (1.81)$$

where

$$\varrho = \sum_j p_j \varrho_A^j \otimes \varrho_B^j. \quad (1.82)$$

States that are of the form Eq. (1.82) are called classical correlated or separable [58]. Notice that for those classical correlated states the expectation value in Eq. (1.81) does not factorize as it was the case in Eq. (1.80). The extension to the multipartite case is straightforward. A mixed quantum state is called  $m$ -separable if it can be written as a convex combination of  $m$ -separable pure states. Notice that the pure states in the

convex decomposition do not need to be  $m$ -separable with respect to the same partition. Clearly, if the state is  $m$ -separable, then it is also  $(m - 1)$ -separable. This yields a hierarchical structure of sets among which 2-separability is the weakest notion of separability, which is also called biseparability. If a multipartite state is not biseparable, it is called genuinely multipartite entangled. In general, in order to determine whether a quantum state  $\rho$  is entangled or not, one has to check whether it admits a separable state decomposition as in Eq. (1.82). This problem is known as the separability problem. Unfortunately, this problem is computationally difficult and it is known to be NP-hard [59,60]. However, this does not imply that there do not exist certain subsets of states for which entanglement could be decided efficiently. As we will see later, a variety of entanglement criteria has been developed aiming exactly for that detection.

### Pure bipartite entanglement and the Schmidt decomposition

Bipartite pure entanglement has a particularly easy structure and a complete classification is possible. This is mostly due to the Schmidt decomposition, a direct consequence of the singular value decomposition, which is an indispensable tool in entanglement theory.

**Theorem 1.** *Let  $|\psi\rangle \in \mathcal{H}_A \otimes \mathcal{H}_B$  be a quantum state. Then there exists a local basis  $\{|j\rangle_A\}_j$  for  $\mathcal{H}_A$  and a local basis  $\{|j\rangle_B\}_j$  for  $\mathcal{H}_B$  such that*

$$|\psi\rangle = \sum_{j=0}^{d-1} \lambda_j |j\rangle_A \otimes |j\rangle_B \quad (1.83)$$

with positive, uniquely determined Schmidt coefficients  $\lambda_j$  and Schmidt rank  $d = \min(d_A, d_B)$  where  $d_j = \dim(\mathcal{H}_j)$ . If the  $\lambda_j$  are pairwise different, then also the Schmidt vectors  $|j\rangle_A$  and  $|j\rangle_B$  are unique up to a phase.

The idea behind the proof of Theorem 1 is as follows. If the state is presented in an arbitrary basis  $|\psi\rangle = \sum_{j,k} \psi_{jk} |jk\rangle$ , one can regard the coefficients as a matrix  $C_{ij} = \psi_{ij}$ . A singular value decomposition then yields  $C = UDV^\dagger$  where  $U, V$  are unitary matrices and  $D$  is rectangular diagonal matrix with non-negative real numbers on the diagonal. It can be shown that the overall cost of computing the singular value decomposition of a matrix  $C \in \mathbb{C}^{m \times n}$  requires  $\mathcal{O}(mn^2)$  operations [61]. From the decomposition of the state in Eq. (1.83) one can directly see whether or not the state is entangled. Consequently, there exists an efficient algorithm for deciding whether a bipartite pure state is entangled.

We will call  $|\psi\rangle \in \mathcal{H} \otimes \mathcal{H}$  with  $d = \dim(\mathcal{H})$  a maximally entangled state, if the Schmidt coefficients are given by  $\lambda_j = \frac{1}{\sqrt{d}}$  for  $1 \leq j \leq d$ . Indeed, the state  $|\Phi_d\rangle$  introduced in Eq. (1.7) is maximally entangled with respect to that definition. It should be noticed that the demand of pairwise different Schmidt coefficients  $\{\lambda_j\}_j$  is crucial for the uniqueness of the Schmidt vectors. For instance, consider the state  $|\phi^+\rangle$  from

Eq. (1.6), which is already given in the form of a Schmidt decomposition. The Schmidt vectors correspond to the eigenstates of  $\sigma_3$  and the Schmidt coefficients are given by  $\lambda_1 = \lambda_2 = \frac{1}{\sqrt{2}}$ . Due to this full degeneracy, this state admits a Schmidt decomposition with respect to *any* spin direction. If we denote by  $\sigma_{\vec{n}} := \vec{\sigma} \cdot \vec{n} = \sum_j n_j \sigma_j$  and by  $|\vec{n}_{\pm}\rangle$  the eigenvector corresponding to eigenvalue  $\pm 1$  of the operator  $\sigma_{\vec{n}}$  then

$$|\phi^+\rangle = \frac{1}{\sqrt{2}}(|z_-\rangle|z_-\rangle + |z_+\rangle|z_+\rangle) = \frac{1}{\sqrt{2}}(|\vec{n}_-\rangle|\vec{n}_-\rangle + |\vec{n}_+\rangle|\vec{n}_+\rangle) \quad (1.84)$$

for any spin direction  $\vec{n} \in \mathbb{R}^3$  with  $\|\vec{n}\| = 1$ . Further, states of that form fulfill many useful identities. For instance, any maximally entangled state  $|\psi\rangle$  is of the form  $|\psi\rangle = (\mathbb{1} \otimes U)|\Phi_d\rangle$  for some unitary operator  $U$  and  $|\Phi_d\rangle$  as defined in Eq. (1.7). Further, for any  $A, B \in \mathcal{B}(\mathcal{H})$  one has

$$\text{Tr}[A \otimes B |\Phi_d\rangle\langle\Phi_d|] = \frac{1}{d} \text{Tr}[A^\top B], \quad (A \otimes \mathbb{1})|\Phi_d\rangle = (\mathbb{1} \otimes A^\top)|\Phi_d\rangle. \quad (1.85)$$

In addition, any bipartite pure state  $|\psi\rangle$  with  $\varrho_B = \text{Tr}_A[|\psi\rangle\langle\psi|]$  can be written as

$$|\psi\rangle = (\mathbb{1} \otimes C)|\Phi_d\rangle, \quad (1.86)$$

where  $C = \sqrt{d} \sqrt{\varrho_B} V$  with  $V : \mathcal{H} \rightarrow \mathcal{H}$  an isometry [39].

There also exists a Schmidt decomposition for bipartite mixed quantum states  $\varrho \in \mathcal{B}(\mathcal{H}_A \otimes \mathcal{H}_B)$ , which is mathematically a direct consequence of Theorem 1. Any density operator  $\varrho$  can be written as

$$\varrho = \sum \lambda_j G_j^A \otimes G_j^B, \quad (1.87)$$

where  $\lambda_j \geq 0$  and the operators  $\{G_j^A\}_j, \{G_j^B\}_j$  form an orthonormal basis for the vector space of observables  $\mathcal{B}_H(\mathcal{H}_A), \mathcal{B}_H(\mathcal{H}_B)$ , respectively. However, apart from the analogy to the pure case, no necessary and sufficient criterion is known to decide whether a quantum state  $\varrho$  is entangled or not solely based on its Schmidt decomposition.

### Multipartite entanglement and its classification

A different approach to gain insight into the entanglement properties of quantum states is to consider their interconvertibility. More precisely, one could ask whether two given quantum states  $|\psi_1\rangle$  and  $|\psi_2\rangle$  could be transformed into each other by means of a set of *local operations*. One popular class of such local operations is the set of local unitary (LU) operations. In this context, one calls the  $n$ -partite states  $|\psi_1\rangle$  and  $|\psi_2\rangle$  LU equivalent if there exist local unitary operators  $U_1, \dots, U_n$  such that  $|\psi_2\rangle = U_1 \otimes \dots \otimes U_n |\psi_1\rangle$ . This notion directly generalizes to mixed quantum states.  $\varrho_1$  and  $\varrho_2$  are LU equivalent if and only if  $\varrho_2 = U \varrho_1 U^\dagger$  with  $U = U_1 \otimes \dots \otimes U_n$ . Such a local unitary operation only changes the local basis of the quantum state and is obviously locally reversible. From a mathematical viewpoint, in the simplified case of

$n$ -qubits, one considers the action of the unitary group  $U(2)^{\times n}$  on the space  $(\mathbb{C}^2)^{\otimes n}$ , where each copy of  $U(2)$  acts on a different spin system, i.e. on the corresponding copy of  $\mathbb{C}^2$ . Then one asks for the orbits under the action of the local transformation group, that is,  $(\mathbb{C}^2)^{\otimes n} / U(2)^{\times n}$ . However, even for the smallest case of a two-qubit system, continuous parameters are needed to label all equivalence classes [62]. Hence, there are infinitely many different forms of entanglement. For pure states, the question of LU equivalence of  $n$ -partite qubit states can in principle be solved via so-called local polynomial invariants [63], which are polynomials that are invariant under local unitary transformations. Even though one can prove that the set of such invariants is finitely generated [64], i.e., it is sufficient to only consider a finite set of them, complete finite sets are only known for very few simple cases.

In general, a simpler classification of multipartite entanglement classes would be advisable. These new equivalence classes would then be given as coarse grainings of the former, fine-grained equivalence classes. Clearly, to obtain less equivalence classes, the set of allowed transformations must be enlarged. One of such larger classes of local transformations is given by the set of local operations and classical communication (LOCC). We will illustrate this for the bipartite case with parties named Alice and Bob. A channel  $\Lambda : \mathcal{S}(\mathcal{H}_A \otimes \mathcal{H}_B) \rightarrow \mathcal{S}(\mathcal{H}_A \otimes \mathcal{H}_B)$  will be called a LOCC channel, if it can be written as a sequence of instruments and channels as well as an exchange of classical communication. In particular, no quantum information between the different parties can be transferred and no entanglement between the parties can be created. Note that the quantum channel that is applied by the parties in the  $n$ -th run of the LOCC protocol can depend on the operations and possible measurement outcomes in the previous  $n - 1$  rounds. Assume that the protocol initializes with Bob performing a measurement with outcome space  $\Omega_1$ , which is described by an instrument  $\{\mathcal{I}_{\omega_1}\}_{\omega_1 \in \Omega_1}$ , such that  $\sum_{\omega_1 \in \Omega_1} \mathcal{I}_{\omega_1} =: \mathcal{E}_1^B$  is a quantum channel. Upon observing outcome  $\omega_1$ , he communicates  $\omega_1$  to Alice. This ends the first round of the LOCC protocol. The second round starts with Alice choosing a measurement with outcome space  $\Omega_2$ , corresponding to an instrument  $\{\mathcal{I}_{\omega_2}\}_{\omega_2 \in \Omega_2}$ . The particular choice can depend on the received information  $\omega_1$ . Upon applying the instrument and obtaining outcome  $\omega_2$ , Alice communicates her outcome  $\omega_2$  to Bob, who continues his action, and so on. After the  $n$ -th information exchange, Alice and Bob each can apply a local channel which is allowed to depend on all the previous communicated outcomes, that is,  $\Lambda_{n+1|\omega_n \dots \omega_1}^A \otimes \Lambda_{n+1|\omega_n \dots \omega_1}^B$ . Finally, as a result of the whole protocol, Alice and Bob have applied a LOCC channel of the form

$$\Lambda = \sum_{\omega_1 \dots \omega_n} (\Lambda_{n+1|\omega_n \dots \omega_1}^A \otimes \Lambda_{n+1|\omega_n \dots \omega_1}^B) \cdots (\mathcal{I}_{\omega_2|\omega_1}^A \otimes \mathcal{I}_{\omega_1}^B). \quad (1.88)$$

It can be shown that for the case of pure quantum states deciding LOCC equivalence of quantum states reduces to deciding LU equivalence [65].

It is often convenient to distinguish between different subclasses of LOCC protocols.

For instance, the class  $\text{LOCC}_0$  is the class of local operations where no communication between the parties is allowed. In this case, the general LOCC channel in Eq. (1.88) collapses to  $\Lambda = \Lambda_A \otimes \Lambda_B$  where  $\Lambda_A, \Lambda_B$  are quantum channels acting on the respective systems. A larger LOCC class is the set  $\text{LOCC}_{1a}$ , where a single communication round from Alice to Bob is allowed. Alice makes a measurement described by an instrument  $\{\mathcal{I}_\omega\}_{\omega \in \Omega}$  and obtains outcome  $\omega$ , which she communicates to Bob. Upon receiving Alice's output, he applies a channel  $\Lambda_\omega^B$  that can depend on  $\omega$ . This results in a so-called LOCC-1a channel [66] and is of the form

$$\text{LOCC}_{1a} \ni \Lambda = \sum_{\omega \in \Omega} \mathcal{I}_\omega \otimes \Lambda_\omega^B. \quad (1.89)$$

Yet another different class of local transformations is the set of stochastic local operations and classical communication (SLOCC). This class is conceptually the same as LOCC with the difference that the state conversion has not to be achieved with certainty. Mathematically, one can show that an equivalent definition is the following [67]: Two  $n$ -partite quantum states  $|\psi_1\rangle, |\psi_2\rangle \in \otimes_{j=1}^n \mathcal{H}_j$  will be SLOCC equivalent, if and only if there exist  $A_j \in \text{GL}(\mathcal{H}_j)$  for  $1 \leq j \leq n$ , such that

$$|\psi_2\rangle = A_1 \otimes \cdots \otimes A_n |\psi_1\rangle, \quad (1.90)$$

where  $\text{GL}(\mathcal{H})$  denotes the group of all invertible operators acting on  $\mathcal{H}$ . In difference to the case of LU equivalence, the number of orbits for the bipartite case is finite. Indeed, as SLOCC transformations cannot increase the Schmidt rank of a bipartite quantum state, the Schmidt rank is a SLOCC invariant. However, it turns out that the Schmidt rank of the quantum state is the only invariant for SLOCC orbits and consequently, when considering a system of the form  $\mathbb{C}^{d_A} \otimes \mathbb{C}^{d_B}$ , the number of SLOCC orbits is simply given by  $\min(d_A, d_B)$ . If in the bipartite case the Schmidt rank of the state is larger than 1, then the state is automatically genuine entangled. Consequently, the maximally entangled state  $|\Psi_d\rangle$  is a representative of the class of genuine entangled states and it is sufficient to only consider this particular state.

Interestingly, for the case of three qubits the situation changes. Here, there are six inequivalent entanglement classes and it turns out that there are two inequivalent classes of genuine multipartite entanglement. The first class can be represented by the so called GHZ state [67] given by

$$|\text{GHZ}\rangle = \frac{1}{\sqrt{2}}(|000\rangle + |111\rangle) \quad (1.91)$$

and the second class by the W state

$$|\text{W}\rangle = \frac{1}{\sqrt{3}}(|001\rangle + |010\rangle + |100\rangle). \quad (1.92)$$

The other classes are given by the three possible classes of bipartite entanglement, i.e.,  $AB|C, A|BC$  and  $AC|B$ , and the class of fully separable states. However, it turns out the

class of  $W$  states is much smaller than the class of GHZ states. Indeed, by observing that any three-qubit state  $|\psi\rangle$  can be transformed by means of LU operations into the form [68]

$$|\psi\rangle = \lambda_0|000\rangle + \lambda_1 e^{i\vartheta}|100\rangle + \lambda_2|101\rangle + \lambda_3|110\rangle + \lambda_4|111\rangle \quad (1.93)$$

with  $\lambda_j \geq 0$ ,  $\sum_j \lambda_j^2 = 1$  and  $\vartheta \in [0, \pi]$ , one can see that any state within the  $W$  class has to fulfill  $\lambda_4 = \vartheta = 0$ . As this yields a set of states which is of lower dimension than the set of all states, this shows that the  $W$  class is a set of measure zero in the set of all pure states.

The two classes are also different from a physical perspective. The GHZ state in Eq. (1.91) is maximally entangled and can be seen as a generalization of the Bell states of two qubits. In addition, this state plays an important role in the context of Bell inequalities [69], see Section 1.2.3. The entanglement of the  $W$  state turns out to be more robust against particle losses. Indeed, if one particle is lost in the GHZ state the resulting state  $\text{Tr}_A[|\text{GHZ}\rangle\langle\text{GHZ}|]$  equals the maximally mixed state of the remaining parties  $B$  and  $C$  and is thus separable. However, for the  $W$  state, the reduced density operator  $\text{Tr}_A[|W\rangle\langle W|]$  is entangled.

### The PPT criterion

The positive partial transpose (PPT) or Peres-Horodecki criterion is a necessary separability criterion which can be very efficiently evaluated. It states that if a bipartite state  $\rho$  given by

$$\rho = \sum_{jkmn} \rho_{kn}^{jm} |j\rangle\langle m| \otimes |k\rangle\langle n| \quad (1.94)$$

is separable, then the so-called partial transpose of  $\rho$

$$\rho^{TA} := \sum_{jkmn} \rho_{kn}^{mj} |j\rangle\langle m| \otimes |k\rangle\langle n| \quad (1.95)$$

will also be a valid density operator. It also guarantees positive semidefiniteness of  $\rho^{TB}$  what is defined in an analogous way. Consequently, entanglement of the state  $\rho$  can be detected whenever one of the partial transpositions of a state has a negative eigenvalue. In addition, it has the appeal that it is also sufficient for the case of qubit-qubit systems and qubit-qutrit systems. However, for larger systems there exist entangled states which do not violate the PPT criterion. Those states are called PPT entangled states. A classical example of such PPT entangled states that directly comes as a continuous family are the Horodecki states in  $\mathbb{C}^3 \otimes \mathbb{C}^3$  [70]. For  $\lambda \in [0, 1]$ , the states within this family can be written as

$$\rho_H(\lambda) = \frac{8\lambda}{8\lambda+1} \rho_E + \frac{1}{8\lambda+1} |\eta_\lambda\rangle\langle\eta_\lambda|, \quad (1.96)$$

where  $|\eta_\lambda\rangle = \frac{1}{\sqrt{2}}|2\rangle \otimes (\sqrt{1+\lambda}|0\rangle + \sqrt{1-\lambda}|2\rangle)$  and  $\varrho_E = \frac{1}{8}(3|\Phi_3\rangle\langle\Phi_3| + \mathbb{1} \otimes \mathbb{1} - (\sum_j |jj\rangle\langle jj|) - |20\rangle\langle 20|)$ . One can show that the family of states  $\varrho_H(\lambda)$  is PPT entangled for any  $\lambda \in (0, 1)$  [70, 71].

### Entanglement witnesses

This entanglement criterion is effectively based on the fact that the set of separable states offers a convex structure and is a subset of the set of all quantum states. The Hahn-Banach separation theorem then guarantees that for a given convex set and a point outside that set, one can always construct a continuous linear functional separating the point from the set. Consequently, an entangled state lying outside of the convex set of separable states can be detected by means of such a linear functional, also called witness. Any linear functional  $f : \mathcal{S}(\mathcal{H}) \rightarrow \mathbb{R}$  is of the form  $f(\varrho) = \text{Tr}[\mathcal{W}\varrho]$  for some operator  $\mathcal{W} \in \mathcal{B}_H(\mathcal{H})$ . Therefore, we call a hermitian operator  $\mathcal{W}$  an entanglement witness if

$$\text{Tr}[\mathcal{W}\varrho] \geq 0 \quad \text{for all separable states } \varrho \in \mathcal{S}(\mathcal{H}), \quad (1.97)$$

$$\text{Tr}[\mathcal{W}\varrho] < 0 \quad \text{for at least one entangled state } \varrho \in \mathcal{S}(\mathcal{H}). \quad (1.98)$$

One simple construction to obtain entanglement witnesses is by considering so-called projector-based witnesses. For a given pure entangled state  $|\psi\rangle$ , one makes an ansatz for  $\mathcal{W}$  of the form

$$\mathcal{W} = \lambda\mathbb{1} - |\psi\rangle\langle\psi|. \quad (1.99)$$

Now one sets

$$\lambda = \max_{\sigma \in \text{SEP}} \text{Tr}[\sigma|\psi\rangle\langle\psi|] = \max_{|\pi\rangle \in \text{SEP}} |\langle\pi|\psi\rangle|^2, \quad (1.100)$$

where the last equality in Eq. (1.100) follows from the fact that the maximum of a linear function on a convex set is always attained at one of the extreme points. As the extreme points of SEP are given by pure product states the claim follows. As a consequence, if  $\langle\mathcal{W}\rangle_\varrho < 0$  is measured, one can conclude that  $\varrho$  is entangled. There exist different approaches to construct entanglement witness, for instance those that are based on the PPT criterion.

### The CCNR criterion

We have already pointed out that, in contrast to the Schmidt decomposition for pure bipartite states, no necessary and sufficient criterion for entanglement is known which is based on the operator Schmidt decomposition. The computable cross norm or realignment (CCNR) criterion gives at least a necessary condition for the separability of



$\rho \in \mathcal{B}(\mathcal{H}_A \otimes \mathcal{H}_B)$ . Let

$$\rho = \sum_j \lambda_j G_j^A \otimes G_j^B \quad (1.101)$$

be the Schmidt decomposition as defined in Eq. (1.87). The CCNR criterion states that the operator Schmidt coefficients  $\{\lambda_j\}_j$  of a separable quantum state  $\rho$  fulfill

$$\sum_j \lambda_j \leq 1. \quad (1.102)$$

Consequently, if a state  $\rho$  yields  $\sum_j \lambda_j > 1$  one can conclude that it is entangled. Obviously, if the state is a pure product state, i.e.,  $\rho = |ab\rangle\langle ab|$  the operator Schmidt decomposition will already be given and one has  $\lambda_1 = 1$ , thus fulfilling the criterion.

For the proof of the CCNR criterion it is crucial to notice that the sum over the Schmidt coefficients of a state defines a norm on the set of positive semidefinite operators. Let us denote this norm by  $\|\cdot\|_{\text{CN}}$ . As a valid norm, it has to fulfill the triangle inequality and thus gives for any separable state  $\rho$

$$\|\rho\|_{\text{CN}} = \left\| \sum_j p_j |a_j b_j\rangle\langle a_j b_j| \right\|_{\text{CN}} \leq \sum_j p_j \| |a_j b_j\rangle\langle a_j b_j| \|_{\text{CN}} \leq 1. \quad (1.103)$$

### Majorization criterion

The majorization criterion establishes a connection between the entanglement properties of a state  $\rho$  and the eigenvalues of the reduced density operators  $\rho_A$  and  $\rho_B$ . For any given density operator  $\rho$  we denote by  $\lambda^\downarrow$  the vector of eigenvalues of  $\rho$  in decreasing order. The majorization criterion states that if the state  $\rho$  is separable, then

$$\begin{aligned} \sum_{j=1}^k \lambda_j^\downarrow(\rho) &\leq \sum_{j=1}^k \lambda_j^\downarrow(\rho_A), \\ \sum_{j=1}^k \lambda_j^\downarrow(\rho) &\leq \sum_{j=1}^k \lambda_j^\downarrow(\rho_B) \end{aligned} \quad (1.104)$$

will hold for all  $1 \leq k \leq d$ , where  $d$  is the dimension of the system. The particular type of ordering of the eigenvalues that appear in Eq. (1.104) is called majorization.

### Symmetric extension technique

All entanglement criteria introduced so far have not been able to detect the entanglement of every entangled state, even for the simplest, bipartite case. The symmetric extension technique provides a hierarchy of separability criteria, each of which can be efficiently solved and is in addition complete, i.e., any entangled state will be detected by some instance of the hierarchy. Suppose a bipartite mixed state  $\rho \in \mathcal{S}(\mathcal{H}_A \otimes \mathcal{H}_B)$  is

given. The idea is to consider multiple copies of the system  $\mathcal{H}_A$  and to derive necessary criteria that must be fulfilled if the state were separable. For this purpose, we call a state  $\tilde{\rho}_n \in \mathcal{S}(\mathcal{H}_A^{\otimes n} \otimes \mathcal{H}_B)$  a symmetric extension to  $n$  copies of  $\mathcal{H}_A$  if

$$\text{Tr}_{A_1, \dots, A_{n-1}}[\tilde{\rho}_n] = \rho, \quad (1.105)$$

$$\mathbb{F}\tilde{\rho}_n\mathbb{F} = \tilde{\rho}_n, \quad (1.106)$$

where  $\mathbb{F} : \mathcal{B}(\mathcal{H}_A^{\otimes n} \otimes \mathcal{H}_B) \rightarrow \mathcal{B}(\mathcal{H}_A^{\otimes n} \otimes \mathcal{H}_B)$  performs an arbitrary permutation of the subsystems of the tensor product space  $\mathcal{H}_A^{\otimes n}$ . Further, one will call  $\tilde{\rho}_n$  a PPT symmetric extension to  $n - 1$  copies of  $\mathcal{H}_A$  if the conditions in Eq. (1.105) and Eq. (1.106) are fulfilled and in addition  $\tilde{\rho}_n$  remains positive semidefinite under all possible partial transpositions. Indeed, if  $\rho$  is a separable state it will admit a decomposition of the form

$$\rho = \sum_j p_j |a_j\rangle\langle a_j| \otimes |b_j\rangle\langle b_j| \quad (1.107)$$

and thus the state

$$\tilde{\rho}_n = \sum_j p_j |a_j\rangle\langle a_j|^{\otimes n} \otimes |b_j\rangle\langle b_j| \quad (1.108)$$

is a valid PPT symmetric extension to  $n - 1$  copies of  $\mathcal{H}_A$  of the quantum state  $\rho$ . Therefore, one naturally obtains a countable infinite family of separability criteria. In addition, one can easily see that this family possesses a hierarchical structure, i.e., if  $\rho$  has a PPT symmetric extension to  $n$  copies of  $\mathcal{H}_A$ , then it will also have a PPT symmetric extension to  $n - 1$  copies of  $\mathcal{H}_A$ . Indeed, a simple candidate for such an extension given  $\tilde{\rho}_n$  would be  $\tilde{\rho}_{n-1} = \text{Tr}_A[\tilde{\rho}_n]$ , where  $A$  represents one of the copies of  $\mathcal{H}_A$ . Clearly,  $\tilde{\rho}_{n-1}$  will be symmetric with respect to the remaining copies of  $\mathcal{H}_A$ . Further, it follows from the properties of the partial trace that  $\tilde{\rho}_{n-1}$  is an extension of  $\rho$  to  $n - 1$  copies of  $\mathcal{H}_A$ . Consequently,  $\tilde{\rho}_{n-1}$  is a symmetric extension to  $n - 1$  copies. It remains to show that  $\tilde{\rho}_{n-1}$  is PPT with respect to any subset of parties. Assume the contrary, so there is a subset  $S$  such that  $\tilde{\rho}_{n-1}^{T_S}$  is not a valid quantum state. Then, there must be a negative eigenvalue with a corresponding eigenvector  $|v\rangle$ . If  $\{|j\rangle\}_j$  is a basis for the system  $\mathcal{H}_A$  which has been traced out, then  $\langle v | \langle j | \tilde{\rho}_{n-1}^{T_S} | v \rangle | j \rangle \geq 0$  for all  $1 \leq j \leq \dim(\mathcal{H}_A)$  as the state  $\tilde{\rho}_n$  is PPT. In particular this implies

$$\sum_j \langle v | \langle j | \tilde{\rho}_{n-1}^{T_S} | v \rangle | j \rangle = \langle v | \left( \sum_j \langle j | \tilde{\rho}_{n-1}^{T_S} | j \rangle \right) | v \rangle = \langle v | \text{Tr}_A[\tilde{\rho}_{n-1}^{T_S}] | v \rangle \geq 0. \quad (1.109)$$

As the partial transposition is originally performed on the state  $\tilde{\rho}_{n-1}$ , the set  $S$  cannot contain the system  $A$  which has been traced out. Therefore, one can commute the partial trace  $\text{Tr}_A$  and partial transposition over the subsystem  $S$ . By definition  $\tilde{\rho}_{n-1} = \text{Tr}_A[\tilde{\rho}_n]$  and one arrives at

$$\langle v | \tilde{\rho}_{n-1}^{T_S} | v \rangle = \langle v | \text{Tr}_A[\tilde{\rho}_n]^{T_S} | v \rangle = \langle v | \text{Tr}_A[\tilde{\rho}_n^{T_S}] | v \rangle \geq 0. \quad (1.110)$$

This yields a contradiction and thus  $\tilde{\rho}_{n-1}$  is a PPT symmetric extension to  $n - 1$  copies of  $\mathcal{H}_A$  of the quantum state  $\rho$ .

### Quantifying entanglement by measures

Apart from the approaches of characterizing and detecting entanglement, it is also often of interest to quantify it. This is particularly important in the context when entanglement is regarded as a resource. In addition, entanglement measures often offer an operational interpretation. Further, we have already seen that any attempt to characterize entanglement with respect to SLOCC results in a loss of a total ordering of quantum states, implying that in general a SLOCC based classification of entanglement would be extremely complicated. Indeed, the only systems where in principle a finite number of SLOCC orbits can arise are of the form  $\mathbb{C}^2 \otimes \mathbb{C}^{d_B} \otimes \mathbb{C}^{d_C}$ , thus containing at least one qubit [67]. An entanglement measure  $E$  is a mapping from the set of density matrices into the non-negative real numbers

$$E : \mathcal{S}(\mathcal{H}) \rightarrow \mathbb{R}_{\geq 0}, \quad (1.111)$$

such that all or some of the following properties hold.

- (1) Faithfulness:  $E(\rho) = 0$  if and only if  $\rho$  is separable.
- (2) Monotonicity: Entanglement cannot be created by means of LOCC transformations, that is,  $E(\Lambda(\rho)) \leq E(\rho)$  for any  $\Lambda \in \text{LOCC}$ .
- (3) LU-invariance: Local unitary operations cannot affect the amount of entanglement present in the system, that is,  $E(\rho) = E(U\rho U^\dagger)$  where  $U = U_1 \otimes \cdots \otimes U_n$ .
- (4) Convexity: Randomization of quantum states decreases the entanglement, that is,  $E(p\rho + (1-p)\sigma) \leq pE(\rho) + (1-p)E(\sigma)$  for all  $\rho, \sigma \in \mathcal{S}(\mathcal{H})$ .
- (5) Additivity: The entanglement present in two uncorrelated quantum systems equals the sum of the entanglement of both systems, that is,  $E(\rho \otimes \sigma) = E(\rho) + E(\sigma)$ .
- (5) Pure state reduction: For a pure state  $\rho = |\psi\rangle\langle\psi|$  the measure  $E$  reduces to the entropy of entanglement, that is,  $E(|\psi\rangle\langle\psi|) = (S \circ \text{Tr}_A)(|\psi\rangle\langle\psi|)$  for some subsystem  $A$  and  $S(\rho) = \text{Tr}[\rho \log(\rho)]$ .
- (6) Superadditivity: If  $\rho_{A_1 A_2 B_1 B_2}$  is a state on the system  $A_1 A_2 B_1 B_2$ , where Alice holds the system  $A_1 A_2$  and Bob  $B_1 B_2$ , the measure will satisfy

$$E(\rho_{A_1 A_2 B_1 B_2}) \geq E(\rho_{A_1 B_1}) + E(\rho_{A_2 B_2}). \quad (1.112)$$

The condition of faithfulness is very strong, as a faithful entanglement measure would yield a necessary and sufficient condition for separability. Hence, there are also entanglement measures that may not only vanish on the set of separable states but also on certain entangled subsets. On the other hand, the condition of monotonicity is often replaced by the stronger assumption that  $E$  is non increasing on average under LOCC. This means, that for a given quantum state  $\rho$  one has

$$E(\rho) \leq \sum_j p_j E \left( \frac{K_j \rho K_j^\dagger}{\text{Tr}[K_j \rho K_j^\dagger]} \right), \quad (1.113)$$

where  $\{K_j\}_j$  are the Kraus operators describing some LOCC protocol and the probability of observing outcome  $j$  is given by  $p_j = \text{Tr}[K_j \rho K_j^\dagger]$ . Note that the condition of monotonicity implies LU-invariance [72].

After all, mixed states are just describing our lack of knowledge about the exact behavior of a physical device. From this viewpoint, it seems natural to initially define entanglement measures on the set of pure states and extend them in a certain way to the set of mixed states. The so-called convex roof construction offers the possibility to extend entanglement measures from the pure to the mixed regime. If  $E = E(|\psi\rangle)$  is an entanglement measure for pure states, one defines for a mixed state  $\rho$

$$E(\rho) = \inf_{p_j, |\psi_j\rangle} \sum_j p_j E(|\psi_j\rangle), \quad (1.114)$$

where the infimum runs over all possible convex decompositions of  $\rho$  into pure states, that is,  $\rho = \sum_j p_j |\psi_j\rangle\langle\psi_j|$ . One advantage of the convex roof construction is that the extended entanglement measure  $E(\rho)$  is a convex function. However, an explicit computation of the convex roof in Eq. (1.114) is not an easy task, as the optimization runs over all convex decompositions of  $\rho$ . This difficulty is often paired with the hardness of evaluating the entanglement measure for pure states and typically one is more interested in computing lower bounds on  $E(\rho)$  [73].

In the following, we will discuss important entanglement measures and, if they were defined initially on pure states, their extension to mixed quantum states. We first give examples of entanglement measures for the bipartite case and then proceed with the multipartite case.

- (1) Concurrence: The concurrence is a bipartite entanglement measure which is originally defined for pure states by

$$\mathcal{C}(|\psi\rangle) = \sqrt{2(1 - \mathcal{P}(\text{Tr}_B(|\psi\rangle\langle\psi|))}. \quad (1.115)$$

For mixed states, this definition is extended via the convex roof construction. One advantage of  $\mathcal{C}$  is that for the case of a two-qubit system the convex roof

can be evaluated analytically and is solely determined by the eigenvalues of the related operator  $\tilde{\rho} = \sqrt{\sqrt{\tilde{\rho}}\sigma_2 \otimes \sigma_2\tilde{\rho} \otimes \sigma_2\sqrt{\tilde{\rho}}}$ . Indeed, one can show that [74]

$$\mathcal{C}(\rho) = \max \{0, \lambda_1 - \lambda_2 - \lambda_3 - \lambda_4\}, \quad (1.116)$$

where  $\lambda_j$  denote the ordered eigenvalues of  $\tilde{\rho}$ .

- (2) Entanglement of formation: This is a bipartite entanglement measure and is defined as the convex roof of the von Neumann entropy

$$E_F(\rho) := \inf_{(p_j, |\psi_j\rangle)} \sum_j p_j (S \circ \text{Tr}_A)(|\psi_j\rangle\langle\psi_j|). \quad (1.117)$$

Obviously, this measure fulfills the pure state reduction criterion. Further, it offers an operational interpretation: It yields the minimal number of Bell states that is required to build a single copy of the state [73]. For a long time it was an open problem whether  $E_F$  fulfills the requirement of additivity. It is known that additivity of  $E_F$  is an equivalent statement to superadditivity of  $E_F$  [66]. By relating the problem to a question about the production of entropy by quantum channels it was shown that  $E_F$  is indeed additive [75].

- (3) Entanglement cost: The idea of this measure is to relate entanglement with its usefulness in communication tasks. The entanglement cost yields the minimal rate of Bell states that have to be used to create many copies of  $\rho$  via LOCC [73]. Formally, it is given by

$$E_C(\rho) = \inf_{\text{LOCC}} \lim_{n_{\text{out}} \rightarrow \infty} \frac{n_{\text{in}}}{n_{\text{out}}}, \quad (1.118)$$

where  $n_{\text{in}}$  denotes the minimal number of singlets needed in order to create  $n_{\text{out}}$  copies of  $\rho$  given that arbitrary LOCC operations can be performed.

- (4) Distillable entanglement: Here one wants to quantify how many copies  $n_{\text{in}}$  of the given state  $\rho$  are needed in order to create  $n_{\text{out}}(\Phi_d)$  copies of the desired singlet state, where arbitrary LOCC operations can be performed. Formally, it is given by

$$E_D(\rho) = \inf_{\text{LOCC}} \lim_{n_{\text{out}} \rightarrow \infty} \frac{n_{\text{in}}}{n_{\text{out}}(\Phi_d)}. \quad (1.119)$$

- (5) Geometric measure: This measure is a proper multipartite entanglement measure in the sense that it is not simply a bipartite measure averaged over all possible bipartitions. It is initially defined for pure quantum states  $|\psi\rangle$  and measures how well a state can be approximated by means of product states  $|\pi\rangle$ . Formally, it is given by

$$G(|\psi\rangle) = 1 - \lambda^2(|\psi\rangle) \quad \text{with} \quad \lambda(|\psi\rangle) = \sup_{|\pi\rangle} |\langle\psi|\pi\rangle|. \quad (1.120)$$

The extension to mixed states is obtained via the convex roof construction according to Eq. (1.114). A comparison with Eq. (1.100) directly shows that the geometric measure relates to the construction of entanglement witnesses. Further, it turns out that it offers an operational interpretation in terms of multipartite state discrimination by means of LOCC [76]. It should be noted that in the literature there exists a different but equivalent definition of the geometric measure. Here one studies the quantity  $E_G(|\psi\rangle) = -2 \log(\lambda(|\psi\rangle))$ . In addition, for many families of important quantum states,  $G$  can be computed exactly and efficient algorithms are known to compute upper approximations. Also from a mathematical viewpoint this measure plays a distinguished role as it corresponds to the largest eigenvalue of the coefficient tensor.

- (6) Relative entropy of entanglement: The idea of this measure is to quantify the entanglement via its distance to the set of separable states. The measure is defined via

$$E_R(\rho) = \inf_{\sigma \in \text{SEP}} S(\rho||\sigma), \quad (1.121)$$

where  $S(\rho||\sigma) = \text{Tr}[\rho \log(\rho) - \rho \log(\sigma)]$  is the relative entropy.

- (7) Robustness of entanglement: This measure quantifies how much noise can be added to a given state  $\rho$  until it becomes separable. Formally it is given by

$$\mathcal{R}(\rho) = \inf \{t \geq 0 \mid \frac{1}{1+t}(\rho + t\sigma) \in \text{SEP} \text{ for } \sigma \in \text{SEP}\}. \quad (1.122)$$

Clearly, if  $\rho$  is separable, then  $\mathcal{R}(\rho) = 0$ .

### Monogamy of entanglement

We have already seen that multipartite entanglement offers a rich structure which is difficult to characterize. This fact is also expressed by the phenomenon of monogamy of entanglement. It states that if Alice and Bob are maximally entangled, i.e., they effectively share the state  $|\Phi_d\rangle$ , then they cannot be correlated at all with a third party Charlie. In a less extreme form, this means that there is a trade-off between the amount of entanglement between Alice and Bob and Alice and Charlie. It should be noted that this is a pure quantum effect and has no classical analogue. In the simplest case of three qubits the trade-off can be quantified via the so-called Coffman-Kundu-Wootters monogamy inequality which is given in terms of the concurrence as

$$\mathcal{C}^2(\text{Tr}_C[\rho]) + \mathcal{C}^2(\text{Tr}_B[\rho]) \leq \mathcal{C}^2(\rho_{A|BC}), \quad (1.123)$$

where  $\rho_{A|BC}$  denotes the bipartition of the three-qubit system according to  $A|BC$ . Further, it turns out that the monogamy relation in Eq. (1.123) can be extended to the case of  $n$ -qubits [77].

### 1.2.3 Bell nonlocality

So far, we have a dichotomy of quantum states: Either they are classical correlated or they are entangled. From this it is apparent that one regards entanglement as nonclassical correlations in quantum mechanics. However, so far we have no precise notion of what non-classicality means and different definitions may exist. One such notion was introduced by John Bell as a reaction to the EPR argument and is nowadays called Bell nonlocality. That quantum theory cannot be completed in the EPR sense under the natural assumptions of locality, realism and freedom of choice, is the content of Bell's theorem. Therefore, the experimental observation of a violation of a Bell inequality implies that *any* appropriate description of nature cannot rely on a local realistic theory.

#### Black-box formalism

We have seen that we extract information about the system under investigation by means of measurements. This allows us to learn the corresponding output distribution of a state with respect to that measurement. The aim of the black-box formalism is to formulate this setting in an abstract, theory-agnostic way which only keeps information about the structure of this information-gaining process. Any black-box experiment can be specified by three types of data. First, one has to specify the number of parties or players involved in the experiment to which we refer alphabetically as Alice (A), Bob (B), Charlie (C), etc. Second, each party has access to a certain number of measurement apparatuses among which they can choose in each round of the experiment. The particular choice is called the input. We write  $M_A$  for the number of inputs of Alice and proceed similarly for the other parties. Third, each input will cause an output which is an element of a given set. For simplicity, in this section, we label the outcomes of Alice by  $\{1, \dots, m_A\}$  and proceed similar for the other parties. Further, we assume that all parties have the same number of measurements and all measurements for each party have the same number of outcomes. Hence a scenario is specified by the number of players  $n_P$ , the number of inputs  $M_k$  and the number of outputs  $m_k$  where  $k$  runs over all parties.

#### The no-signaling polytope

To keep the discussion simple, we will focus here on the bipartite case  $n_P = 2$ . Black-box experiments of this kind can be fully characterized by the resulting probability distribution  $p(a, b|x, y)$ , which is also called a behavior. In order to be a proper probability distribution,  $p$  has to fulfill positivity and normalization constraints, that is,

$$p(a, b|x, y) \geq 0 \quad \forall a, b, x, y, \quad (1.124)$$

$$\sum_{a, b} p(a, b|x, y) = 1 \quad \forall x, y. \quad (1.125)$$

At this point,  $p(a, b|x, y)$  does not fulfill any constraint apart from being a valid probability distribution. Therefore,  $M_A M_B (m_A m_B - 1)$  real parameters are needed in order to fully specify  $p$ . Further, there are  $(m_A m_B)^{M_A M_B}$  possible deterministic distributions, i.e., distributions where for each pair of inputs the output on either side is completely determined.

However, if the two parties are space-like separated, then the behavior will be subjected to the so-called no-signaling constraints, which are imposed by special relativity. If the event of choosing the input at Alice's side is space-like separated from obtaining the output on Bob's side, then the choice of the input at Alice's side should not affect the marginal distribution on Bob's side. In particular, if Alice and Bob are space-like separated, the no-signaling constraints will prevent that Alice can use her black-box for instantaneous signaling. However, this possibility relies on the ability of Alice to choose freely among the set of possible inputs, corresponding to an encoding of a message that should be transferred. This assumption is called *freedom of choice* assumption. More formally [78], the no-signaling conditions are

$$p(a|x, y) := \sum_b p(a, b|x, y) = \sum_b p(a, b|x, \tilde{y}) = p(a|x, \tilde{y}) := p(a|x) \quad \forall a, x, y, \tilde{y}, \quad (1.126)$$

$$p(b|x, y) := \sum_a p(a, b|x, y) = \sum_a p(a, b|\tilde{x}, y) = p(b|\tilde{x}, y) := p(b|y) \quad \forall b, x, \tilde{x}, y. \quad (1.127)$$

Behaviors that do not fulfill the constraints in Eq. (1.126) and Eq. (1.127) are called signaling. All others are called no-signaling (NS). Clearly, the set of no-signaling correlations is bounded and constraint by linear inequalities. Therefore, it can be seen as the intersection of hyperplanes with the set of all probability distributions. Consequently, the set of NS correlations forms a polytope, called the NS polytope, and is denoted by  $\mathcal{P}_{\text{NS}}$ . Due to the constraints in Eq. (1.126) and Eq. (1.127) less parameters are needed to specify a NS correlation and they can be easily counted [79]. Indeed, the marginal distributions  $p(a|x)$  and  $p(b|y)$  can be chosen freely, yielding  $M_A(m_A - 1)$  and  $M_B(m_B - 1)$  free parameters. However, this does not fix the correlations between the both parties. For any choice of inputs  $x, y$  and a fixed outcome of Bob,  $b = \beta$ , the NS constraints give

$$p(a, \beta|x, y) = p(a|x) - \sum_{\substack{b=1 \\ b \neq \beta}}^{m_B} p(a, b|x, y). \quad (1.128)$$

Therefore, we can choose the parameters  $\{p(a, \beta|x, y)\}_a$  freely under the condition that they produce the correct marginal distributions, i.e.,  $\sum_a p(a, \beta|x, y) = p(\beta|y)$ . Hence, for fixed  $x, y$  and  $b = \beta$ , there are  $m_A - 1$  independent numbers. Consequently, one



arrives at [78]

$$D_{\text{NS}} = \dim(\mathcal{P}_{\text{NS}}) = M_A M_B (m_A - 1)(m_B - 1) + M_A (m_A - 1) + M_B (m_B - 1). \quad (1.129)$$

This way of organizing the necessary data for specifying the NS correlations reflects the structure of the so-called Collins-Gisin representation [80]. Here only the data of the marginals  $p(a|x)$  with  $a = 1, \dots, m_A - 1$ ,  $x = 1, \dots, M_A$  and  $p(b|y)$  with  $b = 1, \dots, m_B - 1$ ,  $y = 1, \dots, M_B$  as well as the components  $p(a, b|x, y)$  for  $a = 1, \dots, m_A - 1$ ,  $b = 1, \dots, m_B - 1$ ,  $x = 1, \dots, M_A$ ,  $y = 1, \dots, M_B$  are given. These components can be ordered in a correlation table. For example, for the case  $m_A = m_B = M_A = M_B = 2$  the table takes the form

$$p = \left( \begin{array}{c|cc} & p(a=1|x=1) & p(a=1|x=2) \\ \hline p(b=1|y=1) & p(1,1|1,1) & p(1,1|2,1) \\ p(b=1|y=2) & p(1,1|1,2) & p(1,1|2,2) \end{array} \right). \quad (1.130)$$

This way of organizing correlation data will become important in the context of representing Bell inequalities and appears as a crucial tool in Section 4.

### Local realism

Local realism can be seen as a concept that imposes additional restrictions on the set of NS correlations. It entails the idea of the EPR argument that quantum measurements reveal physical properties that are pre-determined, i.e., they have definite values regardless of whether they are measured or not. This is in line with the "classical" viewpoint that a preparation of a physical system should encode all knowledge about all possible subsequent measurements, that is, it should predict the outcomes of the measurements performed by Alice and Bob.

This translates to the black-box scenario as follows: In each round of the experiment, there exists a *hidden* description of a process  $\lambda$  such that each party's output is generated by only taking into account the same party's input. However, this process does not need to be deterministic. If one knows the process  $\lambda \in \Lambda$ , where  $\Lambda$  denotes the set of all possible processes, this means that there exist *response functions*  $p_A(a|x, \lambda)$  for Alice and  $p_B(b|y, \lambda)$  for Bob such that

$$p(a, b|x, y, \lambda) = p_A(a|x, \lambda)p_B(b|y, \lambda). \quad (1.131)$$

As the process is in general unknown and can in principle differ in each round, one assumes that  $\lambda$  is distributed according to some probability density  $\mu$  on  $\Lambda$ . Therefore, the set of all probability distributions allowed by this construction is given by

$$p(a, b|x, y) = \int_{\Lambda} \mu(\lambda) p_A(a|x, \lambda) p_B(b|y, \lambda) d\lambda. \quad (1.132)$$

A behavior that is of the form in Eq. (1.132) is called a local behavior and otherwise nonlocal. If a given behavior is local, one will also say that it admits a local hidden variable (LHV) model. Clearly, local behaviors fulfill the NS constraints and can be described by using at most the number of parameters that were needed to describe a NS behavior. Thus they can be casted in the Collins-Gisin form. Even though local behaviors do not need to be deterministic, the deterministic ones play a distinguished role. In a deterministic model, the local response functions  $p_A(a|x, \lambda)$  and  $p_B(b|y, \lambda)$  only take the values within  $\{0, 1\}$ , that is,

$$p(a|x, \lambda) = \delta(a, f_\lambda(x)), \quad p(b|y, \lambda) = \delta(b, g_\lambda(y)) \quad (1.133)$$

for certain functions  $f_\lambda, g_\lambda$ . Clearly, local deterministic models can be equivalently characterized by just giving a list of all outputs for all possible inputs, that is,

$$\lambda = \{a_1, \dots, a_{M_A}, b_1, \dots, b_{M_B}\}. \quad (1.134)$$

In this case, one would choose the functions  $f_\lambda(x) := a_x \in \lambda$  and  $g_\lambda(y) = b_y \in \lambda$ . Further, it immediately follows that the number of possible local deterministic strategies is given by  $m_A^{M_A} m_B^{M_B}$ . Among all  $(m_A m_B)^{M_A M_B}$  deterministic behaviors, these  $m_A^{M_A} m_B^{M_B}$  deterministic behaviors are the only ones that fulfill the locality condition [81]. All others can only be realized using signaling resources. More generally, the local deterministic behaviors remain extremal points of the NS polytope, but the NS polytope also has nonlocal extremal behaviors [81]. From a practical viewpoint, the characterization of local behaviors in Eq. (1.132) with stochastic response functions and a general probability density  $\mu$  is not amenable. In particular, it is not clear how one can prove that a given behavior  $p$  is of this form. Intuitively, it should be possible to absorb any kind of local randomness present in the response functions  $p_A$  and  $p_B$  into the shared random variable  $\lambda$ . Indeed, for given  $x$  and  $\lambda$ , the process of obtaining output  $a$  is a random variable  $A$ , which can be equivalently characterized by its cumulative distribution function  $F_A(a) = \text{Prob}[\{A \leq a\}] = \sum_{\tilde{a} \leq a} p(\tilde{a}|x, \lambda)$ . If we want to transfer the randomness of  $p_A$  into the hidden variable  $\lambda$ , we must extend  $\lambda$  by additional parameters which reflect this. To do so, introduce the new local parameter  $\lambda_A \in [0, 1]$  and define a new response function  $\tilde{p}_A = \tilde{p}_A(a|x, (\lambda, \lambda_A))$  which assigns an output  $a$  according to the deterministic rule

$$\tilde{p}(a|x, (\lambda, \lambda_A)) := \begin{cases} 1, & \text{if } F_A(a-1) \leq \lambda_A < F_A(a), \\ 0, & \text{otherwise.} \end{cases} \quad (1.135)$$

If  $\lambda_A$  is uniformly chosen among  $[0, 1]$ , one obtains

$$\int_0^1 \tilde{p}(a|x, (\lambda, \lambda_A)) d\lambda_A = \int_0^1 \mathbb{1}([F_A(a-1), F_A(a)]) d\lambda_A \quad (1.136)$$

$$= F_A(a) - F_A(a-1) = p(a|x, \lambda), \quad (1.137)$$

where  $\mathbb{1}$  denotes the indicator function. The same argument works for the response function  $p_B$  of Bob and introduces the additional local parameter  $\lambda_B$ . Thus, any local behavior can be seen as arising from a convex mixture of local deterministic behaviors. Further, given a local deterministic behavior, it fixes the output value for any possible input, that is,  $p(a, b|x, y) = \delta(a, f_j(x))\delta(b, g_k(y))$ . In particular, for fixed functions  $f_j$  and  $g_k$  the output assignment is fixed according to

$$(a_1 = f_j(1), \dots, a_{M_A} = f_j(M_A), b_1 = g_k(1), \dots, g_k(M_B)). \quad (1.138)$$

From this we can infer that any local behavior  $p$  can be obtained by marginalizing the joint distribution over all outcomes, as it decomposes into local deterministic ones. The previous discussion can be summarized in the following theorem, which is due to A. Fine [82]

**Theorem 2** ([82]). *For a given behavior  $p$  the following statements are equivalent.*

- (1)  $p$  is local.
- (2)  $p$  is a convex combination of local deterministic processes

$$p(a, b|x, y) = \sum_{j=1}^{d_A} \sum_{k=1}^{d_B} \mu_{j,k} \delta(a, f_j(x)) \delta(b, g_k(y)), \quad (1.139)$$

where  $d_A = m_A^{M_A}$  and  $d_B = m_B^{M_B}$ .

- (3) There exists a joint probability distribution  $\omega : \mathcal{A}^{\times M_A} \times \mathcal{B}^{\times M_B} \rightarrow [0, 1]$  such that each  $p(a, b|x, y)$  can be obtained by marginalizing  $\omega$ , that is,

$$p(a, b|x, y) = \sum_{\substack{a_j \\ j \neq x}} \sum_{\substack{b_k \\ k \neq y}} \omega(a_1, \dots, a_{M_A}, b_1, \dots, b_{M_B}), \quad (1.140)$$

where  $\mathcal{A}$  and  $\mathcal{B}$  denote the outcome space of Alice and Bob, respectively.

### Bell inequalities

In the following we will denote the set of all local behaviors for a given scenario by  $\mathcal{L}$  and refer to it as the local set. In simple terms, Bell inequalities are conditions that an arbitrary behavior must fulfill in order to belong to the local set  $\mathcal{L}$ . Clearly, the set  $\mathcal{L}$  is convex and bounded, thus compact. Further, it is clear from Theorem 2 that the extremal points of  $\mathcal{L}$  are exactly the deterministic behaviors. One can show [83] that the local set has the same dimension as the no-signaling polytope, i.e., they span the same affine space. Therefore, the local set  $\mathcal{L}$  is a polytope embedded into  $\mathbb{R}^{D_{NS}}$ . As the number of local deterministic behaviors is  $m_A^{M_A} m_B^{M_B}$  and thus larger than  $D_{NS}$ , some of the extreme points must be linearly dependent.

A Bell functional is a linear functional acting on the space  $\mathbb{R}^{D_{NS}}$  with the aim to separate some nonlocal behavior from the set of all LHV behaviors. Any Bell functional  $I$  acting on behaviors  $p = \{p(a, b|x, y)\}$  can be written as

$$I(p) := \sum_{a,b,x,y} c_{a,b,x,y} p(a, b|x, y). \quad (1.141)$$

As the local set  $\mathcal{L}$  is compact, there exists a  $I_L < \infty$  such that  $\max_{p \in \mathcal{L}} I(p) \leq I_L$ . The Bell functional  $I$  together with its local bound  $I_L$  is called a Bell inequality.

The simplest nontrivial Bell scenario corresponds to the case  $M_A = M_B = m_A = m_B = 2$  and can be completely characterized [81, 84]. The local polytope is embedded into  $\mathbb{R}^8$  and comes with  $2^4 = 16$  extreme points. From this, one can obtain that the complete polytope has 24 facets among which 16 correspond to positivity constraints. The remaining 8 are versions of the same inequality up to relabeling. This facet inequality is the so-called Clauser-Horne-Shimony-Holt (CHSH) inequality [6] and is given by

$$S = E_{0,0} + E_{0,1} + E_{1,0} - E_{1,1} \leq 2, \quad (1.142)$$

where  $E_{x,y} = p(a = b|x, y) - p(a \neq b|x, y)$ . The term  $E_{x,y}$  is also called a correlation coefficient. As the CHSH inequality can be expressed in terms of correlation coefficients, it is also called a *correlation inequality*. Note that the inequality in Eq. (1.142) does not depend on the particular labeling of the outputs. However, the CHSH inequality is often formulated in terms of expectation values. Here one chooses the concrete labels  $a, b \in \{\pm 1\}$  for the outputs. As the expectation value is given by  $\langle a_x, b_y \rangle = \sum_{a,b} ab p(a, b|x, y) = E_{x,y}$  one arrives at the equivalent form

$$S = \langle a_0 b_0 \rangle + \langle a_0 b_1 \rangle + \langle a_1 b_0 \rangle - \langle a_1 b_1 \rangle \leq 2. \quad (1.143)$$

### Symmetries and relabelings

A given Bell inequality can appear in multiple versions, yet performing the same test. Behind this degeneracy there are in principle two reasons.

**Freedom due to relabelings:** We have already seen that parties, measurement settings and outcomes simply appear as labels in order to guarantee the distinctness of different events. Therefore, the property of a linear functional to be a test for nonlocality should be preserved under a change of those labels. This results in three different kinds of relabelings. First, one can relabel the party's inputs. If  $\pi_A \in \mathfrak{S}(M_A)$  and  $\pi_B \in \mathfrak{S}(M_B)$  are permutations of the inputs of Alice and Bob, respectively, then

$$c_{a,b,x,y} \mapsto c_{a,b,\pi_A(x),\pi_B(y)} \quad (1.144)$$

defines another version of the same Bell inequality. Second, one can relabel the party's outputs. For a fixed choice of input labels, let  $\pi_A \in \mathfrak{S}(m_A)$  and  $\pi_B \in \mathfrak{S}(m_B)$  be permutations of the outputs of Alice and Bob, respectively. Then

$$c_{a,b,x,y} \mapsto c_{\pi_A(a),\pi_B(b),x,y} \quad (1.145)$$

defines another version of the same Bell inequality. The third type of relabeling is the permutation of the parties. If  $m_A = m_B$  and  $M_A = M_B$ , then

$$c_{a,b,x,y} \mapsto c_{b,a,y,x} \quad (1.146)$$

defines another version of the same Bell inequality. If one wants to check whether a given behavior  $p$  is local or not, all different versions of a Bell inequality  $I$  have to be taken into account. More precisely, one has to verify that  $I(p) \leq I_L$  for all versions of the Bell inequality  $I$ , where  $I_L$  denotes the local bound.

**Freedom due to constraints:** Bell inequalities are intended to act on correlations which respect the normalization constraints in Eq. (1.125) as well as the no-signaling constraints in Eq. (1.126) and Eq. (1.127). As long as the correlations  $\{p(a,b|x,y)\}$  satisfy these constraints, one can rewrite any Bell functional in an infinite number of ways. For instance, the positivity constraint  $p(0,0|0,0) \geq 0$  can be written in the form of a Bell functional with the coefficients  $c_{a,b,x,y} = -\delta_{a,0}\delta_{b,0}\delta_{x,0}\delta_{y,0}$ . However, due to the normalization constraints, the coefficients

$$\tilde{c}_{a,b,x,y} := c_{a,b,x,y} + \mu(\delta_{x,0} - \delta_{x,1}) \quad (1.147)$$

define the same inequality for any  $\mu \in \mathbb{R}$ . Further, by using the freedom introduced by the no-signaling constraints, one can obtain a well-known equivalent form of the CHSH inequality, which is called Clauser-Horne (CH) inequality [7]. The CH inequality is given by

$$S_{\text{CH}} = \sum_{x,y=0}^1 (-1)^{xy} p(0,0|x,y) - p_A(0|0) - p_B(0|0) \leq 0, \quad (1.148)$$

where  $p_A$  and  $p_B$  denote the marginal distributions for Alice and Bob, respectively. Further, using the NS constraints once more, that is, replacing  $p_A(0|0) = p(0,0|0,1) + p(0,1|0,1)$  and  $p_B(0|0) = p(0,0|1,0) + p(1,0|1,0)$ , the CH inequality in Eq. (1.148) can be written in the so-called Eberhard form [85]

$$S_{\text{E}} = p(0,0|0,0) - p(0,1|0,1) - p(1,0|1,0) - p(0,0|1,1) \leq 0, \quad (1.149)$$

which will become important in Section 4.

### The Popescu-Rohrlich box

As the set of NS correlations is a convex polytope, it can be characterized by means of its extreme points. As all deterministic NS behaviors are local, all nonlocal extremal NS behaviors must be nondeterministic. For the simplest setting  $m_A = m_B = M_A = M_B = 2$ , these extremal points can be constructed explicitly. Suppose we want to achieve the algebraic maximum of the CHSH inequality in Eq. (1.142). Clearly, as  $|E_{x,y}| \leq 1$ , this algebraic limit is  $S = 4$  and enforces that the correlation coefficients fulfill  $E_{0,0} = E_{1,0} = E_{0,1} = -E_{1,1} = 1$ . If the measurement settings and the outcomes of Alice and Bob are labeled by  $a, b, x, y \in \{0, 1\}$ , to achieve a value of  $S = 4$ , they have to fulfill the relation

$$a + b \bmod 2 = xy. \quad (1.150)$$

From this condition one can construct 16 deterministic behaviors that are, however, signaling. It turns out, that there exists exactly one convex combination of these behaviors which is contained in the NS polytope. This behavior, denoted by  $p_{\text{PR}}$ , is given by [86, 87]

$$p_{\text{PR}}(a, b|x, y) := \begin{cases} \frac{1}{2}, & a + b \bmod 2 = xy \\ 0, & \text{otherwise.} \end{cases} \quad (1.151)$$

The behavior  $p_{\text{PR}}$ , also called PR box, describes the following strategy: If Alice, Bob or both obtain the input 0, then they always obtain the same outcome, which is uniformly distributed, i.e., with probability  $\frac{1}{2}$  both obtain outcome 0 or outcome 1. If both parties receive input 1, then they will always obtain distinct outcomes. Also in this case, the local outcomes are uniformly distributed, i.e.,  $p_{\text{PR}}(0, 1|1, 1) = p_{\text{PR}}(1, 0|1, 1) = \frac{1}{2}$ .

### Quantum violations of local realism

The concepts of no-signaling and local realism do not refer to a particular physical theory but impose constraints on the observable correlations, independent of a concrete framework. Interestingly, if one assumes that quantum theory is the correct description of nature, i.e., all predictions of quantum theory can in principle be realized in reality, it turns out that local realism can be violated.

More formally, one assumes that Alice and Bob share a bipartite quantum state  $\rho \in \mathcal{B}(\mathcal{H}_A \otimes \mathcal{H}_B)$  and Alice implements generalized measurements  $\{E_{a|x}\}_{a,x}$  and Bob implements  $\{F_{b|y}\}_{b,y}$ . A behavior  $p$  is called a quantum behavior<sup>3</sup> if it can be realized

<sup>3</sup>If one prefers to take observables as the fundamental objects and thus consider the algebra of observables, this gives rise to another possible definition. Here the requirement that the measurement is of the form  $E_{a|x} \otimes F_{b|y}$  is replaced by  $[E_{a|x}, F_{b|y}] = 0$  where  $E_{a|x}, F_{b|y}$  act on the joint Hilbert space  $\mathcal{H}_A \otimes \mathcal{H}_B$ . The set of all quantum behaviors  $\tilde{\mathcal{Q}}$  would then be given by  $p(a, b|x, y) = \text{Tr}[\rho E_{a|x} F_{b|y}]$ . As one has  $[E_{a|x} \otimes \mathbb{1}, \mathbb{1} \otimes E_{b|y}] = 0$ , it directly follows that  $\mathcal{Q} \subset \tilde{\mathcal{Q}}$ . Although both sets coincide for the  $m_A = M_A = m_B = M_B = 2$  scenario [88], recently it turned out that they are in general different [89].

within quantum resources, i.e., there exist  $\varrho$ ,  $\{E_{a|x}\}_{a,x}$  and  $\{F_{b|y}\}_{b,y}$  such that

$$p(a, b|x, y) = \text{Tr}[\varrho E_{a|x} \otimes F_{b|y}] \quad \forall a, b, x, y. \quad (1.152)$$

If the scenario is fixed, that is,  $m_A, M_A$  and  $m_B, M_B$  are given, one can define the quantum set  $\mathcal{Q}$ , which contains all behaviors that can be realized using quantum resources. If the quantum state is fixed, one can define the set of all distributions that one can obtain by performing local measurements on  $\varrho$ . For this, denote by  $\mathcal{A}$  the set of *all* generalized measurements that Alice can perform on her system and define  $\mathcal{B}$  in a similar manner for Bob. Then

$$\mathcal{Q}(\varrho) := \left\{ \text{Tr}[\varrho E_{a|x} \otimes F_{b|y}] : \{E_{a|x}\}_{a,x} \in \mathcal{A}, \{F_{b|y}\}_{b,y} \in \mathcal{B} \right\}. \quad (1.153)$$

We call a quantum state  $\varrho$  local if it cannot violate any Bell inequality, or equivalently, if the set of derived behaviors  $\mathcal{Q}(\varrho)$  is contained in the local polytope  $\mathcal{L}$ . This particularly means that it is not sufficient to only inspect whether  $\varrho$  violates the Bell-inequalities for a specific scenario but one has to check for *all* possible scenarios. It is clear that separable quantum states can only generate local behaviors. Indeed, for any set of local measurements  $\{E_{a|x}\}_{a,x}, \{F_{b|y}\}_{b,y}$  one has

$$p(a, b|x, y) = \text{Tr}[\varrho E_{a|x} \otimes F_{b|y}] = \sum_{\lambda} p(\lambda) \text{Tr}[\varrho_{\lambda}^A E_{a|x}] \text{Tr}[\varrho_{\lambda}^B F_{b|y}]. \quad (1.154)$$

This is a LHV model where the local response functions are given by  $p_A(a|x, \lambda) = \text{Tr}[\varrho_{\lambda}^A E_{a|x}]$  and similar for Bob. This implies that entanglement is a necessary resource for a violation of a Bell inequality.

As the dimensions of  $\mathcal{H}_A$  and  $\mathcal{H}_B$  are not specified, the Naimark dilation theorem allows one to assume the measurements as projective and the state as pure. Clearly, the set of quantum behaviors is a subset of the non-signaling behaviors as the marginals of a quantum distribution  $p_Q \in \mathcal{Q}$  satisfies

$$\begin{aligned} p_Q(b|x, y) &= \sum_a \text{Tr}[\varrho F_{a|x} \otimes E_{b|y}] = \text{Tr}\left[\varrho\left(\sum_a F_{a|x}\right) \otimes E_{b|y}\right] \\ &= \text{Tr}[\varrho \mathbb{1} \otimes E_{b|y}] =: p_Q(b|y) \end{aligned} \quad (1.155)$$

for all possible choices of  $x$ . In addition, the set of quantum correlations  $\mathcal{Q}$  is convex. This can be seen as a consequence of the unboundedness of the dimension of the local systems. Indeed, assume that  $p_k \in \mathcal{Q}$  for  $1 \leq k \leq n$ , each with quantum realization  $\varrho_k$ ,  $\{E_{a|x}^{(k)}\}_{a,x}$  and  $\{F_{b|y}^{(k)}\}_{b,y}$ . If  $(\mu_k)_{k=1}^n$  is a probability distribution, define the new operators

$$\varrho := \bigoplus_{k=1}^n \mu_k \varrho_k, \quad E_{a|x} := \bigoplus_{k=1}^n E_{a|x}^{(k)}, \quad F_{b|y} := \bigoplus_{k=1}^n F_{b|y}^{(k)}. \quad (1.156)$$

It is clear that  $\varrho$  is a valid state in a larger Hilbert space and  $\{E_{a|x}\}_{a,x}$  and  $\{F_{b|y}\}_{b,y}$  are valid POVMs. The induced behavior is then given by  $p(a, b|x, y) = \sum_k \mu_k p_k(a, b|x, y)$ ,

which is a convex combination of the  $p_k$ . Therefore, the quantum set  $\mathcal{Q}$  is convex. Although the set is convex, one can show that it is not closed [90].

Further, the set of quantum behaviors is a proper subset of the non-signaling polytope. Indeed, we have already seen that

$$E_{00} + E_{01} + E_{10} - E_{11} \stackrel{\text{NS}}{\leq} 4, \quad (1.157)$$

where the maximal value is attained for the PR-box. Further we have seen that the maximal value of CHSH in any local theory is upper bounded by 2. It remains to clarify what the largest possible value of CHSH is if one considers quantum behaviors and whether this value allows for a discrimination between LHV models and quantum theory. Suppose that Alice and Bob share some quantum state  $\rho \in \mathcal{S}(\mathcal{H}_A \otimes \mathcal{H}_B)$ , where we can assume without loss of generality that the local dimensions are equal, i.e.,  $\mathcal{H}_A \cong \mathcal{H}_B = \mathcal{H}$ . The measurements of the two parties are  $\{E_{a|x}\}_{a,x}$  and  $\{F_{b|y}\}_{b,y}$  respectively, where  $x, y \in \{0, 1\}$  and  $a, b \in \{\pm 1\}$ . In order to be valid POVMs one needs  $0 \leq E_{a|x}, F_{b|y} \leq \mathbb{1}$ , implying  $\|E_{a|x}\|_\infty, \|F_{b|y}\|_\infty \leq 1$  with equality if and only if the measurements are projective. Within the framework of quantum theory the correlation coefficients  $E_{x,y}$  in the CHSH inequality take the form

$$E_{x,y} = \text{Tr} \left[ \rho (E_{1|x} - E_{-1|x}) \otimes (F_{1|y} - F_{-1|y}) \right] := \text{Tr} [\rho A_x \otimes B_y], \quad (1.158)$$

where  $A_x, B_y$  are hermitian operators whose spectrum is contained in the interval  $[-1, 1]$ . This allows us to interpret the value of the CHSH functional in the quantum state  $\rho$  as an expectation value of an observable, that is,

$$\langle S \rangle = \text{Tr}[\rho S], \quad \text{where } S = A_0 \otimes B_0 + A_0 \otimes B_1 + A_1 \otimes B_0 - A_1 \otimes B_1. \quad (1.159)$$

The operator  $S$  is called a Bell operator and the maximal value of the CHSH functional achievable in quantum theory can be upper bounded by the largest eigenvalue of the Bell operator  $S$ . This was first realized by Tsirelson who showed that a tight upper bound is given by  $2\sqrt{2}$  [78]. First, notice that it is sufficient to consider projective measurements, as we have no limit on the maximal dimension of  $\mathcal{H}$ . Therefore we have  $\|A_x\|_\infty = \|B_y\|_\infty = 1$  and  $A_x^2 = B_y^2 = \mathbb{1}$ . The square of the CHSH operator is then given by

$$S^2 = 4\mathbb{1} \otimes \mathbb{1} - [A_0, A_1] \otimes [B_0, B_1]. \quad (1.160)$$

Further, we have  $\|[A_0, A_1]\|_\infty \leq 2\|A_0\|_\infty \|A_1\|_\infty = 2$  and the same is true for the operators on Bob's side. Using that  $\|A \otimes B\|_\infty = \|A\|_\infty \|B\|_\infty$  for general operators and  $\|A^2\|_\infty = \|A\|_\infty^2$  for normal operators, one arrives at

$$\|S\|_\infty^2 = \|S^2\|_\infty \leq 4\|\mathbb{1} \otimes \mathbb{1}\|_\infty + \|[A_0, A_1]\|_\infty \|[B_0, B_1]\|_\infty \leq 8. \quad (1.161)$$



Consequently,  $\|S\|_\infty \leq \sqrt{8} = 2\sqrt{2}$ . The bound on the CHSH operator in Eq. (1.161) also highlights the role of the local measurements. Indeed, if  $A_0$  and  $A_1$  commute, the second term in Eq. (1.161) will vanish and thus no violation by *any* quantum state will be possible. In order to find a quantum system where we can achieve the maximal violation, consider a two-qubit maximally entangled state  $|\psi^-\rangle$  together with two projective measurements for each party. These measurements can be identified with unit vectors using the Bloch representation. More precisely, we have  $E_{a|x} = E_a(\vec{x}) = \frac{1}{2}(\mathbb{1} + a\vec{x}\vec{\sigma})$  and  $F_{b|y} = F_b(\vec{y}) = \frac{1}{2}(\mathbb{1} + b\vec{y}\vec{\sigma})$ , where  $\vec{x}, \vec{y}$  are unit vectors in  $\mathbb{R}^3$  and  $a, b \in \{\pm 1\}$ . The possible behaviors for this scenario are then given by

$$p(a, b|x, y) = p(a, b|\vec{x}, \vec{y}) = \text{Tr}[|\psi^-\rangle\langle\psi^-| E_a(\vec{x}) \otimes F_b(\vec{y})] = \frac{1}{4}(1 - ab\vec{x}\vec{y}). \quad (1.162)$$

It follows that the correlation coefficients  $E_{x,y}$  in Eq. (1.142) take the form  $E_{x,y} = \vec{x}\vec{y}$ . In total, the value of the CHSH inequality in the state  $|\psi^-\rangle$  is given by

$$S(|\psi^-\rangle) = \vec{x}_0(\vec{y}_0 + \vec{y}_1) + \vec{x}_1(\vec{y}_0 - \vec{y}_1). \quad (1.163)$$

From this it follows that the measurement directions  $\vec{x}_0, \vec{x}_1, \vec{y}_0, \vec{y}_1$  for a maximal violation have to be chosen as follows. While  $\vec{x}_0$  can be arbitrary, we need  $\vec{x}_0\vec{x}_1 = 0$ , what fixes the directions of Alice. The measurements of Bob follow from  $\vec{x}_0$  and  $\vec{x}_1$  via  $\vec{y}_0 = \frac{1}{\sqrt{2}}(\vec{x}_0 + \vec{x}_1)$  and  $\vec{y}_1 = \frac{1}{\sqrt{2}}(\vec{x}_0 - \vec{x}_1)$  [81]. For instance, on the level of operators one can choose

$$A_0 = \sigma_1, \quad A_1 = \sigma_3, \quad B_0 = \frac{1}{\sqrt{2}}(\sigma_1 + \sigma_3), \quad B_1 = \frac{1}{\sqrt{2}}(\sigma_1 - \sigma_3). \quad (1.164)$$

With similar arguments and using the Bloch representation, one can derive a closed expression for the value of the CHSH inequality with respect to any two-qubit mixed state and projective measurements [91]. For a given state  $\rho$ , we can compute the  $3 \times 3$  matrix  $T_{jk} := \text{Tr}[\rho\sigma_j \otimes \sigma_k]$ . The matrix  $T^T T$  is a positive hermitian matrix with eigenvalues  $\lambda_1 \geq \lambda_2 \geq \lambda_3$ . Then, one can show [91] that the largest violation of CHSH with respect to projective measurements is  $2\sqrt{\lambda_1 + \lambda_2}$ . This result has an important consequence for general pure qubit-qubit states, as any two-qubit pure state is up to LU transformations of the form

$$|\psi_\theta\rangle = \cos(\theta)|00\rangle + \sin(\theta)|11\rangle, \quad (1.165)$$

where  $\theta \in [0, \pi/4]$ . For a state of the form in Eq. (1.165) the matrix  $T$  is diagonal and given by  $T = \text{diag}(\sin(2\theta), -\sin(2\theta), 1)$ , hence  $2\sqrt{\lambda_1 + \lambda_2} = 2\sqrt{1 + \sin^2(2\theta)}$ . In particular, this expression is always larger than 2 unless  $\sin(2\theta) = 0$ , which only happens for  $\theta = 0$ . From this we can conclude that *any* pure entangled two-qubit state can violate the CHSH inequality.

### The GHZ argument

As already pointed out, in order to prove that a given behavior does not belong to the local set  $\mathcal{L}$  it is sufficient to show that a Bell inequality can be violated. However, in certain cases it is possible to obtain a direct, logical contradiction between the predictions of quantum theory and those resulting from a local model. Such demonstrations do not involve any inequality. The Greenberger-Horne-Zeilinger (GHZ) argument [92,93] involves three parties, where each has access to two measurements, labeled by 1,2, yielding two outcomes, labeled by  $\pm 1$ . Now one assumes that the three parties observe the following correlations

$$\langle a_2 b_1 c_1 \rangle = \langle a_1 b_2 c_1 \rangle = \langle a_1 b_1 c_2 \rangle = 1. \quad (1.166)$$

In particular, as the random variable  $a_2 b_1 c_1 \in \{\pm 1\}$ , this implies that *in each* run of the experiment where the setting  $a_2 b_1 c_1$  was chosen, the parties have observed  $a_2 b_1 c_1 = 1$ . If local realism holds, then the outputs are predetermined for all measurement settings, regardless of whether they have been measured or not. Therefore, for all measurement settings we have a valid output assignment, that is,

$$a_2 b_1 c_1 = a_1 b_2 c_1 = a_1 b_1 c_2 = 1 \quad (1.167)$$

in each single round. Further, as  $a_1^2 = b_1^2 = c_1^2 = 1$ , multiplying the assignments in Eq. (1.167) yields

$$a_2 b_2 c_2 = 1, \quad \text{thus} \quad \langle a_2 b_2 c_2 \rangle = 1. \quad (1.168)$$

Now consider the three-partite state

$$|\text{GHZ}\rangle = \frac{1}{\sqrt{2}}(|000\rangle + |111\rangle) \quad (1.169)$$

and the measurements  $\sigma_1 \otimes \sigma_2 \otimes \sigma_2$ ,  $\sigma_2 \otimes \sigma_1 \otimes \sigma_2$  and  $\sigma_2 \otimes \sigma_2 \otimes \sigma_1$ . It directly follows that Eq. (1.166) is fulfilled, while  $\langle \sigma_1 \otimes \sigma_1 \otimes \sigma_1 \rangle = -1$ . One should note that the GHZ argument relies on the observation of perfect correlations or anticorrelations.

### Hardy's test

Hardy's test is a Bell test which involves only two parties and is similar to the GHZ argument based on the presence of extreme correlations in the sense that they are deterministic, i.e., equals 0 or 1 and can be achieved in quantum theory. The setting is equal to the CHSH setting such that each party has two inputs 1,2 and two outputs  $\pm 1$ . One now enforces that

$$p(1,1|1,1) = p(-1,1|1,2) = p(1,-1|2,1) = 0. \quad (1.170)$$

Thus, one forbids the occurrence of events like  $(1,1|1,1)$  by construction. However, as the observables are dichotomic, the "complementary" event must happen with certainty, that is,

$$b_2 = 1 \Rightarrow a_1 = 1, \quad a_1 = 1 \Rightarrow b_1 = -1, \quad b_1 = -1 \Rightarrow a_2 = -1. \quad (1.171)$$

In particular, under the assumption of local realism, one can take the implications in Eq. (1.171) together in order to obtain a chain of implications resulting in

$$b_2 = 1 \Rightarrow a_2 = -1. \quad (1.172)$$

However, when these conditions are applied to quantum theory, one obtains a contradiction. More precisely, there exist quantum states, the so-called Hardy states, and measurement settings such that the conditions in Eq. (1.170) hold while  $p(1,1|2,2) > 0$ . The maximal value that can be achieved by using quantum theory is given by  $\frac{1}{2}(5\sqrt{5} - 11) \approx 0.09$  [81]. This means that Hardy's argument works only for 9% of the runs of a certain experiment. Interestingly, the set of suitable states does not contain maximally entangled states.

#### 1.2.4 Quantum steering

The EPR argument relies on the ability of Alice to predict with certainty the measurement outcome of certain observables of the other party, given that the same measurement is performed. From this they concluded that this nonlocality must be a consequence of the incompleteness of quantum theory. Schrödinger observed [94] that, by choosing her measurement direction, Alice can steer the other side into an eigenstate of  $\sigma_1$  or  $\sigma_2$ . This steering of Bob's wave function by Alice is in Schrödinger's own words *magic*, as in this case, Bob has to believe that Alice can influence his particle from a distance. Here it is important to notice that no information can be conveyed by Alice's measurement choice as the reduced state of Bob is independent of that choice. However, unlike EPR, he believed that the quantum state is a correct and complete description for localized, isolated systems [95]. On the other hand, similar to EPR, he could not easily accept the nonlocality of quantum theory. He suggested that quantum theory is incorrect with respect to the description of delocalized, entangled states. In particular, he believed that the quantum system held by Bob has a definite state. Therefore, one can assign a state to Bob's system which is independent of the measurement choice of Alice, i.e., one can speak about the state held by Bob [95].

#### Formalizing the EPR argument

In 2007, steering was formulated in the language of quantum information, that is, it was formulated with respect to a task [95]. Suppose that Alice can prepare a bipartite

quantum state  $\rho$  and sends one part of the state to Bob. She aims to prove to Bob that she is indeed capable of preparing an entangled state, while Bob is skeptical. This process is repeated many times and we refer to each realization as a round of the experiment. In each round, Alice has access to a set of generalized measurements, labeled by  $x$ . Upon request by Bob, she performs one among of those and obtains an outcome  $a$ . For each setting  $x$  and outcome  $a$ , Bob remains with a conditional state, which is described by the unnormalized density operator  $\rho_{a|x}$ . Further, one assumes that Bob believes that quantum theory yields the correct description of his local particle, i.e., he uses the Born rule to compute probabilities of measurement outcomes. Hence he has the ability to perform quantum state tomography on his particle in order to obtain a classical description of  $\rho_{a|x}$ . Here it is important to notice that Bob has to know  $a|x$ . After many runs, Bob obtains a full description of the *assemblage*  $\{\rho_{a|x}\}_{a,x}$  of conditional states. Now Bob may try to explain the appearance of his states without assuming that the state shared with Alice was entangled. He supposes that his particle was initially in a state  $\sigma_\lambda$  unknown to him, which occurs with probability  $p(\lambda)$ . Then, Alice uses her knowledge of  $\lambda$  and announces outcome  $a$  according to some local response function  $p_A(a|x, \lambda)$ . The conditional states  $\rho_{a|x}$  that can be constructed in this manner are of the form

$$\rho_{a|x} = \int_{\Lambda} p(\lambda) p_A(a|x, \lambda) \sigma_\lambda \, d\lambda. \quad (1.173)$$

If a representation of Bob's conditional states of the form in Eq. (1.173) exists, Bob does not need to assume any kind of action at a distance to explain the appearance of  $\rho_{a|x}$  and Alice would have failed to convince Bob that she can prepare an entangled state. In this case, one also says that the state  $\rho$  is unsteerable or has a local hidden state (LHS) model. If such a LHS model does not exist, Bob must admit that Alice can steer the state by some action at distance. Thus, Alice would have convinced Bob that she can prepare an entangled state [95].

As an example consider the assemblage that Bob obtains in the case of the EPR argument, which is formed by  $(1/2)\{|0\rangle\langle 0|, |1\rangle\langle 1|, |x^+\rangle\langle x^+|, |x^-\rangle\langle x^-|\}$ . As the states appearing in this assemblage are rank-1, and thus cannot be mixtures of other states. Hence the hidden states  $\sigma_\lambda$  have to be proportional to the four conditional states. To cast this in the form of Eq. (1.173) one must have  $p(a|x, \lambda) = 1$  if  $\sigma_\lambda$  corresponds to  $\rho_{a|x}$ . Hence one has  $p(\lambda) = 1/2$  for all  $\lambda$ , implying that the distribution is not normalizable [96].

### Relation to Bell nonlocality and entanglement

The LHS model in Eq. (1.173) gives also rise to a class of bipartite correlations on the level of probability distributions. Indeed, if Bob performs quantum mechanical

measurements, one arrives at a hybrid model of the form

$$p(a, b|x, y) = \int_{\Lambda} p(\lambda) p_A(a|x, \lambda) \text{Tr}[F_{b|y} \sigma_{\lambda}] d\lambda. \quad (1.174)$$

If Alice and Bob share a separable state  $\varrho = \int_{\lambda} p(\lambda) \varrho_{\lambda}^A \otimes \varrho_{\lambda}^B d\lambda$  and each party can implement measurements  $E_{a|x}$  and  $F_{b|y}$  respectively, the observed correlations will be of the form

$$p(a, b|x, y) = \int_{\lambda} p(\lambda) \text{Tr}[\varrho_{\lambda}^A E_{a|x}] \text{Tr}[\varrho_{\lambda}^B F_{b|y}] d\lambda. \quad (1.175)$$

From Eq. (1.175) it is immediately clear that entanglement is a necessary resource in order to demonstrate steering. Further, Eq. (1.174) shows that a LHS model is a particular instance of a LHV model, where Bob's local response function  $p_B$  is given by  $p_B(b|y, \lambda) = \text{Tr}[\sigma_{\lambda} E_{b|y}]$ . Therefore, quantum steering relies on quantum correlations which are intermediate between entanglement and Bell nonlocality. More precisely, any state that violates a Bell inequality can be used for steering and any steerable state is entangled. It can be shown that these relations are strict [95]. An important observation is that any state admitting a LHS model automatically has a LHV model and is therefore Bell local.

### Detection of steering using SDP's

In order to detect steering it is important to decide whether an assemblage demonstrates steering or not, i.e., whether it can be explained by means of a LHS model. It follows from the definition of a LHS model that one has to check whether there exist quantum states  $\{\sigma_{\lambda}\}_{\lambda}$  and distributions  $p(\lambda)$ ,  $p_A(a|x, \lambda)$ , such that the assemblage is of the form in Eq. (1.173). This is in general a difficult problem, as the distributions could be continuous. However, in the case of a fixed and finite number of measurements and outcomes the problem becomes much simpler. Similar to Fine's theorem for LHV models, also for steering one can decompose the response function of Alice  $p_A(a|x, \lambda)$  into a finite number of deterministic distributions, i.e., into distributions that yield a fixed output for each measurement [97]. Again, one introduces a new variable  $\tilde{\lambda}$  such that  $p_A(a|x, \tilde{\lambda}) = \delta(a, \tilde{\lambda}(x))$  which can be identified with the corresponding string of outputs

$$\tilde{\lambda} = (\tilde{\lambda}(0), \dots, \tilde{\lambda}(M_A - 1)) = (a(x=0), \dots, a(x=M_A - 1)). \quad (1.176)$$

As there are  $m_A$  outcomes and  $M_A$  measurements, there are  $d_A := m_A^{M_A}$  possible deterministic assignments. Therefore, we can expand the response function  $p_A$  as

$$p_A(a|x, \lambda) = \sum_{j=1}^{d_A} p(\tilde{\lambda}_j|\lambda) \delta(a, \tilde{\lambda}_j(x)), \quad (1.177)$$

where  $p(\tilde{\lambda}_j|\lambda)$  is the weight of the deterministic distribution identified with  $\tilde{\lambda}_j$ . If one inserts Eq. (1.177) into Eq. (1.173) one obtains

$$\varrho_{a|x} = \sum_{j=1}^{d_A} \int_{\Lambda} p(\lambda) p(\tilde{\lambda}_j|\lambda) \delta(a, \tilde{\lambda}_j(x)) \sigma_\lambda \, d\lambda = \sum_{j=1}^{d_A} \delta(a, \tilde{\lambda}_j(x)) \tilde{\sigma}_j, \quad (1.178)$$

where  $\tilde{\sigma}_j = \int p(\lambda) p(\tilde{\lambda}_j|\lambda) \sigma_\lambda \, d\lambda$ . Here it is important to note that Eq. (1.178) only involves a finite number of distributions which are known for a fixed value of  $0 \leq j \leq M_A - 1$ . This is in contrast to the original definition in Eq. (1.173), where the number of local response functions could have been continuous. As we just absorbed the randomness of the response function  $p_A$  into the hidden states at Bob's side, it follows that

$$\sum_{j=1}^{d_A} \text{Tr}[\tilde{\sigma}_j] = \sum_j \int p(\lambda) p(\tilde{\lambda}_j|\lambda) \text{Tr}[\sigma_j] \, d\lambda = \sum_j \int p(\lambda) p(\tilde{\lambda}_j|\lambda) \, d\lambda = 1. \quad (1.179)$$

Taking Eq. (1.178) and Eq. (1.179) together allows us to formulate the question of the existence of a LHS model as a feasibility semidefinite program (SDP), that is,

$$\begin{aligned} &\text{given} && \{\varrho_{a|x}\}_{a,x} \\ &\text{find} && \{\sigma_\lambda\} \\ &\text{subject to} && \sum_j \delta(a, \lambda_j(x)) \sigma_j = \varrho_{a|x} \quad \forall a, x \\ &&& \sigma_j \geq 0 \quad \forall j. \end{aligned} \quad (1.180)$$

This feasibility SDP can be reformulated as an explicit convex optimization problem [98]. For this, one relaxes the constraint  $\sigma_j \geq 0$  to  $\sigma_j \geq \mu \mathbb{1}$ , where  $\mu \in \mathbb{R}$ . Hence, the feasibility problem is equivalent to

$$\begin{aligned} &\text{given} && \{\varrho_{a|x}\}_{a,x} \\ &\text{maximize} && \mu \\ &\text{subject to} && \sum_j \delta(a, \lambda_j(x)) \sigma_j = \varrho_{a|x} \quad \forall a, x \\ &&& \sigma_j \geq \mu \mathbb{1}, \quad 1 \leq j \leq d_A, \end{aligned} \quad (1.181)$$

where  $\mu < 0$  indicates that the assemblage demonstrates steering, while  $\mu \geq 0$  means that there exists a LHS model explaining the appearance of the conditional states  $\{\varrho_{a|x}\}_{a,x}$ . Apart from the fact that the explicit SDP in Eq. (1.181) can be more suitable from a computational perspective, it also allows for a direct application of the duality theory of SDPs. The dual program to the optimization problem in Eq. (1.181)

is given by

$$\begin{aligned}
& \text{given} && \{Q_{a|x}\}_{a,x} \\
& \text{minimize} && \text{Tr} \left[ \sum_{a,x} F_{a|x} Q_{a|x} \right] \\
& \text{subject to} && \sum_{a,x} F_{a|x} \delta(a, \lambda_j(x)) \geq 0 \quad \forall j \\
& && \sum_{a,x,j} \delta(a, \lambda_j(x)) \text{Tr}[F_{a|x}] = 1.
\end{aligned} \tag{1.182}$$

One can use this dual formulation for the construction of steering inequalities, i.e., linear functionals that witness the nonexistence of a LHS model. Indeed, if the given assemblage  $\{\tilde{Q}_{a|x}\}_{a,x}$  demonstrates steering, one obtains not necessarily positive, hermitian operators  $\{F_{a|x}\}_{a,x}$  such that

$$\sum_{a,x} \text{Tr} \left[ F_{a|x} Q_{a|x} \right] \geq \alpha \tag{1.183}$$

for all assemblages  $\{Q_{a|x}\}_{a,x}$  that admit a LHS model. Further for the given assemblage  $\{\tilde{Q}_{a|x}\}_{a,x}$  one has  $\sum_{a,x} \text{Tr} \left[ F_{a|x} \tilde{Q}_{a|x} \right] < \alpha$ , thus witnessing the steerability of the given assemblage.

### 1.2.5 The measurement problem revisited

In Section 1.1.4 we have introduced the thought experiment of Wigner's friend and its extension by Deutsch. Both scenarios allow in principle that each party can verify the state which they assign to the system, i.e., the friend can simply repeat his spin measurement of the  $z$  component and Wigner could make a Bell-type measurement on the joint system, consisting of the spin system and the friend's lab. Importantly, such a measurement does not disturb the information that a definite outcome has been observed by the friend. This particularly means that each observer is in possession of some kind of "fact". The Deutsch proposal additionally offers the possibility that Wigner obtains direct evidence that the friend perceives a definite outcome, which indicates the existence of the friend's "fact". This raises the question whether both facts can be considered to be *real*, in the sense that they coexist and can jointly be regarded as being objective properties [36].

The argumentation of Wigner as well as of Deutsch assume the validity of quantum theory and depend crucially on the prepared state. As it turns out, extending the scenario to two friends and two Wigners that are space-like separated allows one to argue about the objective status of the observations of Wigner and his friend solely based on the observed statistics of the two Wigners. This makes the status of the quantum state irrelevant. In particular, it allows to make statements about the objective status of the outcomes in a theory-independent manner.

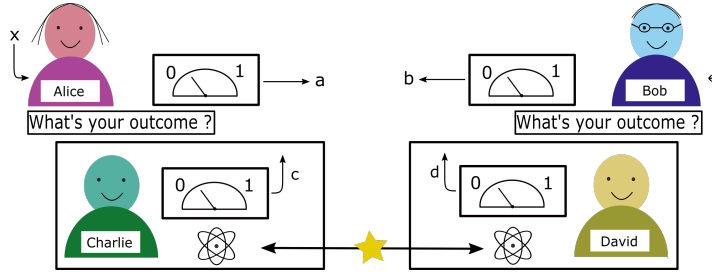


Figure 1.2: Illustration of the extended Wigner's friend scenario (EWFS). Here we assume that the measurements of each party can only yield two outcomes 0, 1.

### The extended Wigner's friend scenario

The extended Wigner's friend scenario (EWFS) was first introduced in Ref. [35] and formalized in Ref. [37]. It describes a bipartite version of the original Wigner's friend thought experiment. It involves two superobservers, named Alice and Bob, as well as two friends, named Charlie and David, see also Fig. 1.2. Each of the friends in their respective laboratories holds one particle from an entangled pair on which they can perform a measurement, yielding the outcome  $c$  and  $d$ , respectively. In each run of the experiment, Alice and Bob choose randomly and independently one of  $N \geq 2$  measurements, which they perform in space-like separated regions subsequent to a space-like hypersurface containing the measurements of the friends Charlie and Debbie [37]. As in the Bell scenario, the measurement choices for Alice and Bob are labeled by  $x, y \in \{1, \dots, N\}$ . Each measurement yields upon performance the corresponding output  $a, b$ . If the first measurement is chosen, i.e.,  $x = 1$ , Alice simply opens Charlie's laboratory and directly asks him for his observed outcome and then assigns her own output accordingly, that is, she sets  $a = c$ . Notice that the process of asking Charlie for his outcome is equivalent to the case where Alice makes herself a measurement on the particle hold by Charlie. If  $2 \leq x \leq N$ , Alice performs a different measurement on Charlie's laboratory as a whole. Bob and David proceed in a similar fashion.

Brukner [36] restricts in the EWFS to the case where  $x, y \in \{1, 2\}$ . The measurement  $A_1$  of Alice corresponds to a direct measurement of the spin system, i.e., asking the friend, and  $A_2$  corresponds to the outcome of a Bell-type measurement that is performed on the entire lab. The measurements  $B_1, B_2$  of Bob are defined in a similar manner. Brukner then argues that if both facts can be regarded as objective, then it should be possible to assign jointly truth values to both, the observable outcome of  $A_1$  as well as to  $A_2$ , independently of which measurement has actually been performed. This is formalized by the assumption of observer-independent facts.

**Definition 3** ([36]). *Observer-independent facts (OIF): The truth values of propositions  $A_i$*



of all observers form a Boolean algebra  $\mathcal{A}$ . Moreover, the algebra is equipped with a (countably additive) positive measure  $p(A) \geq 0$  for all statements  $A \in \mathcal{A}$ , which is the probability for the statement to be true.

If one combines the assumption of OIF with three further assumptions, one can derive the following no-go result.

**Theorem 4** ([36]). *The following statements are incompatible.*

- (1) *Universality of quantum theory, i.e., the predictions of quantum theory hold at any scale, even if the measured system contains objects as large as an observer.*
- (2) *Locality in the sense of parameter independence, i.e., the choice of the measurement settings of one observer has no influence on the outcomes of the other distant observers.*
- (3) *Freedom of choice, i.e., the choice of measurement settings is statistically independent from the rest of the experiment.*
- (4) *Observer-independent facts in the sense of Definition 3.*

### The absoluteness of observed events

The EWFS introduced by Brukner allows one to obtain Bell-type inequalities to probe under certain assumptions the coexistence of the outcomes of the friend and the outcomes of Wigner. However, the derived Bell-inequality, which coincides with the CHSH inequality, can be obtained from the assumptions of freedom of choice and Kochen-Specker non-contextuality [37, 99]. This makes the assumption of locality redundant and in addition does not require to consider the friend's observations. As it happens, the Kochen-Specker theorem [100] already shows that quantum theory does not admit a non-contextual model and from that perspective, Theorem 4 does not provide a novel result. In particular, this raises doubts on the implications of Theorem 4 with respect to the assumption about the objectivity of the friend's observation.

From the above discussion we see that the OIF assumption turns out to be too strong as the set of allowed correlations observed by Alice and Bob coincides with the set of LHV correlations. More generally, OIF entails that one can also assign truth values to statements about hypothetical measurements that were *not actually* performed. From this viewpoint, the apparent incompatibility of the assumptions, resulting from a violation of Brukner's inequality Eq. (1.143), could be resolved by maintaining that unperformed measurements have no results [37, 101]. Consequently, the idea would be to replace the set of assumptions of Brukner by a set of weaker assumptions that only involve statements about the nature of *observed outcomes*, i.e., about measurements that have been actually implemented. As it turns out, the notion of *absoluteness of observed events* (AOE) fulfills exactly that requirement.

**Definition 5** ([37]). *Absoluteness of observed events (AOE): An observed event is a real single event, and not relative to anything or anyone.*

In particular, this implies that in each run of the experiment, where Alice *has* implemented measurement  $x$  and Bob *has* implemented measurement  $y$ , there exists a well-defined value for the outcome observed by each party, i.e., for the values  $a, b, c, d$ . The crucial point is that, according to AOE, there will only exist a measurement outcome to the measurement  $x \geq 2$ , if it is actually performed.

Combining AOE with the assumptions of locality and freedom of choice yields a set of allowed correlations that is a proper superset of the correlations allowed by LHV, yet does not contain all quantum correlations. The conjunction of these three assumptions is called local friendliness (LF) [37].

**Theorem 6** ([37]). *If a superobserver can perform arbitrary quantum operations on an observer and its environment, then no physical theory can satisfy LF.*

## 1.3 Further concepts and applications

### 1.3.1 Graphs and graph states

Given a multi-particle quantum system, it is often possible to align the different constituents in a grid-based structure. If one also has information about the interaction pattern between the different particles, the configuration can be associated with a graph. In the following we will review important concepts from graph theory and also introduce the notion of a graph state. The results and definitions presented in this Section are covered in the book by Diestel [102] and in the review article by Hein [103].

#### Concepts from graph theory

Formally, a graph  $G$  is given by a pair  $G = (V, E)$  where  $V = \{1, \dots, n\}$  and  $E \subset V \times V$ . The elements of  $V$  are called vertices and the elements of  $E$  are called edges. In the following the term graph will refer to a simple graph, i.e., a graph without loops or multiple edges. The complement of a graph  $G$ , denoted by  $\bar{G}$ , is a graph on the same vertices  $V$  with edge set  $V \times V \setminus E$ . This means that two distinct vertices of  $\bar{G}$  will be adjacent if and only if they are not adjacent in  $G$ . It is clear that the number of possible graphs grows quickly with the number of vertices. Indeed, there are  $\binom{n}{2}$  different possibilities for choosing a set of edges  $E$  in a graph consisting of  $n$  vertices. Therefore, there are in total  $2^{\binom{n}{2}}$  different graphs. However, typically one wants to regard two graphs as being the same if they can be transformed into each other by permuting the vertices such that this permutation respects the neighborhood structure of the graph. Permutations of this kind are also called graph isomorphisms. More

formally, two graphs  $G_1 = (V_1, E_1)$  and  $G_2 = (V_2, E_2)$  will be called isomorphic, if there exists a bijective map  $\varphi : V_1 \rightarrow V_2$  such that

$$(v, w) \in E_1 \quad \text{if and only if} \quad (\varphi(v), \varphi(w)) \in E_2. \quad (1.184)$$

This allows one to reduce the number of graphs and only consider those that are not isomorphic. However, one can show that the set of nonisomorphic graphs still grows exponentially with the number of vertices [104]. An isomorphism from  $G$  to itself is called an automorphism of  $G$ . The set of all automorphisms of a graph forms a group, called the automorphism group of  $G$ . Two vertices  $v, w \in V$  that are the endpoints of an edge are called adjacent and we write  $v \sim w$  to indicate this relation. The neighborhood of a vertex  $v$  is given by all vertices that are adjacent, that is,  $N_v = \{w \in V \mid w \sim v\}$ . From the adjacency relation one can build the adjacency matrix  $\Gamma_G$  associated to the graph  $G$ , which is a  $|V| \times |V|$ -matrix with elements

$$\Gamma_{v,w} := \begin{cases} 1, & \text{if } v \sim w \\ 0, & \text{otherwise.} \end{cases} \quad (1.185)$$

The cardinality of the neighborhood of a vertex  $v \in V$ ,  $|N_v|$ , is called the degree of the vertex. A vertex of degree zero is called an isolated vertex and pairwise non-adjacent vertices are called independent. More generally, a set of vertices is independent if no two of its elements are adjacent. An ordered list of vertices  $v_1, \dots, v_n$  is called a  $(v_1, v_n)$ -path if  $(v_i, v_{i+1}) \in E$  for all  $1 \leq i \leq n-1$ . A connected graph is a graph that has a  $(v, w)$ -path for any pair  $v, w \in V$  and is called disconnected otherwise. A path in which only the first and last vertex coincide is called a cycle and a graph that consists of a single cycle is called circular graph. As already mentioned, most often one is interested in graphs up to an isomorphism. Consequently, a map that assigns equal values to isomorphic graphs is called a graph invariant. For instance, the number of vertices and the number of edges are obviously invariants of a graph. In order to introduce further graph invariants, we need more terminology. Given two graphs  $G_1 = (V_1, E_1)$  and  $G_2 = (V_2, E_2)$  one calls  $G_1$  a subgraph of  $G_2$  if  $V_1 \subset V_2$  and  $E_1 \subset E_2$ . In this case, we also write  $G_1 \subset G_2$ . If  $G_1 \subset G_2$  with  $G_1 \neq G_2$ , we also say that  $G_1$  is a proper subgraph of  $G_2$ . If  $G_1 \subset G_2$  such that  $G_1$  contains all edges  $(v, w) \in E_2$  with  $v, w \in V_1$ , then  $G_1$  is an induced subgraph of  $G_2$ . Further,  $G_1 \subset G_2$  is a spanning subgraph of  $G_2$  if  $V_1 = V_2$ . A graph with  $|V|$  vertices is called circulant if the cyclic permutation  $\sigma = (1, \dots, |V|)$  is a graph automorphism. This is equivalent to say that for a circulant graph the  $i$ -th vertex is connected to the  $(i-j)$ -th and the  $(i+j)$ -th vertex for each  $j \in S$  where  $S \subset V$ . Therefore, one can specify a circulant graph by specifying the number of vertices  $n$  and the list  $S$ . We will write  $Ci_n(S)$  to refer to such a graph. A complete graph is a graph in which every pair of distinct vertices is connected by an

edge. Clearly, a complete graph is uniquely specified by its number of vertices and we write  $K_n$  to denote the complete graph with  $n$  vertices.

**Definition 7.** Let  $G = (V, E)$  be a graph.

(1) A clique  $C$  of  $G$  is an induced subgraph that is complete. A maximum clique of  $G$  is a clique such that there is no clique with more vertices. The number of elements in a maximum clique is called the clique number of  $G$  and is denoted by  $\omega(G)$ .

(2) The fractional packing number of  $G$  is

$$\alpha^*(G) := \max \sum_{i \in V} p_i, \quad (1.186)$$

where the maximum is taken over all  $p_i \geq 0$  and for all cliques  $C$  of  $G$ , under the restriction  $\sum_{i \in C} p_i \leq 1$ .

(3) The independence number of  $G$ , denoted by  $\alpha$ , is the largest cardinality of any independent set of  $G$ .

(4) For a subset  $S \subset V$  of vertices we denote by

$$\Xi(S) := \max_{v \in S} |\{w \in S \mid w \sim v\}| \quad (1.187)$$

the highest degree of a vertex within  $S$  with respect to the neighborhood structure induced by  $S$ . For the particular choice of  $S = \mathcal{S}_{\alpha+1}$ , where  $\mathcal{S}_{\alpha+1}$  is the set of all subsets of  $(\alpha + 1)$  vertices of  $G$ , with  $\alpha$  the independence number of  $G$ , the xi number of  $G$  is given by

$$\Xi(G) := \min \{\Xi(S) \mid S \in \mathcal{S}_{\alpha+1}\}. \quad (1.188)$$

(5) The chromatic number of  $G$ , denoted by  $\chi$ , is the smallest number of colors needed to color the vertices of  $G$  such that no two adjacent vertices share the same color.

It is known that calculating the chromatic number  $\chi$  or the clique number  $\omega$  is a NP-complete problem [105]. Clearly, for the complete graph  $K_n$  one has  $\chi(K_n) = n$  and for a generic graph on  $n$  vertices a trivial bound is given by  $1 \leq \chi(G) \leq n$ . If the graph  $G$  contains a clique of size  $k$ , then  $k$  colors will be needed in order to color that clique. Therefore, the chromatic number is at least as large as the clique number, i.e.,  $\chi(G) \geq \omega(G)$ . The graph  $G$  is called perfect if  $\chi(G) = \omega(G)$ . Further, the clique number of a graph  $G$  is equal to the independence number of the complement graph  $\overline{G}$ , i.e.,  $\omega(G) = \alpha(\overline{G})$ . In addition, it follows directly from the definition that for a graph  $G = (V, E)$  and  $S_1, S_2$  with  $S_1 \subset S_2 \subset V$  one has  $\Xi(S_1) \leq \Xi(S_2)$ . In particular, this implies that for  $S \subset V$  containing no fewer than  $\alpha + 1$  vertices, where  $\alpha$  is the independence number of  $G$ , one has  $\Xi(S) \geq \Xi(G)$ . In addition,  $\Xi(G) \geq 1$  since there is at least one edge among any set of  $\alpha + 1$  vertices. For example, let us consider the complete graph  $K_3$  of three vertices, see also Fig. 1.3. Obviously, the independence

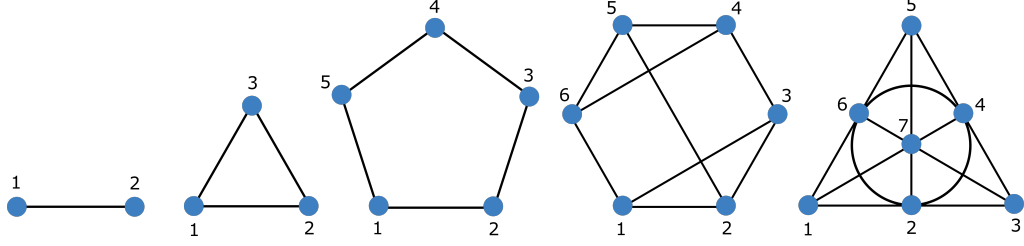


Figure 1.3: Examples of graphs for a small number of vertices. For instance, the first three graphs are  $K_2$ ,  $K_3 = \text{Xi}_3(\{1\})$  and  $C_5$ . From the perspective of the corresponding graph states, the first four graphs yield the known AME states on two/three/five/six qubits respectively. The last graph state, the Fano graph state [106], is a 2-uniform state on seven qubits where 32 of its 35 three-body marginals are maximally mixed. The figure is taken from Ref. [D].

number of this graph is  $\alpha = 1$  and therefore  $\mathcal{S}_2 = \{\{1,2\}, \{2,3\}, \{1,3\}\}$ . For the choice  $S = \{1,2\}$  we find  $\Xi(S) = \max\{|\{2\}|, |\{1\}|\}$ . As one finds similar results for all other sets in  $\mathcal{S}_2$ , we can conclude that  $\Xi(G) = 1$ . Notice that the complete graph on three vertices is also a circulant graph, that corresponds to  $\text{Ci}_3(\{1\})$ . A more advanced example is the circulant graph  $\text{Ci}_{10}(\{2,3\})$ , i.e., a graph with 10 vertices in which vertex  $i$  is adjacent to vertices  $i+1$  and  $i+3$ . Here one finds  $\Xi(G) = 2$ .

It is often the case that graphs which are highly symmetric are of particular interest. The amount of symmetry that a graph possesses is measured by means of its automorphism group. In this context, so-called vertex-transitive graphs play a distinguished role. A graph  $G$  will be called vertex-transitive if, for every pair of vertices, there exists an automorphism of the graph mapping one vertex to the other. More formally, this means that the automorphism group of the graph is transitive, i.e., the group orbit of one vertex equals the set of all vertices  $V$ . Consequently, every vertex of  $G$  has the same local environment and thus one cannot distinguish a vertex from another based on the vertices and edges surrounding it. Taking again the example of the circulant graph  $\text{Ci}_3(\{1\})$ , one directly sees that the automorphism group is generated by the cyclic permutations  $\sigma_1 = (1,2)$  and  $\sigma_2 = (2,3)$ . Therefore, for any pair of two vertices, there exists a graph automorphism mapping the one vertex to the other. Hence  $\text{Ci}_3(\{1\})$  is a vertex-transitive graph. In a similar manner, the circulant graph on 5 vertices  $\text{Ci}_5(\{1\})$  is a vertex-transitive graph, see Fig. 1.3. Its automorphism group is generated by the cycles  $\sigma_1 = (2,5)$ ,  $\sigma_2 = (3,4)$  and  $\sigma_3 = (1,2,3,4,5)$ .

**Definition 8.** Let  $G = (V, E)$  be a graph.

- (1) An orthonormal representation of  $G$  in  $\mathbb{C}^d$  is an assignment of a unit vector  $|v_j\rangle \in \mathbb{C}^d$  to each vertex  $j \in V$  satisfying that  $\langle v_j | v_k \rangle = 0$  for all pairs  $j, k \in V$  of adjacent vertices.

- (2) The orthogonal rank of  $G$ , denoted  $\xi$ , is the smallest positive integer  $d \geq 1$  for which there is an orthonormal representation in  $\mathbb{C}^d$  of  $G$ .

It is important to notice that the definition of an orthonormal representation neither requires that different vertices are assigned different vectors, nor that nonadjacent vertices correspond to nonorthogonal vectors. It is often the case that a further unit vector  $|\psi\rangle \in \mathbb{C}^d$  is specified together with the orthonormal representation. In this context, such a vector is called a handle. For instance, the graph  $\text{Ci}_3(\{1\})$  admits an orthonormal representation in  $\mathbb{C}^3$  by virtue of the assignment  $|v_1\rangle = \vec{e}_1$ ,  $|v_2\rangle = \vec{e}_2$  and  $|v_3\rangle = \vec{e}_3$ . Further, as this graph is fully connected, there cannot exist an orthonormal representation in  $\mathbb{C}^2$ . Hence the orthogonal rank of  $\text{Ci}_3(\{1\})$  is 3.

**Definition 9.** Let  $G = (V, E)$  be a graph. The Lovász number of  $G$  is

$$\vartheta(G) := \max \sum_{i \in V} |\langle \psi | v_i \rangle|^2, \quad (1.189)$$

where the maximum is taken over all orthonormal representations  $\{|v_j\rangle\}_{j \in V}$  of  $G$  and all handles  $|\psi\rangle$  in any dimension.

It can be shown that the calculation of the Lovász number of a graph  $G$  can be rephrased as a so-called semidefinite problem [107] and can thus be numerically approximated in time bounded by a polynomial in the number of vertices of  $G$ . Further, the so-called "sandwich theorem" [108] states that for a graph  $G$  the Lovász number  $\vartheta$  can be bounded via

$$\omega(G) \leq \vartheta(\overline{G}) \leq \chi(G). \quad (1.190)$$

This implies that for a perfect graph  $G$  one has  $\vartheta(\overline{G}) = \omega(G) = \chi(G)$ .

### Quantum states from graphs

We now want to associate a pure multipartite quantum state to a given graph  $G$ . The idea is that the vertices correspond to qubits and the edges represent an Ising-type interaction. In the literature, one can find two a priori different definitions of a graph state, which turn out to be equivalent. Given a graph  $G$  one can construct the corresponding graph state  $|G\rangle$  as follows. For each vertex  $v \in V$  one initializes the system in the state  $|+\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$ , which is the +1 eigenstate of the Pauli-X operator. Consequently, the whole system is initially in the pure state  $|+\rangle^{\otimes n}$ , where  $n = |V|$ . Then, for each edge  $(v_1, v_2) = e \in E$  one applies an Ising-type interaction between the qubits, represented by vertices  $v_1$  and  $v_2$ , which is given by the controlled-Z interaction CZ. The CZ interaction between two qubits is described by the unitary

operator

$$\text{CZ} = |00\rangle\langle 00| + |01\rangle\langle 01| + |10\rangle\langle 10| - |11\rangle\langle 11| = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}. \quad (1.191)$$

For a graph  $G$  with  $n$  vertices and  $e = (v_1, v_2)$  an edge, we denote by  $\text{CZ}_e$  the unitary operator acting on  $(\mathbb{C}^2)^{\otimes n}$ , which implements CZ on the qubits  $v_1$  and  $v_2$ , and is the identity on the other tensor factors. With this notation, we obtain the corresponding graph state as

$$|G\rangle = \prod_{e \in E} \text{CZ}_e |+\rangle^{\otimes n}. \quad (1.192)$$

Consequently, the graph encodes a summary of the interaction history of the particles which transforms the initial state  $|+\rangle^{\otimes n}$  to  $|G\rangle$ . From a physical viewpoint, graphs that contain isolated vertices represent a system where certain constituents are completely uncorrelated to the others. Therefore, the first nontrivial example of a graph state appears for the complete graph  $K_2$  with  $n = 2$  vertices, see also Fig. 1.3. In this case one has

$$|K_2\rangle = \text{CZ}_{12}|+\rangle|+\rangle = \frac{1}{2}(|00\rangle + |01\rangle + |10\rangle - |11\rangle) \quad (1.193)$$

$$= \frac{1}{\sqrt{2}}(|0\rangle|+\rangle + |1\rangle|-\rangle) \stackrel{\text{LU}}{=} \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle). \quad (1.194)$$

Hence this state is up to applying  $\mathbb{1} \otimes U$  for some unitary  $U$  equivalent to the maximally entangled state. The fact that the graph  $K_2$  is invariant under relabeling the vertices is mirrored by the graph state  $|K_2\rangle$  by being permutation symmetric. Let us consider now the graph state associated to the complete graph  $K_3$  with three vertices. Here one has

$$|K_3\rangle = \text{CZ}_{12}\text{CZ}_{23}\text{CZ}_{31}|+\rangle_1|+\rangle_2|+\rangle_3 \stackrel{\text{LU}}{=} |\text{GHZ}\rangle, \quad (1.195)$$

where  $|\text{GHZ}\rangle$  is defined in Eq. (1.91) as a representative of one of the two classes of genuine tripartite entangled states with respect to SLOCC. Here it is important to notice that the state  $|W\rangle$ , which is a representative of the other class of genuine tripartite entangled states, is not a graph state. Again, the fact that the graph  $K_3$  is invariant under relabeling the vertices is reflected by the permutation symmetry of the state  $|\text{GHZ}\rangle$ .

As already mentioned, there exists an equivalent definition of a graph state which is based on finding the common eigenstate to a specific set eigenvalues of a set of commuting operators. This directly relates to the so-called stabilizer formalism, which has a wide range of applications, for instance in quantum error correcting codes. For a

given graph  $G$  one associates to each vertex  $j \in V$  a correlation or stabilizing operator  $g_j$  defined by

$$g_j := X_j \bigotimes_{k \in N_j} Z_k, \quad (1.196)$$

where the index indicates on which qubit the corresponding operator acts. As an example, for the complete graph  $K_2$ , the stabilizing operators are given by  $g_1 = XZ$  and  $g_2 = ZX$ . Similarly, for  $K_3$  we obtain  $g_1 = XZZ$ ,  $g_2 = ZXZ$  and  $g_3 = ZZX$ . The graph state  $|G\rangle$  associated to the graph  $G$  is then the pure  $|V|$ -qubit state fulfilling

$$g_j |G\rangle = |G\rangle \quad \text{for all } j = 1, \dots, |V|. \quad (1.197)$$

In other words,  $|G\rangle$  is the common eigenstate to the eigenvalue  $+1$  of all stabilizing operators  $g_j$ . Clearly, if  $|G\rangle$  is a common eigenstate to the eigenvalue  $+1$  for all the  $g_j$ , then  $|G\rangle$  is also an eigenstate to the eigenvalue  $+1$  for all possible products of the  $g_j$ . The set of all products that can be formed from the stabilizing operators  $g_j$  forms a commutative subgroup of the Pauli group  $\mathcal{P}_n$ , which is the group generated by the  $n$ -fold tensor products of Pauli operators, i.e.,  $\mathcal{P}_n := \langle \{X, Y, Z\}^{\otimes n} \rangle$ . This subgroup is called the stabilizer of the graph state  $|G\rangle$ .

For a given graph  $G$  with associated graph state  $|G\rangle$  one can construct a basis for the space  $(\mathbb{C}^2)^{\otimes |V|}$  such that any basis element  $|G_\sigma\rangle$  is an eigenstate for each of the stabilizing operators  $g_j$ . In particular, the state  $|G_\sigma\rangle$  is an eigenstate of the stabilizing operator  $g_j$  with the eigenvalue  $(-1)^{\sigma_j}$ , where  $((-1)^{\sigma_1}, \dots, (-1)^{\sigma_{|V|}})$  is called the signature of the state  $|G_\sigma\rangle$ . More precisely, one has

$$g_j |G_\sigma\rangle = (-1)^{\sigma_j} |G_\sigma\rangle. \quad (1.198)$$

Indeed, consider for  $\sigma \in \{0, 1\}^{|V|}$  the state

$$|G_\sigma\rangle := \prod_j Z_j^{\sigma_j} |G\rangle. \quad (1.199)$$

Clearly,  $Z_k$  commutes with all stabilizing operators  $g_j$  as long as  $j \neq k$  and by the property in Eq. (1.16) it anticommutes with  $g_k$ . Therefore, for all  $k \in V$  we have

$$g_k |G_\sigma\rangle = g_k \prod_{j=1}^{|V|} Z_j^{\sigma_j} |G\rangle = (-1)^{\delta(\sigma_k, 1)} \prod_{j=1}^{|V|} Z_j^{\sigma_j} |G\rangle = (-1)^{\sigma_k} |G_\sigma\rangle. \quad (1.200)$$

Since there are  $2^{|V|}$  possible configurations for the eigenvalues  $(-1)^{\sigma_1}, \dots, (-1)^{\sigma_{|V|}}$  and  $\langle G_\sigma | G_{\tilde{\sigma}} \rangle = \delta_{\sigma, \tilde{\sigma}}$  holds, the pure states  $\{|G_\sigma\rangle\}_\sigma$  form an orthonormal basis for  $(\mathbb{C}^2)^{\otimes |V|}$ .

From this one obtains

$$\langle G_\sigma | \sum_j g_j |G_{\tilde{\sigma}}\rangle = 2^{|V|} \delta(\sigma, \vec{0}) \delta(\tilde{\sigma}, \vec{0}) \quad (1.201)$$



for any of the states  $|G_\sigma\rangle$  and  $|G_{\bar{\sigma}}\rangle$ . Therefore, one can express the graph state as the sum over the different elements of the stabilizer group

$$|G\rangle\langle G| = \frac{1}{2^{|V|}} \sum_{g \in \mathcal{S}} g. \quad (1.202)$$

Independent of the used definition, it is important to notice that different graphs can yield the same graph state in the sense that the associated states coincide up to local unitary transformations. Therefore, it would be desirable to have a criterion that determines whether two different graphs yield up to LU equivalence the same graph state. Although this problem is not fully resolved yet, when restricting to a subclass of operations, the so-called local Clifford operations, there exists a simple scheme called local complementation, which decides equivalence with respect to that class. Here it is important to notice that the technique of local complementation is directly applied to the graph, while LU equivalence refers to a property of the associated graph state  $|G\rangle$ . As already mentioned in Section 1.2.2, two pure states are LU equivalent if and only if they are LOCC equivalent. For the class of graph states one can make the even stronger statement that two graph states are LU equivalent if and only if they are SLOCC equivalent [109]. From this it is directly clear that in the case of two and three vertices there exists, up to LU equivalence, only one graph state.

Let us now explain the technique of local complementation in more detail. First, we need the notion of the Clifford group on one qubit, which is the set of all operators which map the Pauli group  $\mathcal{P}_1$  to itself under conjugation. The definition of the  $n$ -qubit local Clifford group is analogous. For a given graph  $G$  choose a vertex  $j \in V$ . The local complementation of  $G$  with respect to the vertex  $j$  is obtained by complementing the subgraph in  $G$  consisting of all vertices in its neighborhood  $N_j \subset V$  and their associated edges. To complement a subgraph means to erase all the edges in the subgraph, but instead connect the vertices which were originally disconnected. If two graphs belong to the same local complementation orbit, i.e., they can be transformed into each other by local complementation, their corresponding graph states will be in the same local Clifford orbit. As local Clifford equivalence is a particular instance of LU equivalence the graph states are also LU equivalent. However, this conclusion only works in one direction, i.e., there exist LU equivalent graph states that are not local Clifford equivalent [110].

### 1.3.2 Quantum computing

It has already been realized by the founders of information science that *information is physical* [23]. In this vein, the overall idea of quantum computing is to replace the classical bits and gates in classical computation by quantum systems and quantum operations. One of the key promises of quantum computers is to allow for new algorithms which can solve problems efficiently, while any known classical algorithm requires

exorbitant resources [29]. These problems range from the simulation of quantum systems [24] to purely mathematical problems, like the factorization of integers [30]. However, apart from these theoretical advantages, the quantum information, i.e., the physical system carrying the information, appears to be extremely fragile. In particular, the sensitivity to noise puts an ultimate time limit and size limit for any quantum computation, which may nullify any quantum speedup [111]. Although the theory of quantum error correction can in principle avoid that quantum information is corrupted, it comes at the cost of a large qubit overhead. This renders the application to the available noisy intermediate scale devices impossible [31]. In this sense, the field of quantum error mitigation can be understood as a temporary replacement of full error correction [32]. In the following, we introduce the concepts related to quantum computation that are needed for the subsequent chapters of this Thesis.

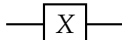
### Quantum gates

Classical and quantum computing relies on algorithms, which can be seen as a prescription for performing a particular task, which typically depends on a given input. Even though the processing of this input data during the computation can be arbitrarily complex, the operations that can be implemented are restricted to a very finite set of basic operations. These are the so called gates. Classically, the unit of information is the bit, taking values in  $\{0, 1\}$  and a logic gate is a function  $f : \{0, 1\}^m \rightarrow \{0, 1\}^n$ , where  $m$  refers to the number of inputs and  $n$  to the number of outputs. An example of such a logical gate is the NOT gate having one input and one output bit, turning 0 to 1 and 1 to 0, i.e.,  $f(x) = 1 \oplus x$  where  $\oplus$  is addition in  $\mathbb{Z}_2$ . Further important logical gates are the AND and the OR gate, taking two values as an input while only having one output bit. The AND gate acts on the inputs as  $(x, y) \mapsto xy$  while OR acts as  $(0, 0) \mapsto 0$ ,  $(0, 1) \mapsto 1$ ,  $(1, 0) \mapsto 1$  and  $(1, 1) \mapsto 1$ . Here it is important to notice that the action of AND and OR is not reversible, in contrast to the action of NOT.

In quantum computing we replace the classical bit by a two-level system, i.e., a qubit. According to the formalism of quantum theory we have to replace classical gates by unitary operations. The NOT gate directly translates to the quantum regime via an implementation of the Pauli X gate

$$X|0\rangle = |1\rangle, \quad X|1\rangle = |0\rangle, \quad (1.203)$$

which is represented in the circuit model as follows:

Pauli-X gate: 

Apart from the Pauli-X gate, also the corresponding Pauli-Y and Pauli-Z gates play an important role. Further, single qubit operations that will appear frequently are the Hadamard gate H, the phase gate S and the T gate (also called  $(\pi/8)$  gate). With

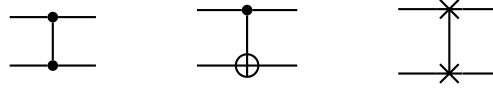
respect to the computational basis they can be written as

$$H = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}, \quad S = \begin{pmatrix} 1 & 0 \\ 0 & i \end{pmatrix}, \quad T = \begin{pmatrix} 1 & 0 \\ 0 & e^{i\frac{\pi}{4}} \end{pmatrix} \quad (1.204)$$

and they are denoted in the circuit representation in a similar manner as the Pauli-X gate. In addition, there are also the two-qubit gates

$$\text{CNOT} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \quad \text{and} \quad \text{SWAP} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \quad (1.205)$$

In the circuit representation, the gates CZ, CNOT and SWAP are denoted by



Here CZ and CNOT are controlled gates, as the action on the second qubit depends on the value of the first one. Indeed, the action of the CNOT gate is given by  $|x\rangle|y\rangle \mapsto |x\rangle|x \oplus y\rangle$ , i.e., if the controlled qubit is set to  $|1\rangle$ , then the target qubit (in the second register) will be flipped and otherwise there is no action on the target qubit at all. More generally, if  $U$  is a single qubit unitary operation, one can turn it into a controlled unitary operation, that is, if the control qubit is in the state  $|1\rangle$ , then  $U$  will be applied to the second register, and otherwise there will be no action. Finally, at the end of each quantum computation, there is an ideal projective measurement in the computational basis  $\{|0\rangle, |1\rangle\}$  on each qubit. By means of this measurement, we obtain a classical output string, whose length depends on the number of qubits being present at the end of the computation. Typically, in order to implement an algorithm, the composition of many elementary gates is needed and such a composition is called a circuit. Many powerful quantum algorithms like Shor's algorithm [30] or the quantum principle component analysis [112] require many gates for their implementation if the input grows. In general, the length of a quantum computation is expressed by the depth of a quantum circuit implementing this computation.

### Universal gate sets

In classical computation a logic gate is an arbitrary function  $f : \{0,1\}^m \rightarrow \{0,1\}^n$ . However, it turns out that a small set of gates is sufficient to compute an arbitrary function  $f$ . If this is the case, we call such a set of gates a *universal gate set*. For instance, for classical computation the set {AND, OR, NOT} is universal [29]. Here it is important to notice that universal gate sets are not uniquely defined and there exist other gate

sets that also turn out to be universal. A similar result can be obtained for quantum computation. However, as the set of operations is continuous the definition has to be adapted. One says that a set of gates will be universal for quantum computation, if any unitary operation may be approximated to an arbitrary accuracy by a quantum circuit involving only those gates. For example, it can be shown that the gate set  $\{\text{CNOT}, \text{H}, \text{T}\}$  is universal for quantum computation [113]. Further, also the gate set consisting of CNOT and *all* qubit operations turns out to be universal for quantum computation [29]. Even though these results ensure that one can restrict to a finite gate set, they are silent about the efficiency of the procedure of approximating a quantum circuit using that discrete set. Indeed, given the small coherence times on NISQ devices [31], it is important that the decomposition does not result in too deep circuits.

**Theorem 10** ([29, 114]). *Let  $S \subset SU(2)$  be a finite set, closed under taking the inverse such that the generated group  $\langle S \rangle$  is dense in  $SU(2)$ . Let the accuracy  $\epsilon > 0$  be given. Then the set  $S_n$  of all words of length  $n$  that can be built from elements of  $S$ , that is,*

$$S_n := \left\{ \prod_{j=1}^n s_{i_j} \mid s_{i_j} \in S \right\} \subset SU(2) \quad (1.206)$$

is an  $\epsilon$ -net in  $SU(2)$  if  $n = \mathcal{O}\left(\log\left(\frac{1}{\epsilon}\right)^c\right)$  for  $c \approx 2$ .

Thus Theorem 10 implies that an arbitrary single qubit gate can be approximated up to accuracy  $\epsilon$  using only  $\mathcal{O}\left(\log\left(\frac{1}{\epsilon}\right)^2\right)$  gates from a discrete set.

### Quantum error correction

One of the largest problems of practical quantum computing is its sensitivity to errors and noise. While also classical computation suffers from noise, e.g., bit flips, the situation for quantum computation is much more complicated. Classical error correction is based on encoding information in a redundant way such that even in the presence of noise the encoding process would recover the original data. For instance, suppose that a classical bit should be sent through a noisy channel, which flips the bit with a probability of  $p > 0$  while with a probability  $1 - p$  the bit arrives without error at the receiver. In order to make the communication more robust against noise, i.e., to decrease the probability that the wrong message is decoded, consider the scheme

$$0 \mapsto 000 \quad \text{and} \quad 1 \mapsto 111. \quad (1.207)$$

Here the bit strings 000 and 111 are also called logical 0 and logical 1, respectively and one says that the original message was encoded in the logical subspace. Now all three bits are sent through the communication channel and one assumes that each of the bits is affected independently by the noise. Therefore, the receiver probably obtains a noisy message that he has to decode. For this, majority voting is employed, that is, if

the message is 001 it will be decoded as 0. This simple error correction code is called a repetition code.

However, if one would like to adapt schemes of that kind for quantum computation one has to deal with at least three difficulties which directly relate to fundamental aspects of quantum theory.

- (1) **Quantum measurements:** While in classical error correction the particular bit can be read out at any state of the computation, quantum measurements disturb the quantum state and can destroy the information encoded into the state. However, in order to correct an error, it is necessary to detect it first.
- (2) **Continuous errors:** In classical computation errors can be modeled by bit flips, which yields a discrete set. However, due to the superposition principle of quantum theory, vastly different errors can occur for quantum computation, e.g. a phase flip error.
- (3) **No-cloning:** Encoding the information in a similar manner as in the repetition code is not possible as an unknown quantum state can not be copied. In addition, even the cloning of a quantum state would be allowed, it would not be possible to measure and compare the three quantum states from the output.

A first attempt to address these problems was presented by Peres, although the introduced "code" is not able to correct Pauli-Y errors [115]. We will outline the idea of this scheme by the so-called bit flip code [29]. Suppose that the system is initially in the state  $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$  with  $\alpha, \beta \in \mathbb{C}$ . This state is encoded into the three-qubit state  $|\tilde{\psi}\rangle = \alpha|000\rangle + \beta|111\rangle$ . The logical subspace in  $(\mathbb{C}^2)^{\otimes 3}$  is spanned by the logical states  $|0\rangle_L = |000\rangle$  and  $|1\rangle_L = |111\rangle$  and we refer to  $|0\rangle_L$  as the logical zero and to  $|1\rangle_L$  as the logical one. This three-qubit state is sent through a noisy quantum channel, acting on each qubit independently and performs with probability  $p$  a bit flip and acts as the identity with probability  $1 - p$ . Now suppose that, after the encoded system has passed through the channel, the bit flip occurred on at most one qubit. Then, one can devise a two-state error-correction procedure able to correct the state. In the first stage, we aim to *detect* the error. For this, we have to perform a measurement on the encoded state to obtain information which error has occurred. The outcome of the measurement is called the error syndrom and determines the operation that has to be performed in order to *correct* the error. Under the assumption that at most one of the qubit's was affected by the bit flip, there are four possible types of error: no error occurred, an error on the first qubit occurred, an error on the second qubit occurred and an error on the third qubit occurred. To each of these events, we assign a projection operator

$$\begin{aligned} \Pi_0 &= |000\rangle\langle 000| + |111\rangle\langle 111|, & \Pi_1 &= |100\rangle\langle 100| + |011\rangle\langle 011|, \\ \Pi_2 &= |010\rangle\langle 010| + |101\rangle\langle 101|, & \Pi_3 &= |001\rangle\langle 001| + |110\rangle\langle 110|. \end{aligned} \tag{1.208}$$

First, notice that all projection operators  $\Pi_0, \dots, \Pi_3$  commute and form a resolution of the identity. Therefore, if one wants to determine the error syndrom, it is possible to perform the measurement at once, i.e., a four-outcome projective measurement with effects  $(\Pi_0, \dots, \Pi_3)$  or sequentially, that is, one measures the two-outcome projective measurements  $(\Pi_j, 1 - \Pi_j)$  for  $0 \leq j \leq 3$ . Suppose that the bit flip occurred on the first system. The noise transforms the encoded state  $|\tilde{\psi}\rangle$  into the noise corrupted state  $|\tilde{\psi}_N\rangle = \alpha|100\rangle + \beta|011\rangle$ . If one now implements the syndrom measurement, i.e., the measurement of the four projectors in Eq. (1.208), one finds that  $\langle \tilde{\psi}_N | \Pi_1 | \tilde{\psi}_N \rangle = 1$ . This means that the outcome associated with  $\Pi_1$  appears with certainty. Because  $|\tilde{\psi}_N\rangle$  lives in the subspace which is spanned by  $\Pi_1$ , a measurement of  $\Pi_1$  does not alter the state of the system, i.e., the system is in the state  $|\tilde{\psi}_N\rangle$  before and *after* the measurement. Consequently, we can infer from the outcome of the measurement which particular error has occurred. Now we are in the position to *correct* the state. As we know that the first qubit was corrupted, one simply has to apply  $X \otimes 1 \otimes 1$  to recover the encoded state  $|\tilde{\psi}\rangle$ .

Although this simple example of a quantum error correcting code (QEC) does not take into account general and continuous errors, it can be seen as an evidence that all the difficulties quantumness brings into error-correction can be coped with.

### Quantum error mitigation

The theory of quantum error correction offers the possibility of fault tolerant quantum computation given that the errors in the components are smaller than a certain threshold  $p_{\text{th}}$ . However, besides the difficulty in achieving this particular threshold, the implementation of a fault-tolerant universal gate set with current codes introduces another challenge due to the large qubit overhead. Indeed, one can show that with the current technology a classically intractable computation requires hundreds of thousands of qubits [116]. Quantum error mitigation (QEM) takes a complementary approach of accepting hardware imperfections limiting the complexity of quantum algorithms and asks for how much information can be recovered by a purely classical post-processing of the data obtained from a noisy device. There is an important difference between QEC and QEM: While the former attempts to reduce the effect of noise on the output in every single circuit run, the latter only aims to reduce the effective damage for the whole ensemble of circuit runs.

## 1.4 Computational and algorithmic aspects

The question of whether a solution to a given problem is optimal with respect to a certain task is ubiquitous in science. Indeed, it appears in explicit form, e.g., finding the largest value of a function within a certain set or implicit via asking for the existence

of a solution with some desired properties. This turns the theory of optimization into a very powerful tool. Here we will introduce different classes of optimization techniques that are important in the subsequent chapters of this Thesis. The results and definitions presented in this Section are covered in the book of Nesterov [117] or Boyd and Vandenberghe [118].

### 1.4.1 General form and terminology

The most general formulation of an optimization problem we consider is given by

$$\begin{aligned} & \text{minimize} && f_0(\vec{x}) \\ & \text{subject to} && f_j(\vec{x}) \leq 0, \quad 1 \leq j \leq m \\ & && h_j(\vec{x}) = 0, \quad 1 \leq j \leq p. \end{aligned} \tag{1.209}$$

Consequently, Eq. (1.209) describes the task of finding a point  $\vec{x} \in \mathbb{R}^n$  that minimizes the *objective function*  $f_0 : \mathbb{R}^n \rightarrow \mathbb{R}$  among the set of all other  $\vec{x}$  that satisfy the conditions  $f_j(\vec{x}) \leq 0$  for  $1 \leq j \leq m$  and  $h_j(\vec{x}) = 0$  for  $1 \leq j \leq p$ . In this context we refer to  $\vec{x} \in \mathbb{R}^n$  as the optimization variable. Further, we call the functions  $f_j : \mathbb{R}^n \supset \text{dom}(f_j) \rightarrow \mathbb{R}$  the inequality constraint functions and  $h_j : \mathbb{R}^n \supset \text{dom}(h_j) \rightarrow \mathbb{R}$  the equality constraint functions. If no constraint functions are given, i.e.,  $m = p = 0$ , the optimization problem in Eq. (1.209) is called unconstrained. In order to give a meaning to Eq. (1.209), there should exist points which can simultaneously fulfill all the constraints. A necessary condition for such a point is that it is contained in the domain  $\mathcal{D} \subset \mathbb{R}^n$  of the optimization problem given by

$$\mathcal{D} := \bigcap_{j=1}^m \text{dom}(f_j) \cap \bigcap_{j=1}^p \text{dom}(h_j). \tag{1.210}$$

Further, we call  $\vec{x} \in \mathcal{D}$  feasible if it satisfies all the constraints  $f_j(\vec{x}) \leq 0$  as well as  $h_j(\vec{x}) = 0$ . In this case, the optimization problem in Eq. (1.209) is called feasible, otherwise unfeasible. The collection of all feasible points is called the feasible set and denotes by  $\mathcal{F}$ . A point  $\vec{x} \in \mathcal{F}$  will be called strictly feasible if  $f_j(\vec{x}) < 0$  for  $1 \leq j \leq m$ . Geometrically, this means that a strictly feasible point  $\vec{x} \in \mathbb{R}^n$  lies in the interior of  $\mathcal{F}$  with respect to the subspace topology that is induced on the affine hull of  $\mathcal{F}$ . To the problem in Eq. (1.209) one associates the optimal value  $p^*$  via  $p^* := \inf_{\vec{x} \in \mathcal{F}} f_0(\vec{x})$ . Typically one allows the optimal value  $p^*$  to take values in  $\mathbb{R} \cup \{\pm\infty\}$  and in the case that the problem is infeasible one sets  $p^* = \infty$ . In the case that there exists a sequence of feasible points  $(\vec{x}_k)_{k \geq 1} \subset \mathcal{F}$  such that  $f_0(\vec{x}_k) \rightarrow -\infty$ , one sets  $p^* = -\infty$  and says that the problem is unbounded from below. If  $f_0 \equiv 0$  one clearly has  $p^* \in \{0, \infty\}$ . Optimizing  $f_0 \equiv 0$  over a (probably empty) feasible set  $\mathcal{F}$  is also called a feasibility problem as it determines whether the constraints are consistent, and if so, finds a point that satisfies them.

### 1.4.2 Lagrange duality

For a given optimization problem of the form Eq. (1.209) assume that  $\mathcal{D} \neq \emptyset$  and denote by  $p^*$  its optimal value. The main idea of Lagrange duality is to incorporate the constraint functions  $f_j$  and  $h_j$  into the objective function  $f_0$ . More precisely, one associates the so-called Lagrangian  $\mathcal{L} : \mathbb{R}^{n \times m \times p} \rightarrow \mathbb{R}$  given as

$$\mathcal{L}(\vec{x}, \vec{\lambda}, \vec{\mu}) := f_0(\vec{x}) + \sum_{j=1}^m \lambda_j f_j(\vec{x}) + \sum_{j=1}^p \mu_j h_j(\vec{x}) \quad (1.211)$$

to the optimization problem. Here  $\vec{\lambda}$  and  $\vec{\mu}$  are called the Lagrange multipliers or dual variables associated with the constraints. The Lagrange dual function  $g : \mathbb{R}^{m \times p} \rightarrow \mathbb{R}$  is then the minimal value of the Lagrangian  $\mathcal{L}$  with respect to the variable  $\vec{x} \in \mathbb{R}^n$ ,

$$g(\vec{\lambda}, \vec{\mu}) := \inf_{\vec{x} \in \mathcal{D}} \mathcal{L}(\vec{x}, \vec{\lambda}, \vec{\mu}) = \inf_{\vec{x} \in \mathcal{D}} [f_0(\vec{x}) + \sum_j \lambda_j f_j(\vec{x}) + \sum_j \mu_j h_j(\vec{x})]. \quad (1.212)$$

Clearly, if  $\mathcal{L}$  is unbounded from below, then one has  $g = -\infty$ . Here it is important to notice that  $\mathcal{L}$  is an affine function in the variables  $\vec{\lambda}$  and  $\vec{\mu}$ . As the Lagrange dual function  $g$  is obtained as the pointwise infimum of the Lagrangian  $\mathcal{L}$ , it follows that  $g$  is a concave function, even the original problem is not convex. This by itself is already a pleasant property. This raises the question how the values of  $g$  relate to the optimal value  $p^*$ . It turns out that the dual function  $g$  yields lower bounds on the optimal value  $p^*$  of the original problem, which is by itself defined via a minimization. More precisely, for any choice of  $\vec{\lambda}$  and  $\vec{\mu}$  with  $\lambda_j \geq 0$  for all  $1 \leq j \leq m$  one has  $g(\vec{\lambda}, \vec{\mu}) \leq p^*$ . Indeed, if  $\vec{z} \in \mathcal{D}$  is a feasible point, it follows that

$$\mathcal{L}(\vec{z}, \vec{\lambda}, \vec{\mu}) = f_0(\vec{z}) + \sum_j \lambda_j f_j(\vec{z}) + \sum_j \mu_j h_j(\vec{z}) \leq f_0(\vec{z}), \quad (1.213)$$

as all functions  $h_j$  vanish for  $\vec{z}$  and  $f_j(\vec{z}) \leq 0$ . Consequently, we obtain for the dual function

$$g(\vec{\lambda}, \vec{\mu}) = \inf_{\vec{x} \in \mathcal{D}} \mathcal{L}(\vec{x}, \vec{\lambda}, \vec{\mu}) \leq \mathcal{L}(\vec{z}, \vec{\lambda}, \vec{\mu}) \leq f_0(\vec{z}). \quad (1.214)$$

As this argument works for any feasible point  $\vec{z}$ , one obtains that  $g(\vec{\lambda}, \vec{\mu})$  yields a lower bound on the minimal value that  $f_0$  can attain on the feasible set. Obviously, a necessary criterion on the bound  $g(\vec{\lambda}, \vec{\mu})$  that one obtains for  $p^*$  via Eq. (1.214) to be useful in practice is that  $g(\vec{\lambda}, \vec{\mu}) \neq -\infty$ . This means that the dual function can only yield nontrivial lower bounds when all  $\lambda_j \geq 0$  and  $(\vec{\lambda}, \vec{\mu}) \in \text{dom}(g)$ . If this is the case we call the pair  $(\vec{\lambda}, \vec{\mu})$  dual feasible. It is in the nature of the problem that we are not satisfied by obtaining some lower bound on  $p^*$ . We are interested in the optimal lower bound we can obtain via the dual function  $g(\vec{\lambda}, \vec{\mu})$ , which yield parameter dependent



bounds. This directly leads to another optimization problem

$$\begin{aligned}
& \text{maximize} && g(\vec{\lambda}, \vec{\mu}) \\
& \text{subject to} && \lambda_j \geq 0, \quad 1 \leq j \leq m \\
& \text{with respect to} && \vec{\lambda}, \vec{\mu}.
\end{aligned} \tag{1.215}$$

The optimization problem in Eq. (1.215) is called the Lagrange dual problem, or short, dual problem, associated with the original problem in Eq. (1.209). A pair  $(\vec{\lambda}^*, \vec{\mu}^*)$  is called dual optimal if it is optimal with respect to Eq. (1.215). The optimal value will be denoted by  $d^*$ . From the previous discussion it is clear that  $d^* \leq p^*$ . This property is called weak duality and the difference  $p^* - d^* \geq 0$  is called the duality gap of the original problem. If one finds that  $d^*$  and  $p^*$  coincide,  $d^* = p^*$ , one says that strong duality holds. Even though strong duality does not hold in general, there are important subclasses of optimization, e.g., convex or semi-definite programming, where it in (almost) all cases hold. Further, for these subclasses easy to evaluate conditions on the problem are known under which strong duality holds. These conditions are called constraint qualifications.

### 1.4.3 Convex optimization

In convex optimization one assumes that the objective function  $f_0$  as well as all constraint functions  $f_j, h_j$  are convex. Problems of this form turn out to have a wide field of applications and especially the subclass of semidefinite programs (SDPs) are of paramount importance in quantum information. For instance, they can be used to decide the separability of a multipartite quantum state [119], bounding the maximal violation of a Bell inequality in quantum theory [120], upper bounds on the distillable entanglement [121], relate to invariants of graphs [107] or can be used to decide the irreducibility of a positive-operator values measure [122].

#### The formulation of the problem

In convex optimization one is concerned with problems of the form

$$\begin{aligned}
& \text{minimize} && f_0(\vec{x}) \\
& \text{subject to} && f_j(\vec{x}) \leq 0, \quad 1 \leq j \leq m \\
& && A\vec{x} = \vec{b},
\end{aligned} \tag{1.216}$$

where  $f_j : \text{dom}(f_j) \subset \mathbb{R}^n \rightarrow \mathbb{R}$  convex and  $A \in \mathbb{R}^{p \times n}$ .

In this context, it turns out to be useful to introduce the concept of a  $\alpha$ -sublevel set  $S_\alpha(f)$  of a function  $f : \mathbb{R}^n \rightarrow \mathbb{R}$ . It is defined via

$$S_\alpha(f) := \{\vec{x} \in \mathbb{R}^n \mid f(\vec{x}) \leq \alpha\}. \tag{1.217}$$

The first important thing to notice is that if a function  $f$  is convex also the  $\alpha$ -sublevel sets are convex. Indeed, for  $\vec{x}, \vec{y} \in S_\alpha$  one directly verifies that for  $\vec{z} = \lambda\vec{x} + (1 - \lambda)\vec{y}$  one has  $f(\vec{z}) \leq \lambda f(\vec{x}) + (1 - \lambda)f(\vec{y}) \leq \alpha$ , thus  $\vec{z} \in S_\alpha$ . This implies that for a convex optimization problem the feasible set itself is convex as it arises from the intersection of the sublevel sets  $S_0(f_j)$  and the hyperplane given by  $\{\vec{x} \in \mathbb{R}^n \mid A\vec{x} - \vec{b} = 0\}$ . From this one can derive one of the most important and useful properties of convex optimization problems. It turns out, that any *local* optimal point is also a *global* optimal point. Let  $\mathcal{F}$  be the feasible set with respect to the constraints in Eq. (1.216),  $\vec{x} \in \mathcal{F}$  a local optimal point, that is  $f_0(\vec{x}) = \min \{f_0(\vec{z}) : \|\vec{z} - \vec{x}\|_2 \leq \delta, \vec{z} \in \mathcal{F}\}$  for some  $\delta > 0$ . Now assume that  $\vec{x}$  is not globally optimal, i.e., there must exist a feasible point  $\vec{y} \in \mathcal{F}$  such that  $f_0(\vec{y}) < f_0(\vec{x})$ . Clearly one has  $\|\vec{y} - \vec{x}\|_2 > \delta$  as otherwise this would contradict the assumption of local optimality of  $\vec{x}$ . As  $\mathcal{F}$  is a convex set it follows that  $\mathcal{F} \ni \vec{z} = (1 - \lambda)\vec{x} + \lambda\vec{y}$  for  $\lambda = \delta / (2\|\vec{y} - \vec{x}\|_2) < 1$ . Obviously  $\|\vec{x} - \vec{z}\|_2 = \delta/2 < \delta$ , i.e.,  $\vec{z}$  lies in the ball around  $\vec{x}$  for which  $\vec{x}$  is by assumption locally optimal. However, the convexity of  $f_0$  yields

$$f_0(\vec{z}) = f_0((1 - \lambda)\vec{x} + \lambda\vec{y}) \leq (1 - \lambda)f_0(\vec{x}) + \lambda f_0(\vec{y}) < f_0(\vec{x}), \quad (1.218)$$

in contradiction to the assumption. Therefore, the locally optimal point  $\vec{x}$  is also globally optimal. Further, in the context of convex optimization problems, one has an easy criterion for strong duality, called Slater's condition. It states that for the optimization problem in Eq. (1.216) strong duality holds if there exists a point  $\vec{x} \in \mathcal{F}$  that is strictly feasible. In fact, Slater's condition does not only imply strong duality for convex problems. It also implies that the dual optimal value is attained when  $d^* > -\infty$ .

### Conic programming

The theory of convex programming allows for the optimization of a convex function over an arbitrary convex set, which is described via the intersection of 0-sublevel sets of convex functions. A special subclass of convex sets are convex cones which appear to yield a robust and rich theory. A set  $C \subset \mathbb{R}^n$  is called a cone if  $\vec{x} \in C$  implies  $\lambda\vec{x} \in C$  for all  $\lambda \geq 0$ . If the set  $C$  is in addition convex, we will call  $C$  a convex cone. A convex cone  $C$  will be called proper, if it is closed, has a nonempty interior and is pointed, i.e., if  $\vec{x} \in C$  and  $-\vec{x} \in C$  then  $\vec{x} = \vec{0}$ . Given such a cone  $C$ , one can simply take the function

$$f_1(\vec{x}) := \begin{cases} 0, & \text{if } \vec{x} \in C \\ \infty, & \text{otherwise} \end{cases} \quad (1.219)$$

as the inequality constraint function in Eq. (1.216), thus reproducing the form of a convex program. These proper cones have the property that they induce a partial ordering on  $\mathbb{R}^n$  that can be used to define a so-called generalized inequality. The associated ordering to the cone  $C$  is given by  $\vec{x} \leq \vec{y}$  if and only if  $\vec{y} - \vec{x} \in C$ .

### Semidefinite programming

A semidefinite program (SDP) is a conic optimization program where the convex cone is given by the set of positive semidefinite (PSD) matrices. Therefore, a SDP has the form

$$\begin{aligned} & \text{minimize} && \vec{c}^\top \vec{x} \\ & \text{such that} && \sum_j x_j F_j + G \geq 0 \\ & && A\vec{x} = \vec{b}, \end{aligned} \tag{1.220}$$

where  $G, F_j$  are symmetric matrices and  $\vec{c} \in \mathbb{R}^n$ .

It often appears that the problem at hand does not directly offer a formulation as a SDP but can be refined in that form. Typical examples include the fidelity between two states, the trace distance, the infinity norm, as well as many of the smooth entropies. For instance

$$\|\varrho\|_\infty = \sup \{ \text{Tr}[\varrho X] \mid 0 \leq X \leq \mathbb{1} \} = \min \{ \alpha \mid \varrho \leq \alpha \mathbb{1} \}. \tag{1.221}$$

### Linear programming

A linear program can be seen as a semidefinite program where all matrices are diagonal and therefore the linear matrix inequalities reduce to a set of linear inequalities. Consequently, a linear program (LP) is of the form

$$\begin{aligned} & \text{minimize} && \vec{c}^\top \vec{x} \\ & \text{subject to} && A\vec{x} = \vec{b} \\ & && \vec{x} \geq 0. \end{aligned} \tag{1.222}$$

However, one can also interpret a linear program as a conic optimization problem, where the cone is given by the nonnegative orthant, i.e., the set  $\{\vec{x} \in \mathbb{R}^n \mid x_j \geq 0 \forall j\}$ .

## 2 Certifying irreducible measurements in a prepare-and-measure scenario

The number of outcomes is a defining property of a quantum measurement. In particular, there exist measurements which cannot be simulated by randomizing simpler measurements with fewer outcomes. These measurements are called irreducible and provide advantages for many quantum information processing tasks. In this Chapter we show that in a prepare-and-measure scenario the minimal scheme for certifying an irreducible three-outcome qubit measurement requires three state preparations and an auxiliary two-outcome measurement. Further, we provide experimentally feasible examples for this minimal certification scheme. In addition, we discuss the dimension assumption which is characteristic for the prepare-and-measure approach and to what extent it can be mitigated. This Chapter is based on Project [I].

### 2.1 Motivation

The most general description of a quantum measurement is given by positive operator-valued measures. The set of measurements described in this way contains instances which are neither projective nor obtainable by combining projective measurements [40]. This class of genuinely nonprojective measurements has important applications, for instance, in quantum computing [123, 124], quantum cryptography [125, 126], randomness certification [127], and quantum tomography [128]. However, since genuinely nonprojective measurements cannot be combined from projective measurements, their experimental implementation is difficult and typically requires control over additional degrees of freedoms, what can be seen as the content of the Naimark extension theorem.

For example, the implementation of the so-called symmetric informationally complete (SIC) POVM (see Fig. 2.3) via post-processing a projective measurement can only be achieved by coupling the system to at least one auxiliary qubit [129]. A quantum circuit implementing such a dilation is depicted in Fig. 2.1. It is hence of interest to verify whether an experiment has successfully implemented a nonprojective measurement.

Recently, semi-device-independent certification schemes have become the focus of theoretical investigations [130–132] as well as experimental implementations [133–135].

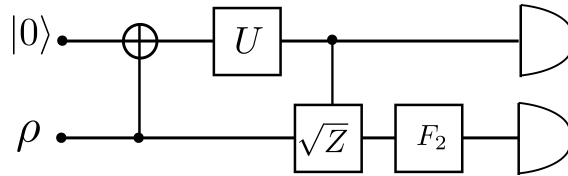


Figure 2.1: Quantum circuit implementing a SIC-POVM on a qubit. Here,  $U$  is a unitary operator depending on the particular orientation of the SIC effect vectors and  $F_2$  refers to the Fourier gate acting on the system. At the end of the circuit, both particles are measured in the computational basis, i.e., a projective measurement is performed.

Here the employed certification schemes can be divided into two classes. Those based on Bell-like scenarios [133] and those using a prepare-and-measure scenario [130–132, 134]. In the former case, an entangled state is distributed to two spatially separated measurement stations and the correlations between the different measurements at each station can be used to certify the presence of a genuinely nonprojective measurement. In the latter case, the certification consists of several preparation procedures, possibly intermediate transformations, and subsequent measurements on the same system. Both scenarios are in the sense device-independent as only very rudimentary assumptions need to be made about the particular implementation details of the state preparations and measurement devices.

For the Bell-like scheme, it turns out that those certification procedures can be formulated such that they are agnostic to the dimension of the prepared system. Even more, by employing the Collins-Gisin-Linden-Massar-Popescu inequalities [136, 137] it is possible to derive certificates for any number of outcomes. This pleasant property can be seen as a consequence of the convexity of the set of all  $n$ -chotomic non-signaling correlations [133]. This is in stark contrast to the prepare-and-measure approach. First, nonprojective measurements can always be implemented on a system with enlarged Hilbert space, hence an upper limit on the dimension of the prepared system is needed. Second, the dimension constraint typically renders the set of correlations nonconvex which makes its analysis more complex. Indeed, no family of certification criteria for  $n$ -chotomic measurements in the prepare-and-measure scenario is known.

In this Chapter we study the structure of correlations produced by irreducible three-outcome qubit measurements in the prepare-and-measure scenario. We begin by revisiting the concept of prepare-and-measure scenarios, the simulability of measurements and unambiguous state discrimination in Section 2.2. In Section 2.3 we define the operational setup in which a three-outcome qubit measurement can be certified. In particular, we prove the necessity of one auxiliary measurement and three preparation procedures. We proceed by investigating the robustness of genuine trichotomic

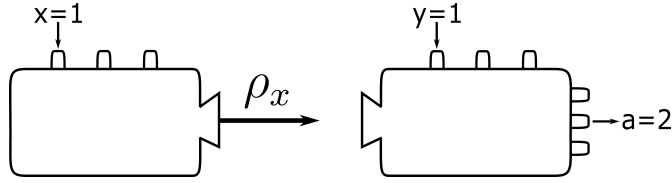


Figure 2.2: The structure of a prepare-and-measure scenario. The experimenter can choose a preparation labeled by  $x$  and a measurement labeled by  $y$ . This yields the quantum state  $\rho_x$  on which a measurement  $M_y$  is performed, resulting in the outcome  $a$ . Over many runs, the collected data allows one to compute the statistics  $p(a|x, y)$ .

correlations in Section 2.4. Afterwards, we discuss in Section 2.5 how the dimension assumption on the prepared system can be partly mitigated using a variant of theory-agnostic tomography. Finally, we analyse in Section 2.6 the experimental feasibility of the minimal scheme and provide concrete bounds on the preparation fidelity necessary.

## 2.2 Concepts and notation

### 2.2.1 The prepare-and-measure scenario

A prepare-and-measure scenario can be understood as a setup that is composed solely of a preparation device and a measurement device, see Fig. 2.2. An experimenter can choose among  $s$  different preparation procedures labeled by  $x \in \{1, \dots, s\}$  and  $m$  different measurements labeled  $y \in \{1, \dots, m\}$ . After choosing one particular pair of preparations and measurements  $(x, y)$  the experimenter produces the state  $x$  on which the measurement  $y$  is performed. Consequently, experiments of this kind can be fully characterized by the experimentally accessible correlations  $p(a|x, y)$  which give the probability of obtaining outcome  $a$  when performing measurement  $y$  on preparation  $x$ . Here it is important to note that the preparations as well as the measurements are considered as a black box, that is, no assumption is made about the prepared state  $\rho_x$  and the measurement description  $M_y$ . However, in order to render the prepare-and-measure approach useful for discrimination tasks, one typically assumes that the dimension of the underlying Hilbert space is fixed. In this sense, properties of the system that can be deduced from the experimental data alone are semi-device-independent.

### 2.2.2 The structure of measurements

As already mentioned in Section 1.1.2, any  $n$ -outcome quantum measurement  $M$  with  $n < \infty$  on a  $d$ -dimensional Hilbert space can be seen as a collection of positive semidef-

inite operators  $M = (M_1, \dots, M_n)$  satisfying  $\sum_a M_a = \mathbb{1}$ . The set of all POVMs is convex, that is, it is closed with respect to taking probabilistic mixtures and efficient algorithms are known in order to decompose a given POVM into extremal POVMs [138].

It is interesting to notice that a similar notion of POVMs can also be introduced in classical probability theory. Here a general  $n$ -outcome measurement on a  $d$ -dimensional classical system is given by  $n$  vectors from the  $d$ -dimensional unit cube  $C_d = \{\vec{x} \in \mathbb{R}^d \mid 0 \leq \vec{x}_j \leq 1\}$  such that they sum up to  $(1, \dots, 1) \in \mathbb{R}^d$ . Consequently, the set of all classical  $n$ -outcome measurements is equivalent to the set of all right stochastic  $d \times n$  matrices. We mention that this classical case is identical to the quantum case when one restricts all effects and states to be diagonal in some fixed basis. In the following we want to distinguish between different kinds of measurements [122].

**Definition 11.** *Let  $M$  be a generalized measurement.*

- (1) *We call  $M$  dichotomic if it has only two nonzero outcomes, trichotomic for three nonzero outcomes and  $n$ -chotomic in the case of  $n$  nonzero outcomes.*
- (2) *If  $M$  has  $n$ -outcomes we say that it can be simulated with  $n'$ -chotomic POVMs  $(N_\ell)_\ell$  if there exists a probability distribution  $(p_\ell)_\ell$  such that*

$$M = \sum_{\ell} p_{\ell} N_{\ell}. \quad (2.1)$$

*Otherwise the measurement is called irreducibly  $n$ -chotomic.*

For the particular case of  $n = 3$  and  $n' = 2$  the simulation reduces to the randomization of three dichotomic measurements, that is,

$$(M_1, M_2, M_3) = p_1(N_{1|1}, N_{2|1}, 0) + p_2(0, N_{2|2}, N_{3|2}) + p_3(N_{1|3}, 0, N_{3|3}), \quad (2.2)$$

where we write  $N_{a|\ell}$  for the outcome  $a$  of the measurement  $N_\ell$ . These reducible three-outcome measurements form a convex subset of the set of all measurements. While in  $d$ -dimensional classical probability theory all measurements are reducible to  $d$ -outcome measurements, this is not the case for quantum theory [40]. An archetypical counterexample is the trine POVM  $S$  which is composed of three qubit effects given by  $S_a = \frac{2}{3}|S_a\rangle\langle S_a|$  where the  $|S_a\rangle$  for  $a = 1, 2, 3$  are located in a plane of the Bloch sphere and are rotated by an angle of  $(2/3)\pi$  against each other, see also Fig. 2.3.

### 2.2.3 Unambiguous state discrimination

We have seen in Section 1.1.2 that unambiguous state discrimination (USD) is a special instance of quantum state estimation. In what follows we formulate the task of USD for the special case where the system subject to discrimination is prepared in one of two pure states. Then one party (called Alice) randomly but with equal probability chooses one of the two pure states  $|\psi_1\rangle$  and  $|\psi_2\rangle$  which are known to both parties

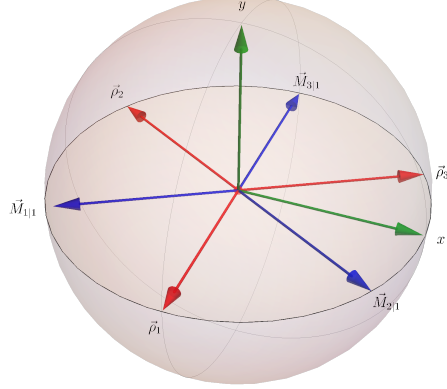


Figure 2.3: Representation of the state- and effect configuration for generating trichotomic correlations with the trine POVM. The states  $\rho_x$  are obtained from the vectors labeled by  $\vec{\rho}_x$  via  $\rho_x = \frac{1}{2}(\mathbb{1} + \sum_k [\vec{\rho}_x]_k \sigma_k)$  for  $x \in \{1, 2, 3\}$ , and the effects  $M_{a|1}$  are obtained from the vectors labeled by  $\vec{M}_{a|1}$  via  $M_{a|1} = \frac{1}{3}(\mathbb{1} + \sum_k [\vec{M}_{a|1}]_k \sigma_k)$  for  $a \in \{1, 2, 3\}$ . The figure is taken from Ref. [I].

and sends it to a receiver (called Bob). Again, as long as Alice's two states are not perfectly distinguishable, that is  $\langle \psi_1 | \psi_2 \rangle \neq 0$ , Bob cannot achieve a unit success rate in Eq. (1.46). Bob's best measurement in order to discriminate these states with maximal success probability is given by the irreducibly trichotomic POVM  $M = (M_1, M_2, M_3)$  with

$$M_1 = \frac{\mathbb{1} - |\psi_2\rangle\langle\psi_2|}{1 + |\langle\psi_1|\psi_2\rangle|}, \quad M_2 = \frac{\mathbb{1} - |\psi_1\rangle\langle\psi_1|}{1 + |\langle\psi_1|\psi_2\rangle|} \quad (2.3)$$

and  $M_3 = \mathbb{1} - M_1 - M_2$ . With this construction it is directly clear that if the measurement yields outcome  $a = 1$  or  $a = 2$ , then one can conclude that the received state was  $|\psi_a\rangle$ , while outcome  $a = 3$  does not allow a definite statement.

## 2.3 Certification from correlations

Motivated by the above considerations, we define now a family of correlations that is particularly useful for our later analysis. We consider a prepare-and-measure scenario with three preparations,  $\rho_1, \rho_2$ , and  $\rho_3$ , and two measurements,  $M_1 = (M_{1|1}, M_{2|1}, M_{3|1})$  and  $M_2 = (M_{1|2}, M_{2|2})$ . The states and the measurements are chosen in such a way, that if  $M_2$  yields the outcome 1 this implies that the received state was not  $\rho_1$  and if the received state is  $\rho_2$ , then  $M_2$  produces with certainty outcome 2. For  $M_1$  we impose that from outcome 1 (2) it follows that the state was not  $\rho_3$  ( $\rho_2$ ). Since any POVM has to obey the normalization condition, the last effect can always be calculated from the previous ones. Hence we arrange the correlations  $p(a|x, y) = \text{tr}(\rho_x M_{a|y})$  in a  $3 \times 3$



matrix  $\mathcal{P}$ , where the rows correspond to the states and the columns to the effects  $M_{1|1}, M_{2|1}, M_{1|2}$ , that is,

$$\mathcal{P} = \begin{pmatrix} p(1|1,1) & p(2|1,1) & 0 \\ p(1|2,1) & 0 & 1 \\ 0 & p(2|3,1) & p(1|3,2) \end{pmatrix}. \quad (2.4)$$

Since correlations of this form are motivated by USD, we refer to them as USD correlations. In particular, if the dimension of the system is known to be  $d$ , we write  $\text{USD}_d$  for the set of all USD correlation achievable under this constraint. Obviously the sets  $\text{USD}_d$  obey the inclusion  $\text{USD}_d \subset \text{USD}_{d+1}$  and furthermore, as we will see following, dimension 3 is already sufficient to achieve all USD correlations,  $\text{USD}_d = \text{USD}_3$  for all  $d > 3$ . We are therefore particularly interested in the qubit case, for which we have the following characterization.

**Theorem 12.** *The set  $\text{USD}_2$  consists exactly of all correlations of the form*

$$\mathcal{P}(p, q, \xi) = \begin{pmatrix} p\xi & q & 0 \\ p(1-\xi) & 0 & 1 \\ 0 & q(1-\xi) & \xi \end{pmatrix}, \quad (2.5)$$

with  $p, q, \xi \in [0, 1]$  such that  $(1-p)(1-q) \geq pq\xi$ . In addition, for a given correlations matrix  $\mathcal{P}$ , the states and measurements realizing  $\mathcal{P}$  are unique, up to a global unitary transformation.

*Proof.* Suppose that  $\mathcal{P} \in \text{USD}_2$ . This has consequences for the measurements  $M_1 = (M_{1|1}, M_{2|1}, M_{3|1})$ ,  $M_2 = (M_{1|2}, M_{2|2})$  and the states  $\rho_1, \rho_2, \rho_3$  that can realize the correlations. In particular,  $\mathcal{P}_{2,3} = \text{tr}[\rho_2 M_{1|2}] = 1$  and  $\mathcal{P}_{1,3} = 0$  imply  $\rho_2 = M_{1|2} = |\psi\rangle\langle\psi|$  and  $\rho_1 = |\psi^\perp\rangle\langle\psi^\perp|$ , where  $|\psi\rangle$  and  $|\psi^\perp\rangle$  are two orthonormal vectors. With a similar argument we obtain  $M_{2|1} = q|\psi^\perp\rangle\langle\psi^\perp|$  with  $0 \leq q \leq 1$ . It remains to consider the consequences of  $\mathcal{P}_{3,1} = 0$  for  $\rho_3$  and  $M_{1|1}$ . This requires  $M_{1|1} = p|\eta^\perp\rangle\langle\eta^\perp|$  and  $\rho_3 = |\eta\rangle\langle\eta|$  for some orthonormal vectors  $|\eta\rangle$  and  $|\eta^\perp\rangle$  and  $0 \leq p \leq 1$ .

Without loss of generality, we can assume

$$|\eta\rangle = \sqrt{\xi}|\psi\rangle + \sqrt{1-\xi}e^{i\phi}|\psi^\perp\rangle, \quad (2.6)$$

$$|\eta^\perp\rangle = \sqrt{1-\xi}|\psi\rangle - \xi e^{i\phi}|\psi^\perp\rangle, \quad (2.7)$$

where  $0 \leq \xi \leq 1$  and  $\phi \in \mathbb{R}$ . This yields immediately Eq. (2.5) together with the conditions  $p, q, \xi \in [0, 1]$ . For  $M_1$  to form a POVM it remains to verify that  $M_{3|1} = \mathbb{1} - M_{1|1} - M_{2|1}$  is positive semidefinite. This reduces here to  $\text{tr}(M_{3|1}) \geq 0$  and  $\det(M_{3|1}) \geq 0$  and can be equivalently expressed as the single condition  $(1-p)(1-q) \geq pq\xi$ .

From the above construction it is also immediately clear that conversely, any choice of  $p, q, \xi$  satisfying the constraints in Lemma 12 is in  $\text{USD}_2$ . Finally, given  $\mathcal{P}$ , all effects

and states are fixed by the above considerations, except for the choice of the orthonormal basis  $\{|\psi\rangle, e^{i\phi}|\psi^\perp\rangle\}$ , proving the claim of a unique representation up to a unitary transformation.  $\square$

### 2.3.1 Certifying trichotomic measurements

In the following we are concerned with the question whether the difference between reducible and irreducible  $n$ -outcome measurements can be observed using only the (empirically accessible) correlations, that is, whether there exists an irreducible  $n$ -outcome POVM  $M = (M_1, \dots, M_n)$  together with auxiliary measurements  $M_2, \dots, M_m$  and quantum states  $\rho_1, \dots, \rho_s$  such that the correlations  $p(a|x, y) = \text{Tr}[\rho_x M_{a|y}]$  cannot stem from a reducible measurement. Correlations of this type are genuinely  $n$ -chotomic, otherwise simulable  $n$ -chotomic. Clearly, if such correlations exist, they enable us to certify that the measurement  $M_1$  is indeed irreducible. In order to illustrate the difference between the concept of genuinely irreducible measurements and genuinely irreducible correlations consider the following example. Suppose that one implements the trine POVM  $S$  on a qubit system, which is an irreducible three-outcome measurement. As we will see in the following, this measurement alone can never yield correlations which cannot be explained by a reducible measurement. Therefore, an irreducible three-outcome measurement does not necessarily define genuine trichotomic correlations.

### 2.3.2 The minimal scenario

Suppose that we want to certify that a given measurement apparatus implements an irreducible  $n$ -outcome POVM  $M_1$ . What is the minimal scenario in which one can conclude from the output statistics alone that the POVM is irreducible? More precisely, what is the minimal number of state preparations  $s$  and auxiliary measurements  $m - 1$ ?

Clearly, if one has access to  $s$  preparations and  $m - 1$  auxiliary measurements, the set of all possible correlations  $p(a|x, y)$  that can be obtained in a scenario without any constraint on the dimension of the system yields a convex set. As it turns out, this convex set is a polytope, whose extremal points correspond to deterministic correlations [139], that is, where all  $p(a|x, y)$  are either 0 or 1. If the dimension  $d$  of the system is at least  $s$  then these extremal points can be obtained from a fixed choice of  $s$  orthogonal states  $\rho_x = |\psi_x\rangle\langle\psi_x|$  and at most  $s$ -chotomic measurements with effects

$$M_{a|y} = \sum_x p(a|x, y) \rho_x \quad (2.8)$$

Therefore all correlations can be written as convex combination of deterministic strategies. Since all extreme points use the same states, the convex coefficients can be ab-

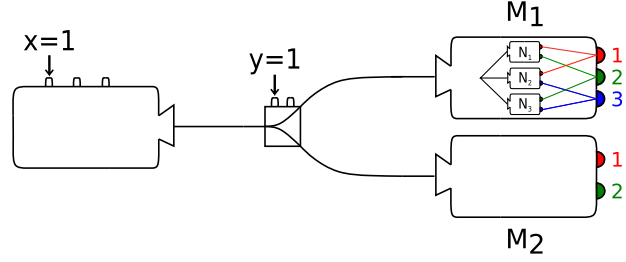


Figure 2.4: Prepare-and-measure setup of the minimal scenario producing simulable trichotomic correlations. On the left hand side one can choose between three different preparation procedures  $x \in \{1, 2, 3\}$  and the corresponding qubit state is sent to one of two measurement devices (right hand side). The measurement device is chosen by the experimenter and this choice is denoted by  $y \in \{1, 2\}$ . For  $y = 1$  the measurement can yield three different outcomes, but governed by a specific inner mechanism. It consists of three two-outcome measurements, one of which is chosen at random. Depending on the outcome of this measurement, an outcome is assigned to the overall three-outcome measurement. If  $y = 2$  the measurement device is a simple two-outcome measurement. The figure is taken from Ref. [I].

sorbed into the effects yielding at most  $s$ -chotomic POVMs. For the case of an irreducible three-outcome measurement on a qubit,  $n = 3$  and  $d = 2$ , this implies that at least  $s = 3$  different states are required. From this it also follows that  $\text{USD}_d \subset \text{USD}_3$  for all  $d \geq 3$  since only three different states are used in the USD correlations.

Regarding the number of auxiliary measurements  $m - 1$ , we consider the case where no auxiliary measurements are used. In this scenario, we denote by  $\mathcal{C}_d(s, n)$  the convex hull of all correlations on a  $d$ -dimensional classical system and by  $\mathcal{Q}_d(s, n)$  the convex hull of all correlations on a  $d$ -dimensional quantum system.

**Theorem 13** ([140]). *The sets  $\mathcal{C}_d(s, n)$  and  $\mathcal{Q}_d(s, n)$  coincide for all values of  $s, n, d$ . In particular, the set of all quantum correlations  $\mathcal{Q}_d(s, n)$  is a polytope.*

Since all correlations in  $\mathcal{C}_d(s, n)$  can be obtained from  $d$ -chotomic measurements, these property must also hold for all correlations in  $\mathcal{Q}_d(s, n)$ . Therefore, the smallest scenario which allows us to certify an irreducible three-outcome measurement on a qubit includes at least two measurements  $M_1 = (M_{1|1}, M_{2|1}, M_{3|1})$  and the auxiliary measurement  $M_2 = (M_{1|2}, M_{2|2})$ . An illustration of this minimal scenario can be found in Fig. 2.4.

The set of correlations achievable in the minimal scenario with fixed  $d$  is subsequently denoted by  $\text{COR}_d$ . Furthermore, we write  $\text{SIM}_d$  for the set of all simulable trichotomic correlations within  $\text{COR}_d$ . Clearly, the correlations  $\text{USD}_d$  are a subset of  $\text{COR}_d$  and  $\text{SIM}_s \cap \text{USD}_d$  are the simulable trichotomic correlations within  $\text{USD}_d$ .

### 2.3.3 The geometry of trichotomic correlations

We now want to investigate how the sets  $\text{SIM}_2$  and  $\text{COR}_2$  are related. In contrast to Bell-scenarios, where the convex hull of the correlations is taken, here we restrict the study to the bare sets  $\text{COR}_2$  and  $\text{SIM}_2$ . These sets are not convex, as can be seen by considering the subset  $\text{USD}_2$ , characterized by Theorem 12. In fact it is evident that the correlations matrix  $\mathcal{D}_1 = \mathcal{P}(1, 0, 1)$  and  $\mathcal{D}_2 = \mathcal{P}(1, 1, 0)$  can be realized with dichotomic measurements, that is,  $\mathcal{D}_1, \mathcal{D}_2 \in \text{USD}_2 \cap \text{SIM}_2$ . However, no convex combination  $\mathcal{D}_\lambda = \lambda \mathcal{D}_1 + (1 - \lambda) \mathcal{D}_2$  with  $0 < \lambda < 1$  can be written in the form  $\mathcal{P}(p, q, \xi)$  as given by Eq. (2.5). Here it is important to note that the correlations  $\mathcal{D}_\lambda$  are still valid USD correlations, that is, they are of the correct form. Hence  $\mathcal{D}_\lambda \in \text{COR}_2$  already implies  $\mathcal{D}_\lambda \in \text{USD}_2$ . The relation of the points  $\mathcal{D}_1$  and  $\mathcal{D}_2$  and their convex mixture  $\mathcal{D}_\lambda$  to the sets  $\text{USD}_2$  and  $\text{SIM}_2$  is illustrated in Fig. 2.5. As  $\text{USD}_2$  is an affine section of  $\text{COR}_2$ , it follows that neither  $\text{COR}_2$  nor  $\text{SIM}_2$  is convex. Our next step is to establish that not all qubit USD correlations are simulable trichotomic.

**Theorem 14.** *There exist correlations in  $\text{USD}_2$  that are not contained in the convex hull of  $\text{USD}_2 \cap \text{SIM}_2$ . In particular, even the convex hull of the simulable trichotomic qubit USD correlations does not cover all qubit USD correlations*

*Proof.* The first step is to parametrize the correlations  $\mathcal{D} \in \text{SIM}_2 \cap \text{USD}_2$  with  $\mathcal{D}_{3,3} \neq 0$ . By virtue of Eq. (2.2) and the proof of Theorem 12 above, the effects of the simulated trichotomic POVM  $M_1$  can be written as

$$M_{1|1} = p|\eta^\perp\rangle\langle\eta^\perp| = \kappa_1 F_1 + \kappa_3(\mathbb{1} - F_3), \quad (2.9)$$

$$M_{2|1} = q|\psi^\perp\rangle\langle\psi^\perp| = \kappa_1(\mathbb{1} - F_1) + \kappa_2 F_2, \quad (2.10)$$

where, compared to Eq. (2.2) we write  $\kappa_j$  in place of  $p_j$  and  $F_j$  in place of  $N_{j|j}$ . For  $\kappa_1 \neq 0$  it follows that  $F_1 \propto |\eta^\perp\rangle\langle\eta^\perp|$  and  $\mathbb{1} - F_1 \propto |\psi^\perp\rangle\langle\psi^\perp|$ , yielding  $F_1 = |\eta^\perp\rangle\langle\eta^\perp| = |\psi\rangle\langle\psi|$ . This is in contradiction to the assumption  $\mathcal{D}_{3,3} \neq 0$  by virtue of  $\mathcal{D}_{3,3} = \text{Tr}[M_{1|2}\varrho_3]$  with  $\varrho_3 = |\eta\rangle\langle\eta|$  and  $M_{1|2} = |\psi\rangle\langle\psi|$ . Therefore we can conclude that  $\kappa_1 = 0$ . This allows us to write  $\kappa := \kappa_2 = 1 - \kappa_3$ . This together with Eq. (2.10) implies  $F_2 = f_2|\psi^\perp\rangle\langle\psi^\perp|$  with  $0 \leq f_2 \leq 1$  and  $\kappa f_2 = q$ . Similarly, one can obtain from Eq. (2.9) that  $\mathbb{1} - F_3 = f_3|\eta^\perp\rangle\langle\eta^\perp|$  with  $(1 - \kappa)f_3 = p$ . Consequently,  $\mathcal{D} \in \text{USD}_2 \cap \text{SIM}_2$  with  $\mathcal{D}_{3,3} \neq 0$  if and only if

$$\mathcal{D} = \begin{pmatrix} f_3(1 - \kappa)\xi & f_2\kappa & 0 \\ f_3(1 - \kappa)(1 - \xi) & 0 & 1 \\ 0 & f_2\kappa(1 - \xi) & \xi \end{pmatrix} \quad (2.11)$$

with  $f_2, f_3, \kappa, \xi \in [0, 1]$  and  $\xi \neq 0$ .

In the next step we show that for any  $0 < \xi < 1$  there exist correlations  $\mathcal{P} \in$

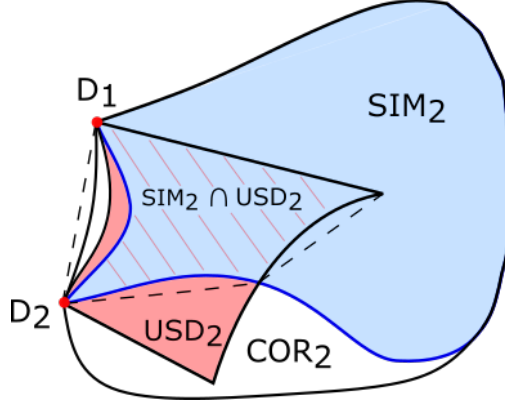


Figure 2.5: Schematic illustration of the relations among the sets  $\text{USD}_2$ ,  $\text{SIM}_2$  and  $\text{COR}_2$ . The set  $\text{COR}_2$  of all correlations that can be obtained with a qubit in the minimal scenario is obviously a superset of  $\text{SIM}_2$ , the set of all simulable correlations in the minimal scenario as well as  $\text{USD}_2$ , the set of all correlations that can be obtained with a qubit and are of the form (2.4). The points  $D_1$  and  $D_2$  represent deterministic correlations contained in  $\text{USD}_2 \cap \text{SIM}_2$ , hence located at the boundary of any of the three sets. The dashed line represents the convex hull of the set  $\text{USD}_2 \cap \text{SIM}_2$ . The figure is taken from Ref. [1].

$\text{USD}_2 \setminus \text{SIM}_2$ . For this purpose we consider the linear map

$$\mathcal{W} : \mathcal{P} \mapsto -\mathcal{P}_{1,1} - \mathcal{P}_{1,2} + \mathcal{P}_{3,2} + \mathcal{P}_{3,3}. \quad (2.12)$$

If one now chooses  $\mathcal{P} = \mathcal{P}(p, q, \xi)$  as in Eq. (2.4) with the choice  $p = \frac{1}{2}$  and  $q = 1/(\xi + 1)$ , we verify that for any  $0 < \xi < 1$  the constraint  $(1 - p)(1 - q) \geq pq\xi$  is satisfied. Further one finds that in addition  $\mathcal{W}(\mathcal{P}) < 0$  holds. However, for any  $\mathcal{D}$  as in Eq. (2.11) we have  $\mathcal{W}(\mathcal{D}) = \xi[1 - f_2\kappa - f_3(1 - \kappa)] \geq 0$  and in addition for any  $\mathcal{T} \in \text{USD}_2$  with  $\mathcal{T}_{3,3} = 0$  we have immediately  $\mathcal{W}(\mathcal{T}) = 0$ . Consequently  $\mathcal{W}(\mathcal{S} - \mathcal{P}) > 0$  for our choice of  $\mathcal{P}$  and any  $\mathcal{S} \in \text{SIM}_2 \cap \text{USD}_2$ , that is  $\mathcal{S} \neq \mathcal{P}$ .

From this observation it readily follows that the convex hull of  $\text{SIM}_2 \cap \text{USD}_2$  does not contain all of  $\text{USD}_2$ . This is because  $\mathcal{W}$  is a linear map and thus its minimum over the convex hull of  $\text{SIM}_2 \cap \text{USD}_2$  is attained already for some  $\mathcal{S} \in \text{SIM}_2 \cap \text{USD}_2$ . However, for this set we just proved that  $\mathcal{W}(\mathcal{S}) > \mathcal{W}(\mathcal{P})$  for certain  $\mathcal{P} \in \text{USD}_2$ .  $\square$

This might raise the expectation that there exists a linear inequality separating  $\text{COR}_2$  and  $\text{SIM}_2$ , despite their nonconvexity discussed above. However, note that the statement of Theorem 14 only concerns the subset of USD correlations and one cannot conclude directly that the convex hull of  $\text{SIM}_2$  is a proper subset of  $\text{COR}_2$ .

## 2.4 Robustness of trichotomic correlations

For an experimental certification of an irreducible three-outcome measurement it is essential to find correlations  $\mathcal{P} \in \text{COR}_2$  such that the closest simulable correlations  $\mathcal{P}' \in \text{SIM}_2$  have a distance  $r$  of reasonable size. This distance can in principle be measured with respect to various norms. Here we measure the distance either in terms of the supremum norm, yielding  $r_\infty$ , or in terms of the Euclidean norm, yielding  $r_2$ , that is,

$$r_\infty = \max_{i,j} |\mathcal{P}_{i,j} - \mathcal{P}'_{i,j}|, \quad (2.13)$$

$$r_2 = \left( \sum_{i,j} (\mathcal{P}_{i,j} - \mathcal{P}'_{i,j})^2 \right)^{1/2}. \quad (2.14)$$

According to Theorem 14, we can preliminarily focus on the family of USD correlations, as it guarantees the existence of USD correlations  $\mathcal{P} \in \text{USD}_2$  such that  $r > 0$ .

### 2.4.1 The computation of upper bounds

In order to compute  $r_2$  and  $r_\infty$  we rely on numerical optimization over the set  $\text{SIM}_2$ . The optimization is nonlinear and involves three, possibly mixed quantum states as well as four dichotomic POVMs. Here it is important to note that if the states or effects are fixed, the problem can be rephrased as a SDP and becomes thereby easy to solve numerically. However, technically this optimization algorithm only yields guaranteed upper bounds on the distances, because it is based on finding the correlations in  $\text{SIM}_2$  closest to  $\mathcal{P}$ . From a formal viewpoint, we want to compute the maximal radius  $r$  of a ball  $B_r(\mathcal{P})$  around a given correlation  $\mathcal{P} \in \text{COR}_2 \cap \text{SIM}_2$  such that  $B_r(\mathcal{P}) \cap \text{SIM}_2$  is empty. For the relevant norms, this can be rephrased as the optimization problem to minimize a real parameter  $t$  over  $\mathcal{Q} \in \text{SIM}_2$  such that

$$-t \leq \mathcal{P}_{i,j} - \mathcal{Q}_{i,j} \leq t \quad \forall i, j, \quad (2.15)$$

$$\sum_{i,j} (\mathcal{P}_{i,j} - \mathcal{Q}_{i,j})^2 \leq t^2. \quad (2.16)$$

We write  $F_1 = M_{1|1}$ ,  $F_2 = M_{2|1}$  and  $F_3 = M_{1|2}$  such that  $\mathcal{Q}_{i,j} = \text{Tr}[\varrho_i F_j]$ . If we keep the effects fixed, then the optimization is a semidefinite program of the following type: Minimize  $t$  under the constraint  $\varrho_i \geq 0$  and  $\text{Tr}[\varrho_i] = 1$  for  $i = 1, 2, 3$  and either the linear constraint given by Eq. (2.15) or the quadratic-convex constraint given by Eq. (2.16). Similar, if we keep the states fixed, then the optimization is again a semidefinite program, however now with the constraint on the states replaced by constraints on the

effects, namely

$$F_1 = F'_1 + F'_0, \quad F_2 = F'_2 + f_0\mathbb{1} - F'_0, \quad (2.17)$$

$$0 \leq F'_0 \leq f_0\mathbb{1}, \quad 0 \leq F'_1 \leq f_1\mathbb{1}, \quad 0 \leq F'_2 \leq f_2\mathbb{1}, \quad (2.18)$$

$$0 \leq F_3 \leq \mathbb{1}, \quad f_0 + f_1 + f_2 = 1. \quad (2.19)$$

As these small semidefinite programs can be solved very fast numerically, this invites a seesaw optimization, where one alternates between the two optimizations until  $t$  converges. We implement this seesaw algorithm using the Python library PICOS with the CVXOPT back-end. As criterion for convergence we take  $t_{n-1} - t_n < 10^{-6}$ , where  $t_n$  is the value after  $n$  seesaw iterations. This convergence happens after at most 300 iterations. We repeat the optimization 4500 times, each time with different starting values for  $\rho_1$ ,  $\rho_2$ , and  $\rho_3$ , where we take pure states chosen randomly according to the Haar measure and then decrease the purity  $\text{Tr}(\rho^2)$  to be uniformly in the interval  $[\frac{1}{2}, 1]$ . The same optimal value is always reached independently of the start values for the Euclidean norm, while it occurs only for about 1% of the start values in the case of the supremum norm. The largest distance we found is realized by the choice of  $\mathcal{P} = (0.577, 0.726, 0.276)$ , where the seesaw algorithm yields  $r_2 \approx 0.0391$  and  $r_\infty \approx 0.0177$ . Larger distances can be achieved using correlations which are not confined to  $\text{USD}_2$ . In particular, choosing an arrangement involving the trine POVM, we find  $r_2 \approx 0.0686$  and  $r_\infty \approx 0.0342$ . The exact states and measurements settings are illustrated in Fig. 2.3. More precisely the involved states are

$$q_1 = \frac{1}{2}(\mathbb{1} - \sigma_3), \quad q_2 = \frac{1}{2}(\mathbb{1} - \frac{\sqrt{3}}{3}\sigma_1 - \frac{1}{2}\sigma_3), \quad q_3 = \frac{1}{2}(\mathbb{1} + \frac{\sqrt{3}}{2}\sigma_1 - \frac{1}{2}\sigma_3), \quad (2.20)$$

and the measurement effects are

$$M_{1|1} = \frac{1}{3}(\mathbb{1} - q_3), \quad M_{2|1} = \frac{2}{3}(\mathbb{1} - q_2), \quad M_{3|1} = \frac{2}{3}(\mathbb{1} - q_1), \quad (2.21)$$

$$M_{1|2} = \mathbb{1} - q_1, \quad M_{2|2} = q_1. \quad (2.22)$$

## 2.4.2 The computation of lower bounds

However, in order to certify a given distance to the set  $\text{SIM}_2$  one has to compute a lower bound on the distance between a given correlation and the set  $\text{SIM}_2$ . Here we will derive a numerical method to obtain lower bounds with respect to the supremum norm and in the following we write  $B_\epsilon(\mathcal{P})$  for the ball with radius  $\epsilon > 0$  in the supremum norm. Similar as for the case of upper bounds, for a given correlation  $\mathcal{P} \in \text{USD}_2$  we are interested in the condition on the states and the measurements under which we have  $\mathcal{Q} \in B_\epsilon(\mathcal{P})$  for  $\epsilon > 0$ . By definition, any  $\mathcal{Q} \in \text{COR}_2$  is of the

form

$$\mathcal{Q} = \begin{pmatrix} \text{Tr} \begin{bmatrix} \varrho_1 M_{1|1} \\ \varrho_2 M_{1|1} \\ \varrho_3 M_{1|1} \end{bmatrix} & \text{Tr} \begin{bmatrix} \varrho_1 M_{2|1} \\ \varrho_2 M_{2|1} \\ \varrho_3 M_{2|1} \end{bmatrix} & \text{Tr} \begin{bmatrix} \varrho_1 M_{1|2} \\ \varrho_2 M_{1|2} \\ \varrho_3 M_{1|2} \end{bmatrix} \end{pmatrix}. \quad (2.23)$$

However, it turns out to be useful to group the free parameters that appear in Eq. (2.23) into two different classes. Define  $a = (\varrho_1, \varrho_2, M_{1|2})$  and  $b = (\varrho_3, M_{1|1}, M_{2|1})$ , which together form the set of parameters of the distribution  $\mathcal{Q} = \mathcal{Q}(a, b)$ . We now want to answer the question how large  $\epsilon > 0$  can be chosen such that there is no  $\mathcal{Q}(a, b) \in B_\epsilon(\mathcal{P})$  given that the trichotomic measurement  $M = (M_{1|1}, M_{2|1}, M_{3|1})$  is limited to be simulable by dichotomic measurements. In the case that  $\mathcal{Q}(a, b)$  is a USD correlation, then the elements  $\mathcal{Q}_{1,3}, \mathcal{Q}_{2,2}, \mathcal{Q}_{3,1}$  and  $\mathcal{Q}_{2,3}$  are all either 0 or 1. This imposes strong constraints on the parameters in the set  $a$ . Indeed, if  $\mathcal{Q}(a, b) = \mathcal{P}$ , then the value of  $a$  is fixed up to a unitary transformation, as we have seen in the proof of Theorem 14. Let us denote this fixed value for  $a$  by  $\alpha$ . In the following we will extend the argument in the proof of Theorem 14 in order to show that  $\mathcal{Q}(a, b) \in B_\epsilon(\mathcal{P})$  implies that  $a \in B_{\mathcal{O}(\sqrt{\epsilon})}(\alpha)$ . We will first outline the idea of the argument before proceeding with the formal treatment. Because the map  $(a, b) \mapsto \mathcal{Q}(a, b)$  is continuous, we can then show that  $a \in B_{\mathcal{O}(\sqrt{\epsilon})}(\alpha)$  implies  $\mathcal{Q}(a, b) \in B_{\mathcal{O}(\sqrt{\epsilon})}(\mathcal{Q}(\alpha, b))$ . Since  $\mathcal{Q}(a, b) \in B_\epsilon(\mathcal{P})$  and  $\mathcal{Q}(a, b) \in B_{\mathcal{O}(\sqrt{\epsilon})}(\mathcal{Q}(\alpha, b))$ , the triangle inequality directly yields that  $\mathcal{Q}(\alpha, b) \in B_{\epsilon + \mathcal{O}(\sqrt{\epsilon})}(\mathcal{P})$ . Therefore, we have reduced the problem of asking for the existence of  $b$  such that  $\mathcal{Q}(a, b) \in B_\epsilon(\mathcal{P})$  to asking for the existence of  $b$  such that  $\mathcal{Q}(\alpha, b) \in B_{\epsilon + \mathcal{O}(\sqrt{\epsilon})}(\mathcal{P})$ . The latter means that, for a fixed value of  $\epsilon$  and fixed values of  $\varrho_1, \varrho_2$  and  $M_{1|2}$  (corresponding to the parameters in the variable  $a$ ), we ask for the feasibility of  $\varrho_3, M_{1|1}$  and  $M_{2|1}$  (corresponding to the parameters in the variable  $b$ ) such that  $\mathcal{Q} \in B_{\epsilon + \mathcal{O}(\sqrt{\epsilon})}(\mathcal{P})$ . Although this is not yet a SDP it can be decomposed into a finite number of SDPs by scanning the values of  $\varrho_3$  and bounding the error in the finite scanning.

For the formal treatment we need to introduce some notation. We parametrize the effects and the states by Bloch coordinates, that is, for an effect  $E$  we write  $(x_0, \vec{x})$  which means  $E = (1/2) \sum_i x_i \sigma_i$ . In particular,  $M_{1|1} = (x_{01}, \vec{x}_1)$ ,  $M_{2|1} = (x_{02}, \vec{x}_2)$  and  $M_{1|2} = (y_0, \vec{y})$ . In a similar manner, for a state  $\varrho_j$  we write  $(1, \vec{r})$ . Due to the unitary freedom, we can always assume  $\vec{x}_1$  and  $\vec{x}_2$  to have no  $\sigma_3$  component and at the same time  $\vec{r}_2$  to have no  $\sigma_2$  component. First, we show that one can bound the trace of the effect from below, given a lower bound on the probability for that effect.

**Lemma 15.** *Let  $\varrho = (1, \vec{r})$  be a state and  $F = (x_0, \vec{x})$  an effect.*

- (1) *If  $\text{Tr}[F\varrho] \geq a$  for  $a \in \mathbb{R}$ , then  $x_0 \geq a$ .*
- (2) *If  $\text{Tr}[F\varrho] \leq b$  for  $b \in \mathbb{R}$ , then  $x_0 \leq 1 + b$ .*

*Proof.* In order to prove (1), notice that  $\text{Tr}[F\varrho] = (1/2)(x_0 + \vec{x} \cdot \vec{r}) \geq a$ . Because  $\vec{x} \cdot \vec{r} \leq$



$\|\vec{x}\| \|\vec{r}\| \leq x_0 \|\vec{r}\| \leq x_0$ , we have  $x_0 \geq a$ . To show (2), we have similarly  $\text{Tr}[F\rho] = (1/2)(x_0 + \vec{x} \cdot \vec{r}) \leq b$  yielding  $(1/2)(2 - x_0 - \vec{c} \cdot \vec{r}) \geq 1 - b$ . Then applying (1) one finds that  $2 - x_0 \geq 1 - b$  or equivalently  $x_0 \leq 1 + b$ .  $\square$

Now we can apply Lemma 15 to the condition  $\mathcal{Q} \in B_\epsilon(\mathcal{P})$ .

**Corollary 16.** (1) For  $\mathcal{Q} \in B_\epsilon(\mathcal{P})$  it is necessary that  $1 - \epsilon \leq y_0 \leq 1 + \epsilon$ ,  $c_1 - \epsilon \leq x_{01} \leq 1 + \epsilon$  with  $c_1 := \max\{\mathcal{P}_{1,2}, \mathcal{P}_{2,2}\}$  and  $c_2 - \epsilon \leq x_{02} \leq 1 + \epsilon$  with  $c_2 := \max\{\mathcal{P}_{1,3}, \mathcal{P}_{3,3}\}$ .

(2)  $\mathcal{Q}_{2,3} \geq 1 - \epsilon$  implies  $\|\vec{y} - \vec{r}_2\| \leq \sqrt{4\epsilon + \epsilon^2}$  and  $\mathcal{Q}_{1,3} \leq \epsilon$  implies  $\|\vec{y} + \vec{r}_1\| \leq \sqrt{4\epsilon + \epsilon^2}$ .

*Proof.* The first claim in (1) is a direct consequence of  $0 \leq \mathcal{Q}_{1,3} \leq \epsilon$  and  $1 \geq \mathcal{Q}_{2,3} \geq 1 - \epsilon$ . Similarly the second claim follows from  $\mathcal{Q}_{1,1} \in B_\epsilon(\mathcal{P}_{1,1})$  and  $\mathcal{Q}_{2,1} \in B_\epsilon(\mathcal{P}_{2,1})$  and the third from  $\mathcal{Q}_{1,2} \in B_\epsilon(\mathcal{P}_{1,2})$  and  $\mathcal{Q}_{3,2} \in B_\epsilon(\mathcal{P}_{3,2})$ . To prove (2) observe that  $\mathcal{Q}_{2,3} = (1/2)(y_0 + \vec{y} \cdot \vec{r}_2) \geq 1 - \epsilon$  implying that  $\vec{y} \cdot \vec{r}_2 \geq 2(1 - \epsilon) - y_0$ . Further one has  $\vec{y} \cdot \vec{r}_2 \geq 2(1 - \epsilon) - y_0$  yielding that  $\|\vec{y}\|^2 + \|\vec{r}_2\|^2 - \|\vec{y} - \vec{r}_2\|^2 \geq 4(1 - \epsilon) - 2y_0$ . Because  $\|\vec{y}\| \leq \min\{y_0, 2 - y_0\}$  and  $\|\vec{r}_2\| \leq 1$ , we obtain that

$$(\min\{y_0, 2 - y_0\})^2 + 2y_0 - 3 + 4\epsilon \geq \|\vec{y} - \vec{r}_2\|^2. \quad (2.24)$$

The left hand side of Eq. (2.24) is maximized at  $y_0 = 1 + \epsilon$ , which leads to  $4\epsilon + \epsilon^2 \geq \|\vec{y} - \vec{r}_2\|^2$ . The proof of the second statement in (2) is analogous to the proof presented above.  $\square$

Now we consider the constraint of the form  $\text{tr}[\rho F] \leq \epsilon$ . We show that if the trace of  $F$  is bounded from below, we can bound the purity of  $\rho$ . This also extends to the consideration of each Bloch components of the Bloch vectors.

**Lemma 17.** Let  $\rho = (1, \vec{r})$  be a quantum state and  $F = (x_0, \vec{x})$  be an effect such that  $x_0 \geq c$  and  $\vec{x}$  having no component with respect to  $\sigma_3$ . Then  $\text{Tr}[\rho F] \leq \epsilon$  implies that  $\|\vec{r}\| \geq \|\vec{r}_{xy}\| \geq 1 - (2\epsilon/c)$  and  $|r_z| \leq \sqrt{1 - (1 - (2\epsilon/c))^2}$ , where  $\vec{r}_{xy}$  denotes the vector  $\vec{r}$  where the  $\sigma_3$  component is set to 0. As a consequence, if  $\vec{n}$  denotes the unit vector pointing into the direction of  $\vec{r}_{xy}$ , then  $\|\vec{r}_{xy} - \vec{n}\| \leq 2\epsilon/c$  and  $\|\vec{r} - \vec{n}\| \leq (2\epsilon/c) + \sqrt{1 - (1 - (2\epsilon/c))^2}$ .

*Proof.* Observe that  $\text{Tr}[\rho F] = (1/2)(x_0 + \vec{r} \cdot \vec{x}) = (1/2)(x_0 + \vec{r}_{xy} \cdot \vec{x}) \leq \epsilon$ . This leads to  $-\vec{r}_{xy} \cdot \vec{x} \geq x_0 - 2\epsilon$ . Then we have  $\|\vec{r}_{xy}\| x_0 \geq x_0 - 2\epsilon$  and thus  $\|\vec{r}\| \geq \|\vec{r}_{xy}\| \geq 1 - (2\epsilon/x_0) \geq 1 - (2\epsilon/c)$ . Also  $1 \geq \|\vec{r}_z\|^2 + \|\vec{r}_{xy}\|^2 \geq |r_z|^2 + (1 - (2\epsilon/c))^2$ , such that  $\|\vec{r}_z\| \leq \sqrt{1 - (1 - (2\epsilon/c))^2}$ . The latter part of the statement is obvious.  $\square$

As a direct consequence of Lemma 17 we can estimate  $\vec{r}_2$  and  $\vec{r}_3$  by the unit vectors of their projection onto the  $xy$ -plane

**Corollary 18.** Let  $\vec{r}_{2xy}$  denote the  $xy$  component and  $\vec{r}_{2z}$  the  $z$  component of  $\vec{r}_2$ . Further, let  $\vec{n}$  be the unit vector in the direction of  $\vec{r}_{2xy}$  and  $\vec{t}$  the unit vector in the direction of  $\vec{r}_{3xy}$ .

- (1)  $Q_{3,1} \leq \epsilon$  implies  $\|\vec{r}_{3xy} - \vec{t}\| \leq 2\epsilon/(c_1 - \epsilon) = \epsilon_{3xy}$ . Further one finds that  $\|\vec{r}_{3z}\| \leq \sqrt{1 - (1 - 2\epsilon/(c_1 - \epsilon))^2} = \epsilon_{3z}$ .
- (2)  $Q_{2,2} \leq \epsilon$  implies that  $\|\vec{r}_{2xy} - \vec{n}\| \leq 2\epsilon/(c_2 - \epsilon) = \epsilon_{2xy}$ . Further one finds that  $\|\vec{r}_{2z}\| \leq \sqrt{1 - (1 - 2\epsilon/(c_2 - \epsilon))^2} = \epsilon_{2z}$ .

Now we can consider the cost of pinning the values of  $\vec{r}_1, \vec{r}_2$  and  $\vec{r}_3$ , where it is useful to start with  $\vec{r}_2$ . We consider the estimation of  $\vec{r}_2$  by  $\vec{n}$ . This leads to a new estimation of the matrix elements involving  $q_1$  and  $q_3$ . For simplicity, in the following we also write  $Q \in \mathcal{P} \pm \epsilon$  as a synonym of  $Q \in B_\epsilon(\mathcal{P})$ . With this notation one finds

$$Q_{1,1} = \frac{1}{2}[x_{01} - \vec{n} \cdot \vec{x}_1 + (\vec{n} - \vec{r}_{2xy}) \cdot \vec{x}_1 + (\vec{r}_2 - \vec{y}) \cdot \vec{x}_1 + (\vec{y} + \vec{r}_1) \cdot \vec{x}_1] \quad (2.25)$$

$$\in \frac{1}{2}[x_{01} - \vec{n} \cdot \vec{x}_1 \pm (\epsilon_{2xy} + 2\sqrt{4\epsilon + \epsilon^2})], \quad (2.26)$$

$$Q_{1,2} = \frac{1}{2}[x_{02} - \vec{n} \cdot \vec{x}_2 + (\vec{n} - \vec{r}_{2xy}) \cdot \vec{x}_2 + (\vec{r}_2 - \vec{y}) \cdot \vec{x}_2 + (\vec{y} + \vec{r}_1) \cdot \vec{x}_2] \quad (2.27)$$

$$\in \frac{1}{2}[x_{02} + (-\vec{n}) \cdot \vec{x}_2 \pm (\epsilon_{2xy} + 2\sqrt{4\epsilon + \epsilon^2})], \quad (2.28)$$

$$Q_{2,1} = \frac{1}{2}[x_{01} + \vec{n} \cdot \vec{x}_1 + (\vec{r}_{2xy} - \vec{n}) \cdot \vec{x}_1] \in \frac{1}{2}[x_{01} + \vec{n} \cdot \vec{x}_1 \pm \epsilon_{2xy}], \quad (2.29)$$

$$Q_{2,2} = \frac{1}{2}[x_{02} + \vec{n} \cdot \vec{x}_2 + (\vec{r}_{2xy} - \vec{n}) \cdot \vec{x}_2] \in \frac{1}{2}[x_{02} + \vec{n} \cdot \vec{x}_2 \pm \epsilon_{2xy}]. \quad (2.30)$$

Let us consider the error one introduces by fixing  $\vec{r}_{3xy}$  to  $\vec{t}$ . Note that this only affects the last row of the correlation matrix given by Eq. (2.23).

$$Q_{3,3} = \frac{1}{2}[y_0 + \vec{r}_2 \cdot \vec{r}_3 + (\vec{y} - \vec{r}_2) \cdot \vec{r}_3] \in \frac{1}{2}[y_0 + \vec{r}_2 \cdot \vec{r}_3 \pm \sqrt{4\epsilon + \epsilon^2}] \quad (2.31)$$

$$\in \frac{1}{2}[y_0 + \vec{n} \cdot \vec{t} \pm (\epsilon_{2xy} + \epsilon_{3xy} + \epsilon_{2xy}\epsilon_{3xy} + \epsilon_{2z}\epsilon_{3z} + \sqrt{4\epsilon + \epsilon^2})], \quad (2.32)$$

$$Q_{3,1} = \frac{1}{2}[x_{01} + \vec{x}_1 \cdot \vec{t} + \vec{x}_1 \cdot (\vec{r}_{3xy} - \vec{t})] \in \frac{1}{2}[x_{01} + \vec{x}_1 \cdot \vec{t} \pm \epsilon_{3xy}], \quad (2.33)$$

$$Q_{3,2} = \frac{1}{2}[x_{02} + \vec{x}_2 \cdot \vec{t} + \vec{x}_2 \cdot (\vec{r}_{3xy} - \vec{t})] \in \frac{1}{2}[x_{02} + \vec{x}_2 \cdot \vec{t} \pm \epsilon_{3xy}]. \quad (2.34)$$

From this we can conclude that, as long as  $Q \in B_\epsilon(\mathcal{P})$ , the conditions Eqs. (2.25)-(2.34) should be satisfied. In particular this implies the constraints on the values on the right-hand sides of Eqs. (2.25)-(2.34). For example, Eq. (2.25) together with  $Q_{1,1} \in B_\epsilon(\mathcal{P}_{1,1})$  implies  $(1/2)(x_{01} + \vec{n} \cdot \vec{x}_1) \in \mathcal{P}_{1,1} \pm [\epsilon + (\epsilon_{2xy} + 2\sqrt{4\epsilon + \epsilon^2})]$ . For the other constraints one can proceed similarly. Therefore, given  $\epsilon$  arbitrary, we have to ask whether there exist feasible  $\vec{n}$  and  $\vec{t}$  and  $(x_{10}, \vec{x}_1)$  and  $(x_{20}, \vec{x}_2)$  which are simulable by dichotomic measurements such that all the constraints are satisfied.

Notice that, due to the unitary freedom mentioned earlier, one can always set  $\vec{n} = (1, 0, 0)$ . If we further parametrize  $\vec{t} = (\cos(\varphi), \sin(\varphi), 0)$ , asking for the existence of  $(x_{10}, \vec{x}_1)$  and  $(x_{20}, \vec{x}_2)$  which are simulable by dichotomic measurement, is a SDP. This yields the following algorithm to obtain lower bounds on the distance. Scanning over

the values of  $\vec{t}$  with a certain finite step size in  $\varphi$ , for any value of  $\vec{t}$  we set  $\vec{r}_2 = \vec{n}$ ,  $\vec{r}_1 = -\vec{n}$ ,  $\vec{y} = \vec{n}$  and  $\vec{r}_3 = \vec{t}$ . Set a value for  $\epsilon$  and test the SDP of finding reducible  $M_{1|1}$  and  $M_{2|1}$  such that all the above mentioned constraints are satisfied. Clearly, at  $\epsilon = 0$  the SDP is infeasible as it corresponds to a point of an irreducible correlation. One can implement a bisection method to find the exact transition point where the SDP is infeasible. We obtain the critical error tolerance  $\epsilon_c(\varphi)$ . By scanning over all values of  $\varphi$ , we find  $\epsilon^* := \min_{\varphi} \epsilon_c(\varphi)$ . For practical purposes, the above described procedure is sufficient. In principle one can object that the number  $\epsilon^*$  may not be reliable due to the finite step size scanning over the values of  $\varphi$ . However, this objection can be addressed with the developed theory. The idea is that the error introduced by the finite step size can be bounded by bounding the variation in the function  $\epsilon_c(\varphi)$ . That is, we can find a number  $C > 0$  such that  $|\epsilon_c(\varphi + x) - \epsilon_c(\varphi)| \leq C\delta$  for all  $\varphi \in [0, 2\pi]$  and  $|x| \leq \delta$ . Here it is interesting to observe that the variation of  $\varphi$  only affects the entry  $Q_{3,3}$  of the correlations. A variation of  $\delta$  in  $\varphi$  gives rise to a variation of  $\delta\vec{t}$  with  $|\delta\vec{t}| \leq \delta$ . One then sees that  $\delta Q_{3k} \leq \delta/2$ . Therefore  $|\epsilon_c(\varphi + x) - \epsilon_c(\varphi)| \leq \delta/2$  for any value of  $\varphi$  and  $|x| \leq \delta$ . If one selects a step in  $\varphi$  with size  $\delta$ , the error in the global minimum  $\epsilon^* = \min \epsilon_c(\varphi)$  is bounded by  $\delta/4$  (since the maximum distance from any point to a computed point is  $\delta/2$ ). Taking  $\delta = 2\pi \times 10^{-5}$  is sufficient to bound the error by  $\pi/2 \times 10^{-5}$ . An adaptive scheme of varying the step sizes over different regimes of  $\varphi$  can be utilized to speed up the computation. Applying the algorithm to the distribution  $\mathcal{P} = \mathcal{P}(0.577, 0.726, 0.276)$ , which seemed optimal according to the seesaw iterations, we can lower bound the distance to the set of simulable correlations by  $r_{\infty} \approx 0.0022$ . Although this value is one order of magnitude smaller than the value of the seesaw optimization, it emphasises the consistency of Theorem 14.

## 2.5 Estimating the state space dimension

As discussed, in order to certify an irreducible  $n$ -outcome measurement, knowledge of the dimension of the prepared system is necessary. While it is possible that the dimension can be convincingly deduced from the experimental setup, for a fully device-independent procedure, the dimension of the system has to be determined from correlation data alone. However, the dimension of a system is a physically ill-defined object, in the sense that any description of a system can always be embedded into a higher dimensional system. For example, we can treat a qubit as a restricted theory of a qutrit. But from an operational point of view, one can still assess the dimension by determining the effective dimension, that is, the minimal dimension which explains the experimental data. A dimension witness [141] might seem to be the appropriate tool for this purpose, since it gives a procedure to determine a lower bound on the dimension. However, this is not sufficient for our purposes, because it does not exclude

that the effective dimension can be higher than the dimension witnessed.

In general we assume that the procedure to determine the effective dimension of a system consists of  $s$  different state preparations and  $m$  measurements. The correlations  $p(a|x, y)$  form a matrix  $(\mathcal{A}_{x,\ell})_{x,\ell}$ , where  $x$  labels the states,  $a$  labels the outcome of measurement  $y$ , and  $\ell$  enumerates all outcomes of all measurements. For a  $d$ -level system, the rank of this matrix can be at most the affine dimension of the state space, that is,  $d^2 - 1$ . Hence determining the rank of the matrix  $\mathcal{A}$  can give an estimate of the effective dimension of the system. In practice, one would choose a large number of preparation procedures and a large number of measurement procedures with the expectation, that an estimate of the rank of  $\mathcal{A}$  produces a reliable estimate of the affine dimension of the state space. We mention that for consistency reasons, the preparation- and measurement procedures should include those required to certify the irreducibility of the  $n$ -outcome measurement.

While this approach can work in principle, it has to be considered with care. For an implementation of an irreducible three-outcome measurement on a qubit, it is typically necessary to dilate the three-outcome measurement to a projective measurement on a higher-dimensional system [40]. Despite of this, it still makes sense to speak about an irreducible three-outcome measurement if the additional dimensions used for the dilated measurement are not accessible due to a physical mechanism that reduces the dimension before entering the measurement station. In a setup using the polarization degree of freedom of a photon, this may be achieved, for example, by means of a single mode fiber. Mathematically, such a mechanism corresponds to a completely positive map  $\Phi$  so that the correlations are obtained through  $p(a|x, y) = \text{tr}[\Phi(\rho_x)M_{a|y}]$ . However, then the rank of the matrix  $\mathcal{A}$  alone is insufficient to establish the effective dimension of the system, as can be seen by considering the dephasing qutrit-qutrit channel  $\Phi: \rho \mapsto \sum_k |k\rangle\langle k| \rho |k\rangle\langle k|$ . Using this channel, the matrix  $\mathcal{A}$  will have rank three, which would suggest an effective dimension of  $d = 2$ , while the actual effective dimension is  $d = 3$ . This can be overcome by certifying that the shape of the state- and effect space corresponds to a qubit. For methods to implement such a certification we here refer to Ref. [142, 143].

## 2.6 Experimental feasibility

So far we have derived distances between given irreducible correlations and the set of simulable correlations. However, in experiments an important quantity or benchmark is the fidelity in the state preparations and the measurements. This quantity is typically smaller than 1 due to experimental imperfections and the presence of noise. This raises the question how precise the state preparations must be in order to allow for a demonstration of irreducibility. In order to obtain the smallest possible fidelity

we choose the distribution generated by the states in Eq. (2.20) and the measurements in Eq. (2.21), yielding the distances  $r_2 \approx 0.0686$  and  $r_\infty \approx 0.0342$ . We will denote the resulting distribution by  $\mathcal{P}_t$ . Suppose that the measurements  $M_1, M_2$  are implemented perfectly and are given by Eq. (2.21). As we can see each entry of the correlation table as originating from a two-outcome measurement, i.e., we forget about the structure of the trichotomic measurement, it is sufficient to consider only one entry of the distribution  $\mathcal{P}_t$ . The corresponding dichotomic POVM can thus be described by a single effect  $E$ . Suppose that  $p = \text{Tr}[\rho E]$ . We now want to compute the minimal fidelity between the experimentally implemented state  $\hat{\rho}$  and the theoretical, ideal state which is needed in order to ensure  $|p - \text{Tr}[\hat{\rho} E]| < \delta$ . Note that the ideal states are pure and can be written as

$$|\psi_1\rangle = |0\rangle, \quad |\psi_2\rangle = \frac{1}{2}(|0\rangle - \sqrt{3}|1\rangle), \quad |\psi_3\rangle = \frac{1}{2}(|0\rangle + \sqrt{3}|1\rangle). \quad (2.35)$$

The idea of the proof relies on the following *operational definition* of the fidelity presented in Ref. [144].

**Theorem 19.** *Let  $\rho$  and  $\sigma$  be two quantum states. The fidelity  $\mathcal{F}$  can be written as*

$$\mathcal{F}(\rho, \sigma) = \min_{\{F_i\}} \left[ \sum_{i=1} \sqrt{\text{Tr}[\rho F_i] \text{Tr}[\sigma F_i]} \right]^2, \quad (2.36)$$

where the minimum is taken over all possible POVMs  $\{F_i\}$ .

**Corollary 20.** *Let  $\rho, \sigma$  be quantum states and suppose that  $1 - \mathcal{F}(\rho, \sigma) < \epsilon$ . Then we have*

$$|\text{Tr}[E\rho] - \text{Tr}[E\sigma]| \leq \sqrt{\epsilon} \quad (2.37)$$

for any effect  $E$ . In particular, to obtain a distribution  $\mathcal{Q}$  such that  $\|\mathcal{P}_t - \mathcal{Q}\|_\infty \leq r_\infty$  we need the fidelity between any target states  $|\psi_k\rangle$  and the corresponding experimental state  $\hat{\rho}_k$  to be

$$\langle \psi_k | \hat{\rho}_k | \psi_k \rangle \geq 1 - r_\infty^2 \approx 99.883\%. \quad (2.38)$$

*Proof.* Since it is sufficient to consider only one entry of the distribution  $\mathcal{P}$ , we can view  $q = \text{Tr}[\hat{\rho} E]$  as part of a two-outcome POVM (also for the three-outcome measurement)  $M = (E, \mathbb{1} - E)$  yielding the distribution  $(q, \bar{q}) = (\text{Tr}[\hat{\rho} E], \text{Tr}[\hat{\rho}(\mathbb{1} - E)])$ . Clearly we have

$$\begin{aligned} \mathcal{F}(\rho, \sigma) &= \min_{n \geq 1} \min_{(F_k)_{k=1}^n} \left[ \sum_{k=1}^n \sqrt{\text{Tr}[\rho F_k] \text{Tr}[\sigma F_k]} \right]^2 \leq \min_{(F_k)_{k=1}^2} \left[ \sum_{k=1}^2 \sqrt{\text{Tr}[\rho F_k] \text{Tr}[\sigma F_k]} \right]^2 \\ &\leq \left[ \sqrt{\text{Tr}[\rho E] \text{Tr}[\sigma E]} + \sqrt{\text{Tr}[\rho(\mathbb{1} - E)] \text{Tr}[\sigma(\mathbb{1} - E)]} \right]^2 \\ &=: \left[ \sqrt{pq} + \sqrt{(1-p)(1-q)} \right]^2 =: f(p, q). \end{aligned} \quad (2.39)$$

Consequently, if we know that  $\mathcal{F}(\rho, \sigma) \geq 1 - \epsilon$ , then we also have  $\sqrt{1 - \epsilon} \leq \sqrt{pq} + \sqrt{(1-p)(1-q)}$ . More precisely, we want to solve the following optimization problem for given  $\epsilon > 0$ .

$$\begin{aligned}
& \text{maximize} && \delta(\epsilon) := |p - q| \\
& \text{subject to} && 0 \leq p, q \leq 1 \\
& && \sqrt{1 - \epsilon} \leq f(p, q) \\
& \text{with respect to} && p, q.
\end{aligned} \tag{2.40}$$

One finds that the optimizers are given by  $q = (1 - \sqrt{\delta})/2$ ,  $p = (1 + \sqrt{\delta})/2$  and the optimal value is  $\sqrt{\epsilon}$ . Inserting the value of  $r_\infty$  found for the distribution  $\mathcal{P}_t$  yields the claim.  $\square$

## 2.7 Outlook and discussion

In this Chapter we have studied the structure of the correlations produced by irreducible three-outcome qubit measurements in a prepare-and-measure scenario. Here we have provided a minimal scenario in terms of the number of experimental devices required. Using only one auxiliary measurement, we found that the genuine trichotomic correlations can be separated in the Euclidean norm by  $r_2 \approx 0.0686$ . Here the similarity of our setup to the one in Ref. [134] should be mentioned, namely, the states and the measurements are the same, but in our setup we omit one of the auxiliary measurements. While this scheme is motivated by symmetry considerations, we also used the USD correlations to systematically describe a subset of trichotomic correlations. However, within these correlations the largest distance we obtained is smaller,  $r_2 \approx 0.0347$ . In addition, we have provided a discussion of the minimal state preparation fidelity needed in order to experimentally certify the irreducibility of a three-outcome measurement. Our results are not based on a linear inequality separating genuine and simulable trichotomic correlations and such an inequality is not necessary for the purpose of an experimental certification. However, we also established in Theorem 14 that such an inequality exists when the analysis is constrained to the USD correlations.

For future research it is an interesting open question whether in our minimal scenario this also holds when one considers the set of all correlations, since this would imply that also in the minimal scenario the convex hull of all simulable correlations does not include all trichotomic correlations. A further interesting question is if it is possible to find a systematic approach to construct families of distributions in the minimal scenario that are able to determine whether an implemented  $n$ -chotomic measurement is irreducible. Such an approach would be highly desirable, as it would unify recent results that actually appear in separate contexts.



## 3 Partially observed measurements in the Wigner's friend scenario

The quantum measurement problem points out that the probabilistic state update rule after a measurement in quantum mechanics is in conflict with the natural assumption that the unitary evolution of isolated quantum systems can be applied at any scale. The rule suggests that the action of measuring a quantity is an *absolute event*, meaning that it is the same for *any* observer and a process that cannot be reversed. Using the so-called Wigner friend scenario, it has been shown that the absoluteness of a measurement event imposes strong constraints on the possible observed statistics, which can be violated by the universal validity of unitary quantum evolution. Here we consider the weaker assumption that the measurement event is realized relatively to *one* observer who only partially observed the outcome of a measurement. We propose a protocol to show that this assumption in conjunction with the natural assumptions of no superdeterminism and locality is also not compatible with the universality of the unitary time evolution in quantum theory. This Chapter is based on Project [G].

### 3.1 Motivation

The quantum measurement problem [50] is often illustrated by a thought experiment called Wigner's friend [54], see Fig. 3.1 (a). Here an observer, called the friend Bob, performs a measurement on a physical system in an isolated laboratory and the outcome is reflected by a certain definite position of a pointer device. Further, a second (super)observer called Alice, is placed outside the laboratory and one assumes that she is capable of performing arbitrary quantum operations on Bob's lab. However, while Bob uses the state update rule for describing the state of his system after the measurement and thus assigning the eigenstate corresponding to the observed outcome to his system, Alice describes the lab and all of its content via a unitarily evolving quantum state. This can yield a contradiction between Bob's and Alice's perspective, who assigns a specific entangled state to the system and thus not ascribes a well-defined value to the outcome of Bob's measurement.

Indeed, the state update rule suggests that the action of measuring a quantity is an absolute event, meaning that it cannot be reversed and is the same for any ob-



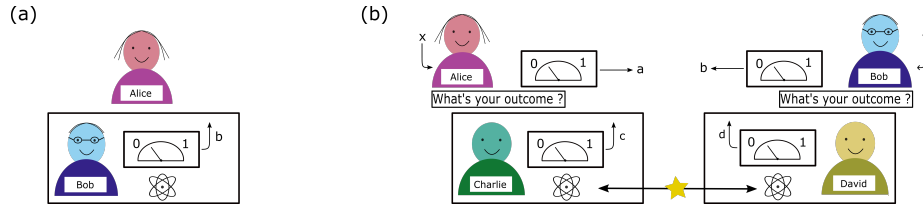


Figure 3.1: (a) Illustration of the Wigner's friend thought experiment. (b) Extended version of the Wigner's friend thought experiment.

server [35–37, 145]. This perception of measurements in quantum theory has been highly debated. While it is supported by collapse models of the measurement process [146, 147], it is not in line with other viewpoints which assume the universal validity of quantum theory [50–53, 148–151]. Assuming universality of quantum mechanics suggests modeling the measurement process by unitary dynamics, which in principle can always be reversed [50]. As a consequence, the measurement may be undone and the value of the measured quantity can be erased, as if it has never existed.

In a series of recent works [35–37], an extended version of the Wigner's friend thought experiment was introduced. It consists of two spatially separated laboratories, each containing an experimenter, accompanied by a superobserver. The two experimenters each hold a half of an entangled particle on which they can perform measurements and it is assumed that the superobservers can perform quantum operations on the respective labs. In this context, Brukner aimed to formalize the assumption of absolute measurement events, which results in the notion of *observer independent facts* (OIF) [36, 37]. This means that measurement results exist with respect to *any* observer regardless of whether the measurement has been actually performed or not. As it turns out, one can put strong constraints on the observable statistics in this extended scenario by combining OIF with two seemingly natural assumptions. First, the assumption of *no superdeterminism*, that guarantees that the choice of measurement settings is uncorrelated with any relevant variables prior to that choice [5]. Second, the assumption of *locality* in the sense of parameter independence [5, 152], which prohibits the influence of a local setting on a distant outcome<sup>1</sup>. However, the set of resulting correlations coincides with those allowed by local realism and could be derived from no superdeterminism and OIF alone without taking into account locality [153, 154].

As it turns out, OIF can be replaced by a weaker assumption [37], such that the set of admissible correlations contains instances which are not allowed by local realism. This assumption is called *absoluteness of observed events* (AOE) and entails that a mea-

<sup>1</sup>More precisely, the assumption of signal locality entails that the probability of Bob observing a particular outcome  $b$  given setting  $y$  is statistically independent of the measurement choice  $x$  of Alice. Note that this locality is a weaker assumption than *local causality* where one additionally assumes that the probability for the event  $(b, y)$  is independent of  $x$  and Alice's outcome  $a$ , yielding an LHV model.

surement result persists with respect to *any* observer only for the measurement that has been *actually* performed [37]. However, one may argue that the absoluteness of a measurement event with respect to *every* observer, that is, all observers exist on par, is too strong of an assumption.

In this Chapter, we further relax the assumption of AOE and require absoluteness of a measurement only with respect to a particular observer, which is distinguished from other observers by partially having access to the realized measurement outcome. The assumption that the event of measurement is realized with respect to *that* observer is called *relative event by incomplete information*, short REII. We show that REII can also be rejected by the universality of quantum mechanics under the assumptions of no superdeterminism and locality. We begin by formalizing the assumptions and introducing a minimal version of the extended Wigner’s friend scenario in Section 3.2. We proceed by combining the assumptions in Section 3.3, which allows us to derive the correlation polytope for the minimal scenario. In addition, we characterize the set of quantum correlations and present possible quantum violations of the correlation polytope. Afterwards, we discuss in Section 3.4 the consequences and subtleties and introduce an extended protocol, where we allow the measurement of Charlie to have four outcomes. Finally, we present three alternative protocols in Section 3.5, which are used to investigate the relation between the different sets of correlations.

## 3.2 A minimal protocol and the assumptions

### 3.2.1 The scenario

To illustrate the idea we consider first a minimal protocol. Note that also in the EWFS it is in principle possible to remove the friend of Bob without affecting the conclusion of the experiment. Consider two parties Alice and Bob sharing two particles at different locations. Alice stores her particle in a laboratory, in which she has another party named Charlie playing the role of Wigner’s friend. In each run of the protocol, Charlie performs a measurement with three possible outcomes  $c \in \{0, 1, 2\}$  on the particle. After the measurement is completed, Alice receives a signal  $x \in \{0, 1, 2\}$ . Given the signal  $x$ , she asks Charlie if his measurement outcome is  $c = x$ . If this is indeed the case, Alice outputs  $a = x$  (and thus  $a = c$ ) as her measurement outcome. In the other case  $c \neq x$  she makes a binary outcome measurement on the particle and uses the obtained outcome to decide the output among  $\{0, 1, 2\}$  with the constraint that  $a \neq x$ . On the other hand, Bob receives a signal  $y \in \{0, 1\}$  in each run of the protocol, based on which he chooses one of two measurements to perform on his particle in order to obtain a binary outcome  $b \in \{0, 1\}$ . Over many runs of the protocol, the collected data allows one to compute the statistics  $p(a, b|x, y)$ , which is observed by Alice and Bob.

In the following we will derive in a similar manner as in the EWFS explicit inequal-

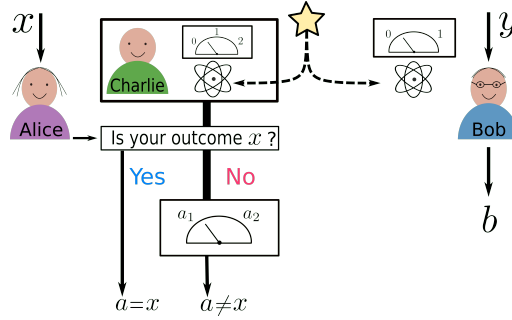


Figure 3.2: Illustration of the protocol. A pair of particles are distributed to Charlie and Bob. Charlie, playing the role of Wigner's friend in a closed laboratory, carries out a measurement with outcomes 0, 1, 2. Alice, playing the role of Wigner, receives a signal  $x \in \{0, 1, 2\}$ . She asks Charlie if his outcome is  $x$  or not. If it is  $x$ , Alice uses it as her output  $a$ . Otherwise, Alice continues to carry out a binary outcome measurement on the system and the whole laboratory to obtain an outcome  $a \in \{0, 1, 2\}$  but  $a \neq x$ . Bob receives a signal  $y \in \{0, 1\}$  and performs correspondingly one of two possible binary outcome measurements on his particle and output the outcome  $b$ . The figure is taken from Ref. [G].

ities based on the assumption that the measurement event is realized to an observer who only partly gains the information about the outcomes (here Alice) combined with other natural assumptions. As we will see, these inequalities can be violated if one assumes that quantum theory is universally valid.

### 3.2.2 Relative event by incomplete information

By construction of our protocol, Alice has obtained some part of information about the outcomes of Charlie's measurement in each of the runs of the experiment. Indeed, even in the case  $x \neq c$  Alice has learned that the outcome of Charlie was not  $x$ . As Alice holds partial information of the outcome, we can directly apply REII to the scenario, which implies that the value of  $c$  is realized in every run of the protocol. Therefore, for any given measurement choice  $x$  and  $y$  in each of the runs, Alice can assume the existence of a joint distribution  $\mathfrak{p} = \mathfrak{p}(a, b, c|x, y)$  such that the observed data  $p(a, b|x, y)$  is given by marginalizing the irrelevant outcome  $c$ , that is,

$$p(a, b|x, y) = \sum_c \mathfrak{p}(a, b, c|x, y). \quad (3.1)$$

Here and in the following we use  $p$  to indicate that  $p(a, b|x, y)$  is experimentally accessible, rather than being hypothetical like  $\mathfrak{p}(a, b, c|x, y)$  in gothic type. It should be noted that the inquiry of Charlie's outcome by Alice is crucial in order to guarantee

the existence of  $p$ . Indeed, if we are to apply our assumption REII to the EWFS, the existence of  $c$  is not implied when Alice does not inquire Charlie at all.

Since  $c = x$  implies  $a = x$ , we directly obtain that  $p(a = x|c, x, y) = \delta_{xc}$ . This means that the existence of the variable  $c$  can be consistently revealed when it is read. By using

$$p(a = x, b|c, x, y) = p(b|a = x, c, x, y)p(a = x|c, x, y), \quad (3.2)$$

this then implies that

$$p(a = x, b|c, x, y) = \delta_{xc}p(b|a = x, c, x, y) = \delta_{xc}p(b|c, x, y). \quad (3.3)$$

Similar to the assumption of AOE or the assumption of realism it is unknown how to reject the thesis of REII purely by its own. However, when combining with two other seemingly natural assumptions, namely freedom of choice (or no superdeterminism) and locality, it put stringent constraints on the observable statistics  $p(a, b|x, y)$ .

### 3.2.3 Freedom of choice and locality

Even we have already discussed these assumption in the context of Bell inequalities and LF-correlations, we will present them here again for the particular scenario under investigation. The assumption of freedom of choice demands that the random inputs  $x$  and  $y$  are independent from the variable  $c$ , that is,  $p(c|x, y) = p(c)$ . This allows us to write

$$p(a, b, c|x, y) = p(c|x, y)p(a, b|c, x, y) = p(c)p(a, b|c, x, y), \quad (3.4)$$

and consequently

$$p(a, b|x, y) = \sum_c p(a, b|c, x, y)p(c). \quad (3.5)$$

The assumption of freedom of choice is justified if Charlie makes the measurement before Alice and Bob make the choices  $x$  and  $y$  respectively. This relies on the honesty of Charlie and the functionality of his device. Interestingly, part of this problem can also be addressed with our protocol, as we will see later.

The locality assumption implies that Bob's measurement result  $b$  does not depend on Alice's input  $x$  and there is an analogous independence between  $a$  and  $y$ . Then

$$p(a|c, x, y) = p(a|c, x) \quad \text{and} \quad p(b|c, x, y) = p(b|c, y). \quad (3.6)$$

It should be emphasized that this notion of signal locality is weaker than the so-called local causality [5] and the probabilities  $p(a|b, c, x, y)$  and  $p(b|a, c, x, y)$  in general cannot be further simplified. We have deliberately used the locality notion from Ref. [37], resembling that of the local friendliness. Had we used the stronger assumption such as local causality, Bell-like correlations were to be obtained [35–37]. We emphasize that, in either case the assumption of REII is used instead of that of AOE.

### 3.3 Combining the assumptions

The combination of the assumptions of freedom of choice, locality and REll may, in accordance with the terminology developed in Ref. [37], be called local friendliness under incomplete information (LFIC). Using this combination, one can put strong constraints on probability distribution  $p(a, b|x, y)$ .

**Theorem 21.** *All probability distribution that satisfy the LFIC assumption are of the form*

$$p(a, b|x, y) = \sum_c \mathfrak{p}_{NS}(a, b|c, x, y)\mathfrak{p}(c), \quad (3.7)$$

$$p(a = x, b|x, y) = \sum_c \delta_{xc}\mathfrak{p}(b|c, y)\mathfrak{p}(c), \quad (3.8)$$

where  $\mathfrak{p}_{NS}(a, b|c, x, y)$  is a probability distribution constrained only by the assumption of locality and freedom of choice.

*Proof.* We will prove the claim for the case where Alice asks Charlie for the confirmation whether his output is 0 when her input was  $x = 0$ . The assumption of REll imposes that

$$p(a, b|x, y) = \sum_c \mathfrak{p}(a, b, c|x, y), \quad \text{and} \quad \mathfrak{p}(a = 0|c, 0, y) = \delta_{0c}. \quad (3.9)$$

In addition, freedom of choice and locality yield

$$\mathfrak{p}(c|x, y) = \mathfrak{p}(c), \quad \mathfrak{p}(a|c, x, y) = \mathfrak{p}(a|a, x), \quad \mathfrak{p}(b|c, x, y) = \mathfrak{p}(b|c, y). \quad (3.10)$$

Taking these equations together yields the claim.  $\square$

On the one hand, Eq. (3.7) can be understood as a direct consequence of Eq. (3.5) and Eq. (3.6). On the other hand, Eq. (3.8) requires the consistency of measurement outcomes when they are read in Eq. (3.3) and Eq. (3.7). Further it is interesting to see that Eq. (3.7) and Eq. (3.8) are a combination of the no-signaling model [78,87] and the local hidden variable model [155]. To mimic the scenario described in Ref. [37] we can adjust the protocol in order to reproduce the original LF model as follows: To respond to Alice's query, her friend Charlie just outputs the outcome of his measurement, which is also Alice's final output in the next step of the protocol. In this case, Eq. (3.7) and Eq. (3.8) reduce to

$$p(a, b|x, y) = \sum_c \delta_{ac}\mathfrak{p}(b|c, y)\mathfrak{p}(c) \quad (3.11)$$

for all values of  $x$ , which is the original LF model. In this case, the LF model is also a local hidden variable model as there is effectively only a single measurement on Alice's side. However, the LFIC correlation polytope is strictly larger than the LF polytope. This inclusion property can be illustrated by choosing a cross section in correlation space, such that both LHV and LF correlations are empty, while the LFIC correlations are not. For more details see Fig. 3.3.

### 3.3.1 The correlation polytope

The correlations that are allowed by Theorem 21 form a polytope, to which there are 60 facets. Among those, 28 coincide with the facets of the no-signaling polytope. The remaining 32 facets can be grouped into four equivalence classes by considering the symmetry between measurements and outcomes, see also Section 1.2.3. Among these facets, two classes only involve at most one measurement of Bob. Consequently, these inequalities cannot separate between the no-signaling model and the local hidden variable model and thus they are omitted. For the remaining two classes, for each we can choose a representative inequality given by

$$\begin{aligned} Z_1 = & p(A_0 = 0, B_0 = 0) + p(A_1 = 1, B_1 = 0) - p(A_2 = 1, B_0 = 0) \\ & + p(A_2 = 1, B_1 = 1) \geq 0, \end{aligned} \quad (3.12)$$

$$\begin{aligned} Z_2 = & p(A_0 = 1, B_0 = 0) + p(A_0 = 2, B_1 = 1) + p(A_1 = 1, B_0 = 1) \\ & - p(A_1 = 1, B_1 = 1) \geq 0, \end{aligned} \quad (3.13)$$

$$Z_3 = p(A_0 = 0, B_0 = 1) + p(A_1 = 1, B_0 = 1) - p(A_2 = 1, B_0 = 1) \geq 0, \quad (3.14)$$

$$Z_4 = p(A_0 = 1, B_0 = 1) + p(A_0 = 2, B_0 = 1) - p(A_1 = 1, B_0 = 1) \geq 0. \quad (3.15)$$

Additionally, the polytope is also constrained by three inequivalent hyperplanes which do not stem from the no-signaling polytope

$$-p(A_0 \neq 0) + p(A_1 = 1) + p(A_2 = 2) = 0, \quad (3.16)$$

$$-p(A_0 \neq 0, B_0 = 1) + p(A_1 = 1, B_0 = 1) + p(A_2 = 2, B_0 = 1) = 0, \quad (3.17)$$

$$-p(A_0 \neq 0, B_1 = 1) + p(A_1 = 1, B_1 = 1) + p(A_2 = 2, B_1 = 1) = 0. \quad (3.18)$$

However, it turns out that the inequalities from Eq. (3.14) to Eq. (3.18) also hold if quantum theory is assumed. Therefore, the only nontrivial inequalities are  $Z_1$  and  $Z_2$ . Note that the inequality in Eq. (3.12) includes all three possible measurements on Alice's side, while Eq. (3.13) only includes two measurements on Alice's side. The inequality Eq. (3.12) resembles the CHSH inequality given in Eq. (1.142), if the event  $A_2 = 2$  is never realized, that is, whenever the query is  $A_2$ , Charlie always replies with a negative answer. Similarly, the inequality in Eq. (3.13) reduces to the CHSH inequality if the event  $A_0 = 0$  is never realized.

In order to visualize the polytope, one can compute a two-dimensional cross section of the polytope with a plane defined by three points  $Q_1, Q_2$ , and  $N_0$ . These points are given by the following probability assignment to the measurement choices and respective outcomes where  $\alpha = (\sqrt{2} - 1)/4\sqrt{2}$ ,  $\beta = (\sqrt{2} + 1)/4\sqrt{2}$ . The point  $N_0$  can be obtained within the no-signaling model while the points  $Q_1, Q_2$  can be realized within

$N_0$	$\mathbb{1}$	$A_0 = 1$	$A_0 = 2$	$A_1 = 1$	$A_1 = 2$	$A_2 = 1$	$A_2 = 2$
$\mathbb{1}$	1	1/2	1/2	1/2	0	1/2	1/2
$B_0 = 1$	1/2	0	1/2	0	0	0	1/2
$B_1 = 1$	1/2	0	1/2	1/2	0	0	1/2
$Q_1$	$\mathbb{1}$	$A_0 = 1$	$A_0 = 2$	$A_1 = 1$	$A_1 = 2$	$A_2 = 1$	$A_2 = 2$
$\mathbb{1}$	1	1/2	0	1/2	0	1/2	0
$B_0 = 1$	1/2	$\alpha$	0	$\alpha$	0	$\alpha$	0
$B_1 = 1$	1/2	$\beta$	0	$\beta$	0	$\alpha$	0
$Q_2$	$\mathbb{1}$	$A_0 = 1$	$A_0 = 2$	$A_1 = 1$	$A_1 = 2$	$A_2 = 1$	$A_2 = 2$
$\mathbb{1}$	1	1/2	1/2	1/2	0	0	1/2
$B_0 = 1$	1/2	$\beta$	$\alpha$	$\alpha$	0	0	$\alpha$
$B_1 = 1$	1/2	$\beta$	$\alpha$	$\beta$	0	0	$\alpha$

Table 3.1: The probability distributions  $N_0$ ,  $Q_1$ ,  $Q_2$  for the cross section in Fig. 3.3. The table is taken from Ref. [G].

quantum theory. In fact, the point  $Q_1$  is one optimal solution for inequality Eq. (3.12) in the sense that it yields the maximal allowed violation among all correlations allowed by quantum theory. In a similar manner the point  $Q_2$  is an optimal solution for inequality Eq. (3.8). In order to obtain a comparison between the different models, we have presented the cross section with the polytope of statistics which is constrained only by the locality condition, see Fig. 3.3. We refer to this polytope as the no-signaling polytope. As expected, this polytope contains the polytope of the LFIC statistics. Interestingly, the LF polytope, which coincides for this choice of cross section with the LHV polytope, has no intersection at all with this plane.

### 3.3.2 The quantum violation

In Fig. 3.3 we also present the set of quantum correlations allowed by quantum mechanics if it is assumed to be universally valid. Its boundary is computed by the NPA hierarchy up to the second level. It clearly indicates that quantum theory violates the LFIC model. Accidentally, although this particular cross section expresses a symmetry between the inequalities  $Z_1$  and  $Z_2$ , the eventual inequivalence between them is revealed by the boundary of the quantum violation.

For a concrete physical realization consider the case where Alice and Bob share a qutrit-qubit system prepared in the state  $|\psi_{t_0}\rangle = (1/\sqrt{2})(|00\rangle + |11\rangle)$ . Alice's qutrit is stored in the laboratory and controlled by her friend Charlie. Charlie performs a measurement in the computational basis  $\{|0\rangle, |1\rangle, |2\rangle\}$ . We further assume that Charlie is in a ready state  $|R\rangle$  before he implements the measurement. If one now assumes

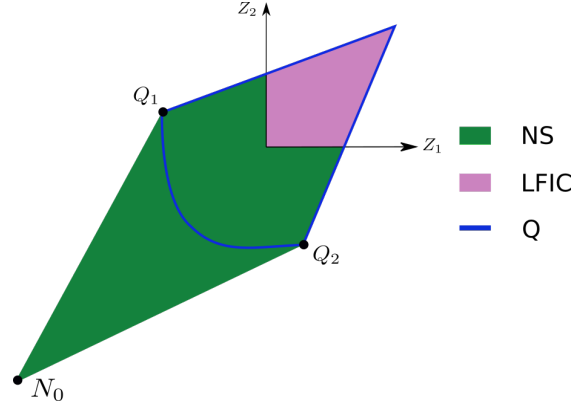


Figure 3.3: Cross sections of different models with a plane containing the points  $Q_1, Q_2$ . The LFIC correlations (purple) are a subset of the NS correlations (green). The blue line bounds the set of correlations allowed by quantum theory within this section. Note that all correlations on this plane cannot be described by a LHV or LF model. The figure is taken from Ref. [G].

that quantum theory is universally valid, i.e., the measurement process is described via the Schrödinger equation, the state of the joint system of Charlie and the qutrit-qubit system after the measurement is given by

$$|\psi_{t_1}\rangle = \frac{1}{\sqrt{2}}(|0\rangle_M \otimes |00\rangle_{AB} + |1\rangle_M \otimes |11\rangle_{AB}), \quad (3.19)$$

where the first term in the tensor products with the subscript  $M$  stands for the measurement device of Charlie, while the latter refers to the qutrit-qubit system shared by Alice and Bob. Upon receiving the random input  $x$ , Alice asks Charlie whether he has observed outcome  $x$ . In the case of confirmation, Alice simply outputs  $x$  as her final output. In the case that Charlie's outcome is not  $x$ , Alice performs measurements depending on the value of  $x$  as follows. In the case  $x = 0$  or  $x = 1$ , Alice can perform an arbitrary measurement and outputs the corresponding outcome. This is due to the fact that the events  $A_0 \neq 0$  and  $A_1 \neq 1$  do not appear in inequality  $Z_1$  and thus the particular choice of those measurements does not affect the violation. However, the event  $A_2 \neq 2$  is present in  $Z_1$  as it contains the term  $A_2 = 1$ . Therefore, if  $x = 2$ , Alice makes a unitary evolution to disentangle Charlie and his device from the qutrit-qubit system, bringing the latter back to

$$|\psi_{t_1}\rangle \mapsto |\psi_{t_2}\rangle = |R\rangle_M \otimes \frac{1}{\sqrt{2}}(|00\rangle_{AB} + |11\rangle_{AB}), \quad (3.20)$$

and thus effectively undoing Charlie's measurement. Alice then performs a measurement of  $\sigma_1$  on the qutrit in the subspace spanned by  $|0\rangle$  and  $|1\rangle$ . Bob is making one of



the two measurements

$$B_0 = \frac{1}{\sqrt{2}}(-\sigma_1 - \sigma_3), \quad B_1 = \frac{1}{\sqrt{2}}(-\sigma_1 + \sigma_3). \quad (3.21)$$

This setup allows for a maximal violation of inequality  $Z_1$  in Eq. (3.12) given by  $Z_1 = (1 - \sqrt{2})/2 \approx 0.2071$ , corresponding to the point  $Q_1$  in Fig. 3.3. This violation is robust with respect to mixing white noise to the state via  $\varrho(p) = p|\psi_{t_0}\rangle\langle\psi_{t_0}| + (1-p)\mathbb{1}/6$  as long as  $p \geq (2/17)(3\sqrt{2} + 1) \approx 0.616781$ . Similar, the point  $Q_2$  can be obtained by using the state  $|\psi_{t_0}\rangle$  and the measurement directions

$$|A_0 = 0\rangle = |A_1 = 2\rangle = |A_2 = 1\rangle = |2\rangle, \quad (3.22)$$

$$|A_0 = 1\rangle = |A_2 = 0\rangle = |0\rangle, \quad (3.23)$$

$$|A_0 = 2\rangle = |A_2 = 2\rangle = |1\rangle, \quad (3.24)$$

$$|A_1 = 0\rangle = |-\rangle, \quad |A_1 = 1\rangle = |+\rangle, \quad (3.25)$$

$$B_0 = \frac{1}{\sqrt{2}}(\sigma_1 - \sigma_3), \quad B_1 = \frac{1}{\sqrt{2}}(-\sigma_1 - \sigma_3). \quad (3.26)$$

### 3.4 Subtleties and an extended protocol

In the protocol introduced we have assumed that Charlie makes the measurement first and then asks for the outcome. In principle, this assumption does not hold if Charlie is not honest or his device does not function properly. For instance, Charlie can wait for the inquiry from Alice and then implement the measurement, or Charlie can answer without referring to the exact measurement outcome. In principle, Alice can always open the box to check whether Charlie's answer is consistent with the outcome of the outcome of his measurement device or not. Thus, we can in principle assume that Charlie's answer is consistent with the outcome of the measurement device. However, this cannot guarantee the no-superdeterminism assumption if the outcome of the measurement, which Charlie supposedly implemented before can be impacted by Alice's inquiry. Therefore one has to propose another protocol to fix this loophole and to make sure that Charlie's answer reflects the outcome of the measurement, which does not depend on Alice's input used in the statistics. The protocol works as follows: For each run of the experiment,

- (1) Charlie makes a measurement and obtains one of four possible outcomes  $c \in \{0, 1, 2, 3\}$
- (2) Alice receives a random number  $t \in \{0, 1, 2, 3\}$  and inquires Charlie whether  $c$  equals  $t$
- (3) Alice receives a random number  $x \in \{0, 1, 2, 3\}$  and inquires Charlie whether  $c$  equals  $x$

(3.1) if  $c = x$ , Alice outputs  $a = x$

(3.2) if  $c \neq x$ , Alice continues to make a measurement and obtains an outcome  $a \in \{0, 1, 2, 3\} \setminus \{x\}$

In each run, Bob receives an input  $y \in \{0, 1\}$ , makes a measurement and obtains an outcome  $b \in \{0, 1\}$ . The statistics after many runs is collected to estimate the outcome probabilities  $p(a, b|x, y)$ .

According to this protocol, although  $c$  could in principle depend on  $t$ , it is independent of  $x$  under the assumption of REll and no-superdeterminism. In the special case that  $c \neq t$  and  $x \neq t$  for any fixed  $t$ , the LFIC model of the extended protocol just reduces to our main protocol. As already discussed, this model cannot describe the statistics predicted by quantum theory.

## 3.5 Further protocols

In addition to the protocols mentioned so far, we have two protocols. In difference to those protocols, here also the LF polytope has an intersection with the plane. It turns out that the LFIC model for the corresponding protocol differs from LHV model, the LF model as well as the NS model.

### 3.5.1 Alternative protocol I

For each of the runs,

(1) Charlie makes a measurement and obtains outcome  $c \in \{0, 1, 2\}$

(2) Alice and Bob receive the inputs  $x, y$  respectively, where  $x, y \in \{0, 1\}$

(3.1) If  $x = 0$ , Alice asks Charlie whether  $c = 0$

– If  $c = 0$ , Alice uses 0 as her outcome  $a$

– If  $c \neq 0$ , Alice continues to make a measurement and obtains an outcome  $a \in \{1, 2\}$

(3.2) If  $x = 1$ , Alice makes a measurement with two outcomes  $\{0, 1\}$  and outputs the outcome  $a$

(4) Independently of  $y$ , Bob makes a measurement with two outcomes  $\{0, 1\}$  and outputs the outcome as  $b$

The statistics after many runs is collected to estimate  $p(a, b|x, y)$ . One finds that up to permutations of the measurements and permutations of the outcomes, the only

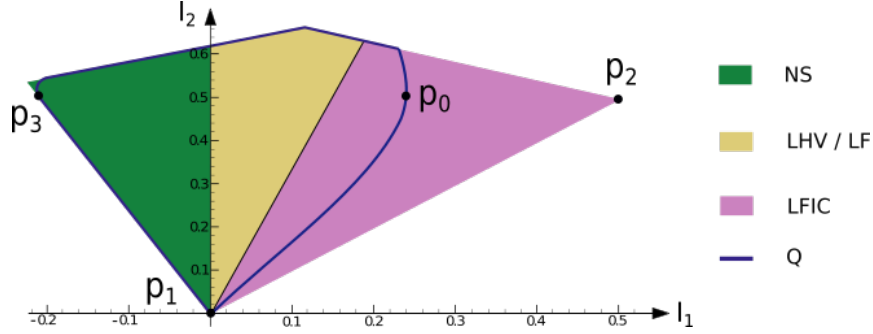


Figure 3.4: Cross sections of different models on the plane spanned by points  $p_1, p_2, p_3$ , where  $I_1$  is the quantity defined in the left hand-side of Eq. (3.27) and  $I_2$  is obtained by switching the measurements  $B_0$  and  $B_1$  in  $I_1$ . Notice that the hierarchical structure of the sets is in accord with our previous discussion. The figure is taken from Ref. [G].

non-trivial tight inequality of the correlation polytope is

$$I_1 = p(A_0 = 0, B_0 = 0) + p(A_0 \neq 0, B_1 = 0) + p(A_1 = 1, B_0 = 0) - p(A_1 = 1, B_1 = 0) \geq 0, \quad (3.27)$$

which is in the form of the CHSH inequality. Therefore the quantum violation can be  $(1/2)(1 - \sqrt{2})$  with the setting

$$|\psi\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle), \quad A_0 = \sigma_3, \quad A_1 = \sigma_1, \quad B_0 = \frac{\sigma_1 + \sigma_3}{\sqrt{2}}, \quad B_1 = \frac{-\sigma_1 + \sigma_3}{\sqrt{2}}, \quad (3.28)$$

where their outputs 0, 1 are mapped to the eigenvalues  $\pm 1$  and the event  $A_0 = 0$  never happens, i.e.,  $p(A_0 = 2) = 0$ . The relation between the different models can be pointed out more clearly in their cross sections with the plane spanned by the correlations  $p_1, p_2, p_3$  as in Fig. 3.4, where

$$\begin{aligned} p_1 &= (1, 1, 1, 1, 0, 1, 0, 1, 0, 1, 1, 1), \\ p_2 &= (1/2)(2, 1, 1, 1, 1, 0, 1, 1, 0, 1, 1, 1), \\ p_3 &= (1/2)(2, 1, 1, 1, 0, a, 0, a, 0, 1, a, b), \end{aligned} \quad (3.29)$$

where  $a = (2\sqrt{2} + 2)/4\sqrt{2}$  and  $b = (2\sqrt{2} - 2)/4\sqrt{2}$ . Here it is important to note that the statistics  $p_3$  can be generated by using the quantum setting mentioned in Eq. (3.28). In an experiment, the observation of the point  $p_3$  directly falsifies the LFIC model. Similarly, the observation of a point such as  $p_0$  falsifies the LF model, but can still be explained by the LFIC model.

### 3.5.2 Alternative protocol II

The main purpose of this protocol is to motivate another interesting protocol (see alternative protocol III). To this end, we add one dichotomic measurement for Alice without querying Charlie. For each of the runs

- (1) Charlie makes a measurement and obtains outcome  $c \in \{0, 1, 2\}$
- (2) Alice and Bob receive the inputs  $x, y$  respectively, where  $x \in \{0, 1, 2, 3\}$  and  $y \in \{0, 1\}$
- (3) If  $x \neq 3$ , Alice inquiries Charlie whether  $c$  is equal to  $x$ 
  - If  $c = x$ , Alice uses  $c$  as her outcome  $x$
  - If  $c \neq x$ , Alice continues to make a measurement and obtains an outcome  $a \in \{0, 1, 2\} \setminus \{c\}$
- (4) If  $x = 3$ , Alice makes a measurement and obtains an outcome in  $\{0, 1\}$  and outputs it as  $a$
- (5) Independently of  $y$ , Bob makes a measurement with two outcomes  $\{0, 1\}$  and outputs the outcome as  $b$

The statistics after many runs is collected to estimate  $p(a, b|x, y)$ . One finds that there are 10 different classes of linear constraints and 7 of them wither do not have two measurements on Bob's side or do have have  $A_3$  on Alice's side. In two of the remaining three classes, only one measurement among  $A_1, A_2, A_3$  appears, which is already the case in the previous alternative protocol. The single left class is

$$- p(A_0 \neq 0, B_0 = 1) + p(A_0 \neq 0, B_1 = 1) + p(A_1 = 1, B_0 = 1) \quad (3.30)$$

$$- p(A_1 = 1, B_1 = 1) + p(A_3 = 1, B_0 = 1) + p(A_3 = 0, B_1 = 0) \geq 0. \quad (3.31)$$

If one identifies the measurements  $A_0, A_1$  and  $A_2$ , i.e., one does not distinguish between the input label, Eq. (3.30) reduces to

$$\begin{aligned} & - p(A_0 = 2, B_0 = 1) + p(A_0 = 2, B_1 = 1) \\ & + p(A_3 = 1, B_0 = 1) + p(A_3 = 0, B_2 = 0) \geq 0, \end{aligned} \quad (3.32)$$

which resembles the CHSH inequality. Therefore, the inequality in Eq. (3.30) can be violated in quantum theory. In comparison to the protocols so far, an extra measurement can indeed more flexibly reveal the difference between the models, e.g., we can even choose the same setting for all the measurements  $A_0, A_1, A_2$  in quantum theory. The inequality in Eq. (3.30) inspires another protocol, where Alice can ignore the information whether she has asked for  $A_0, A_1$  or  $A_2$  and simply group them together as one single measurement which is also called  $A_0$ .

### 3.5.3 Alternative protocol III

For convenience, we relabel the measurement of Alice without querying (former measurement  $A_3$ ) as  $A_1$ . Although this protocol is very similar to that of Ref. [37], here Alice infers in each run only one outcome of Charlie instead of the whole set of outcomes. We denote by LFIC<sub>1</sub>, LFIC<sub>2</sub> and LFIC<sub>3</sub> the protocols where Alice can ask one, two and three outcomes respectively. If Alice can only ask one outcome, this protocol coincides with alternative protocol I.

For each run of the experiment

- (1) Charlie makes a measurement and obtains an outcome  $c \in \{0, 1, 2\}$
- (2) Alice and Bob receive the inputs  $x, y$  respectively, where  $x, y \in \{0, 1\}$
- (3) If  $x = 0$ , Alice chooses  $t = 0$  in the case of LFIC<sub>1</sub>, a random number  $t \in \{0, 1\}$  in the case of LFIC<sub>2</sub> and a random number  $t \in \{0, 1, 2\}$  in the case of LFIC<sub>3</sub>. Then she asks Charlie whether  $c$  is a realization of  $t$ 
  - If  $c = t$ , Alice uses  $t$  as her outcome
  - If  $c \neq t$ , Alice continues to make a measurement and obtains an outcome  $\{0, 1, 2\} \setminus \{t\}$
- (4) If  $x = 1$ , Alice makes a measurement with two outcomes  $\{0, 1\}$  and outputs the outcome  $a$
- (5) Independently of  $y$ , Bob makes a measurement with two outcomes  $\{0, 1\}$  and outputs the outcome as  $b$

The statistics after many runs is collected to estimate  $p(a, b|x, y)$ . Notice that the values of  $t$  are discarded. To compare different models in this scenario, we firstly take the cross section of the polytopes with the plane determined by points  $p_2, p_3$  given by Eq. (3.29) as well as

$$p_4 = (1/2)(2, 1, 1, 0, 1, 0, 1, 0, 1, 1, 0, 1). \quad (3.33)$$

We then express the cross section in the coordinates defined by  $I_2$  and  $I'_1$  where

$$\begin{aligned} I'_1 = & p(A_0 \neq 1, B_0 = 1) + p(A_0 \neq 0, B_1 = 0) + p(A_1 = 1, B_0 = 0) \\ & - p(A_1 = 1, B_1 = 0) \geq 0. \end{aligned} \quad (3.34)$$

In fact  $I'_1$  corresponds to a facet of LFIC<sub>2</sub>. Interestingly, LFIC<sub>1</sub> = NS in this scenario, which is also reflected in Fig. 3.5. In fact, if Alice always queries the realization of 1 (or 2) instead of 0 in the LFIC<sub>1</sub> model, this yields new polytopes which are equivalent to the one of LFIC<sub>1</sub> model up to certain rotation and translation. By definition, LFIC<sub>3</sub> is the convex hull of the union of those three polytopes corresponding to Alice's query

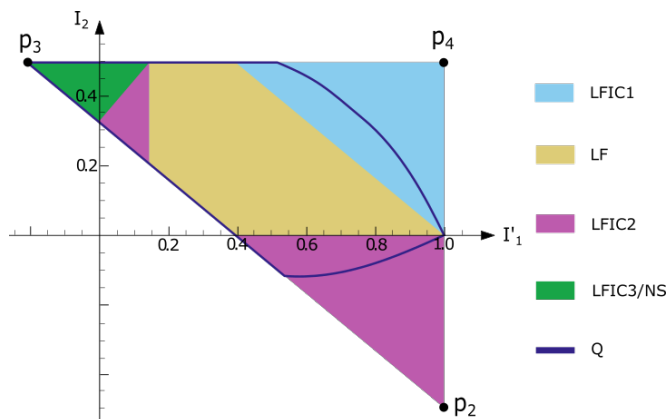


Figure 3.5: Cross sections of different models with the plane spanned by points  $p_2, p_3, p_4$ , where  $I'_1$  is the quantity in Eq. (3.34) and  $I_2$  is obtained by switching the measurements  $B_0, B_1$  in  $I_1$  in Eq. (3.27). Notice that the hierarchical structure of the sets is in accord, i.e.,  $\text{LHV} \subset \text{LF} \subset \text{LFIC} \subset \text{NS}$ . The figure is taken from Ref. [G].

for being 0, 1 and 2. This implies that any point of the NS polytope can be written as a mixture of  $\text{LFIC}_1$  correlations. As it turns out, this is still the case if there is at least one measurement from Alice or Bob having 2 outcomes. Surprisingly, when all the measurements have 3 outcomes, there are points of the NS polytope which cannot be covered by the mixture of those three  $\text{LFIC}_1$  polytopes. This hints that the structure of the NS conditions and NS polytope in this scenario also deserves further investigation. Besides,  $\text{LFIC}_3 = \text{NS}$  implies that there is no quantum violation, as any quantum correlation can be explained by the  $\text{LFIC}_3$  model.

### 3.6 Discussion of the results

The violation of the inequalities given by Eq. (3.12) and Eq. (3.13) by quantum theory is due to the fact that the qutrit at Alice's side still maintains entanglement with Bob's qubit after Charlie responds negatively to the question "is  $x = 2$ ". This points towards a relevant discussion about the nature of the measurement process. Originally, according to von Neumann [41] (see also the discussion in Section 1.1.2), a measurement of a degenerate observable leads to full decoherence in the entire eigenbasis of the observable, such that the post measurement state is diagonal in this basis. It was later recognized by Lüders [42] that this is not the appropriate formulation. If a degenerate measurement is made, then according to the Lüders rule, the coherence within the degenerate subspace is unaffected.

Our analysis of the assumption of relative event by incomplete information also

highlights the difference between the viewpoints of von Neumann and Lüders. If we treat Alice's query about  $j$  and Charlie's measurement as a single measurement  $\tilde{c}_j$  implemented and read by Alice, it constitutes a dichotomic degenerate measurement of the qutrit. The post measurement state following Lüders' rule still has quantum coherence in a two-dimensional subspace and this entanglement with a remote party can still remain. In comparison, we see that von Neumann's rule is similar to the assumption of REII.

The violation of our inequalities can also be seen as related to an interesting remark by Peres in connection with the concept of contextuality [99, 101, 156]. There, in order to justify the contextuality assumption, he argued that it is essential that for a three-outcome measurement, Alice can construct a device to measure whether the system gives some outcome, e.g., whether it is 2 or not, and then, at a later stage or even in a different laboratory, decide how to complete the measurement.

Apart from the more sophisticated protocols that we have discussed, another interesting extension would be to consider this quantum correlation sets with incomplete information beyond the bipartite setting. The investigation of three-partite all-verses-nothing-like proof such as the one discussed in Ref. [36] under incomplete information would also be very interesting. Finally, asking for an experimental test would require the manipulation of an entire laboratory. This brings us back to Bell's famous question [157]: What exactly qualifies some physical system to play the role of a measurer? Does Charlie's laboratory need to contain a physicist with a PhD? Still, proof-of-principle demonstrations, in the spirit of Ref. [37] would be highly desirable.

## 4 Bipartite Bell inequalities with low detection efficiency

Bell inequality tests, where the detection efficiency is below a certain threshold  $\eta_{\text{crit}}$ , can be simulated by means of local hidden variable models. If this is the case, the quantum advantage in many Bell inequality-based protocols will vanish. Here, we introduce a new method to identify Bell tests requiring a low detection efficiency and a relatively low dimension of the local quantum systems. First, we present a family of bipartite Bell inequalities where  $\eta_{\text{crit}}$  can be upper bounded by a function of graph invariants. Second, using a modified version of the so-called Gilbert's algorithm, we optimize the obtained inequality for smaller detection efficiency and better noise robustness. Finally, we illustrate the power of our method by developing an explicit example, yielding a lower  $\eta_{\text{crit}}$  and requiring lower visibility. This Chapter is based on Project [F].

### 4.1 Motivation and previous works

Bell nonlocality refers to the violation of inequalities which are satisfied by *any* local hidden variable (LHV) model and is a fascinating feature of quantum theory. Furthermore, it is a crucial mean to accomplish tasks that are impossible with classical resources. In an ideal Bell test, in every run the two particles being emitted by the source would be detected, one at one party's site and the other at the other party's side. However, in a real Bell test this might *not* be the case. This imperfect detection is quantified by the detection efficiency  $\eta$  of an experimental test of a Bell inequality, which is defined as the ratio between the number of systems detected by one party and the number of pairs emitted by the source. Consequently,  $\eta$  does not only depend on the efficiency of the detectors, but also on all the losses occurring during the distribution of the state.

It was realized [158] that the experimental correlations in a Bell test can be simulated by LHV models if the detection efficiency is below a certain threshold  $\eta_{\text{crit}}$ . As a consequence, if  $\eta$  is not sufficiently high, the quantum advantage in many Bell inequality-based protocols like randomness expansion [159,160] or secure key distribution [17,19,161,162] will vanish. Avoiding this so-called *detection loophole* [163] requires



surpassing  $\eta_{\text{crit}}$ . Here it is important to notice that  $\eta_{\text{crit}}$  depends on the particular targeted quantum correlations, i.e., the prepared state and the measurements performed, as well as the used Bell inequality. Let us consider the case of symmetric Bell tests and perfect visibility, i.e., the targeted (pure) quantum state is exactly prepared. For the CHSH inequality, using the maximally entangled state and the measurements that yield the largest violation, one finds that  $\eta_{\text{crit}} = 2/(\sqrt{2} + 1) \approx 0.828$ . It was then noticed that  $\eta_{\text{crit}}$  can be lowered by considering non maximally entangled pure two-qubit states and measurement directions that are not initially fixed [85]. By optimizing over the states and measurement directions it was shown that  $\eta_{\text{crit}}$  can be lowered down to  $2/3 \approx 0.667$  for the Clauser-Horn (CH) form [7] of the CHSH inequality [85]. For Bell inequalities with four binary settings and maximally entangled states,  $\eta_{\text{crit}}$  is not better than for the CHSH inequality, with one exception that allows a slightly lower value of  $\eta_{\text{crit}} \approx 0.821$  [164, 165].

Although loophole-free Bell tests [13, 166] have proven that it is possible to produce correlations between local quantum systems of dimension  $d = 2$ , which cannot be explained by a LHV model, the value of  $\eta_{\text{crit}}$  required in these experiments is due to noise given by  $\eta_{\text{crit}} \geq 0.720$ . This value is too high for current quantum technology and thus prevents real-life applications and in particular, applications outside laboratories with well controlled losses, or situations involving longer distances, e.g., longer than 5km. It is in this sense that the requirement of  $\eta_{\text{crit}} > 0.667$  without noise and  $\eta_{\text{crit}} > 0.720$  with noise acts as a bottleneck. As it turns out [167], high-dimensional systems can tolerate a detection efficiency that decreases with the dimension  $d$  of the local quantum systems. However, an improvement over the qubit case only exists for  $d > 1600$ . For ququad systems, that is, for systems of local dimension  $d = 4$ , the detection efficiency can be lowered down to  $\eta_{\text{crit}} = 0.770$  for maximally entangled states and to  $\eta_{\text{crit}} = 0.618$  for non maximally entangled states [168]. For the case of  $N$  copies of the two-qubit maximally entangled state and local Pauli measurements acting on the corresponding qubit subsystems, one can upper bound the critical detection efficiency by  $\eta_{\text{crit}} < 0.809$  for  $N = 2$ ,  $\eta_{\text{crit}} < 0.740$  for  $N = 3$  and  $\eta_{\text{crit}} < 0.693$  for  $N = 4$  [169]. Furthermore, for local dimension  $d = 512$ , one can reduce  $\eta_{\text{crit}}$  to 0.469 [170].

This Chapter is organized as follows. In Section 4.2 we introduce our notation and explain different strategies to obtain  $\eta_{\text{crit}}$  for a particular scenario. We proceed in Section 4.3 by introducing a family of bipartite Bell inequalities where the local bound and maximal quantum value are connected to certain graph invariants. Afterwards, we discuss in Section 4.4 examples of correlations with low detection efficiency. As our approach is based on graphs, we discuss in Section 4.5 how to obtain the independence number and the Lovász number for so-called Newman graphs and proceed in Section 4.6 by incorporating possible experimental noise. In Section 4.7 we explain how our methods can be used to construct Bell inequalities with an arbitrary low detection efficiency. Finally, we introduce in Section 4.8 a method that allows for an optimization

of a Bell inequality towards smaller  $\eta_{\text{crit}}$  and higher resistance to white noise.

## 4.2 Strategies to compute the detection efficiency

In an ideal Bell test, every run would end up with the result that the two particles emitted by the source are detected. However, in a realistic Bell test this may not be the case due to the existence of propagation losses and the imperfection of the detectors. Consequently, a detection at one side may not be accompanied with a detection at the other side and also, in some runs, both particles may be undetected. In some cases, in addition to the local losses and imperfect detectors, the particles are not emitted at well-known times. Then, the number of emitted pairs and thus the number of runs of the Bell test will be unknown. Therefore, one typically divides the methods for calculating  $\eta_{\text{crit}}$  into two classes. First, there is the class of Bell tests where the number of runs is known. This is the case in event-ready experiments [166, 171] and in experiments with heralded detection [172–175]. Second, there is the class of Bell tests where the number of runs is not known. This is the case in existing photonic Bell tests with high detection efficiency. In this context, one typically uses the CH Bell inequality [7], inequalities that can be written in terms of the CH functional [168] and Bell inequalities that can be rewritten similarly [176].

In the following we will assume that the number of runs is known. One approach to incorporate no-detection events into the analysis of the obtained data is to associate the no-detection event with a new outcome of the measurement. Then, one has to find a *new* Bell inequality with the same number of settings but with one more outcome per setting than in the original Bell inequality. The experiment will be detection loophole-free as soon as the new Bell inequality is violated. However, finding this new inequality typically turns out to be difficult. For instance, if one adds a new outcome to all the measurements in the  $(2, m, 2)$  Bell scenario, i.e., two parties,  $m$  settings per party and two outcomes, then one ends up in the  $(2, m, 3)$  Bell scenario. As a consequence, the number of deterministic LHV assignments changes from  $2^{2m}$  to  $3^{2m}$ . Meanwhile, the dimension of the affine space spanned by the LHV assignments changes from  $2m + m^2$  to  $4m + 4m^2$ . To illustrate the hardness of the problem of finding such Bell inequalities, notice that we do not know all Bell inequalities for any  $(2, m, 3)$  Bell scenario.

Another possibility is to associate the no-detection event with one of the existing outcomes of each measurement and thus use the original Bell inequality. Clearly, the experiment will be detection loophole-free as soon as the original Bell inequality is violated. In an ideal Bell test in which we could achieve the maximum quantum value  $\mathcal{Q}$  of a Bell functional  $I$ , whose bound for LHV models is given by  $\mathcal{C}$  and in which the experimental detection efficiency is  $\eta$ , the experimental value  $I_{\text{exp}}$  of  $I$  would be

$$I_{\text{exp}} := \eta^2 \mathcal{Q} + \eta(1 - \eta)(\mathcal{Q}_A + \mathcal{Q}_B) + (1 - \eta)^2 \mathcal{X}, \quad (4.1)$$

where  $Q_A$  denotes the value of  $I$  resulting of what the parties output when Alice has detected the particle but not Bob. The definition of  $Q_B$  is analogous and  $\mathcal{X}$  is the value that the parties output when both Alice and Bob have not detected the particles. Usually, the outputs are chosen such that  $\mathcal{X} = \mathcal{C}$ . The critical detection efficiency is then defined as the smallest value of  $\eta$  such that  $I_{\text{exp}}$  still violates the local bound of the original Bell inequality  $I$ . More precisely,

$$I_{\text{exp}} > \mathcal{C} \Leftrightarrow \eta_{\text{crit}} > \frac{2\mathcal{C} - Q_A - Q_B}{\mathcal{C} + Q - Q_A - Q_B}. \quad (4.2)$$

### 4.3 Constructing Bell inequalities from graphs

In this Section, we will introduce a family of bipartite Bell inequalities, in which each inequality is associated to a graph  $G = (V, E)$ , such that the number of settings of each party coincides with the number of vertices  $|V|$  of  $G$  and the number of outcomes is two. An interesting feature of this family is that the LHV bound of each inequality coincides with the independence number of  $G$ . This allows us to take advantage of the vast literature on independence numbers of families of graphs to construct Bell inequalities whose local bounds would be difficult to compute otherwise.

Recall the definition of the independence number  $\alpha$  and the xi number  $\Xi$  of a graph  $G$  from Section 1.3.1. The independence number of  $G$  is the largest cardinality of any independent set of  $G$  and for the definition of  $\Xi$  we refer to Eq. (1.188).

**Theorem 22.** *Let  $G = (V, E)$  be a graph with independence number  $\alpha$  and xi number  $\Xi$ . Then the linear functional*

$$I = \sum_{j \in V} p(1, 1|j, j) - \frac{1}{2\Xi} \sum_{(k,l) \in E} [p(1, 1|k, l) + p(1, 1|l, k)] \stackrel{\text{LHV}}{\leq} \alpha \quad (4.3)$$

is a Bell inequality.

*Proof.* To obtain the upper bound of  $I$  for LHV models, we only need to consider deterministic probability assignments. From the definition of  $I$  in Eq.(4.3) it is easy to see that the bound cannot be less than  $\alpha$ . Therefore, the bound can only be obtained when the events  $\{(1, 1|j, j)\}_{j \in S}$  have been assigned the value 1, where  $S$  contains no fewer than  $\alpha$  vertices. Let us now assume that  $S$  contains no fewer than  $\alpha + 1$  vertices and let us call  $v$  the vertex in  $S$  such that

$$|\{u \mid u \sim v, u \in S\}| = \Xi(S). \quad (4.4)$$

By changing the assignment of the event  $(1, 1|v, v)$  to be zero, the increment of  $I$  is  $-1 + \Xi(S)/\Xi$ . This is because for all  $i, j \in S$

$$p(1, 1|i, j) = p(1, 1|j, i) = 1 \quad (4.5)$$

for the current assignment, especially for  $i = v$  or  $j = v$ . Therefore, in the case that  $S$  contains no fewer than  $(\alpha + 1)$  vertices, we can always set the assignment of one event  $(1, 1|v, v)$  to be zero, such that the value of  $I$  does not decrease. This implies that the upper bound can be obtained in the case that  $S$  contains exactly  $\alpha$  vertices, which implies that the upper bound can be no more than  $\alpha$ . Consequently, the upper bound for LHV models is exactly  $\alpha$ .  $\square$

There is a second reason why the Bell inequalities which are constructed according to Eq. (4.3) are interesting. They allow us to establish a one-to-one connection between a quantum value for  $I$  and another graph invariant of  $G$ . Moreover, this connection also gives us the initial state and the local observables that provide the quantum value for  $I$ . Recall that an orthonormal representation in  $\mathbb{C}^d$  of a graph  $G = (V, E)$  is an assignment of a nonzero unit vector  $|v_j\rangle \in \mathbb{C}^d$  to each vertex  $j \in V$  satisfying  $\langle v_j | v_k \rangle = 0$  for all pairs  $(j, k) \in E$ . Here it is important to note that the definition of an orthonormal representation does not require that different vertices are assigned different vectors, nor that nonadjacent vertices correspond to nonorthogonal vectors. In certain situations it can be useful to specify an additional unit vector  $|\psi\rangle \in \mathbb{C}^d$ , called handle, together with the orthonormal representation. Further it should be noticed that in many works in graph theory the usual definition of orthonormal representation assigns orthogonal vectors to nonadjacent - instead of adjacent - vertices. The smallest positive integer  $d$  for which there exists an orthonormal representation of  $G$  in  $\mathbb{C}^d$  is called the orthogonal rank of the graph and is denoted by  $\xi$ . As pure quantum states are represented by rays,  $\xi$  is also the minimum dimension that a quantum system must have such that adjacent vertices in  $G$  can be assigned orthogonal quantum states, or equivalently, orthogonal rank-one projectors. However, as already mentioned, it can be the case that the same ray is assigned to different vertices.

**Theorem 23.** *For any graph  $G = (V, E)$ , the maximum quantum value for the inequality  $I$  constructed according to Eq. (4.3) fulfills*

$$\mathcal{Q} \geq \frac{|V|}{\xi}, \quad (4.6)$$

where  $|V|$  is the number of vertices and  $\xi$  is the orthogonal rank of  $G$ . The value  $I = |V|/\xi$  is achieved by preparing the maximally entangled quantum state

$$|\psi\rangle = \frac{1}{\sqrt{\xi}} \sum_{j=0}^{\xi-1} |j\rangle |j\rangle \quad (4.7)$$

and using as local settings on Alice's side the observables represented by the projectors  $|v_j\rangle\langle v_j| \otimes \mathbb{1}$ , with  $\{|v_j\rangle\}$  an orthonormal representation of dimension  $\xi$  of  $G$ . The local observables on Bob's side are represented by the projectors  $\mathbb{1} \otimes |v_j^*\rangle\langle v_j^*|$ , where  $|v_j^*\rangle$  is the complex conjugate of  $|v_j\rangle$ .

*Proof.* Let  $\{|v_j\rangle\}_{j=1}^{|V|} \subset \mathbb{C}^{\xi}$  be an orthonormal representation of the graph  $G$ . Recall that for the maximally entangled state  $|\psi\rangle \in \mathbb{C}^{\xi}$  one has  $\text{Tr}[|\psi\rangle\langle\psi|A \otimes B] = (1/\xi) \text{Tr}[A^\top B]$  for any operator  $A, B \in \mathcal{B}(\mathbb{C}^{\xi})$ . In particular, if  $A$  is hermitian  $A^\top = A^*$  holds and therefore Eq. (4.3) becomes

$$\begin{aligned}
I &= \sum_{j \in V} p(1, 1|j, j) - \frac{1}{2\Xi} \sum_{(k,l) \in E} [p(1, 1|k, l) + p(1, 1|l, k)] \\
&= \sum_{j \in V} \text{Tr} \left[ |\psi\rangle\langle\psi| (|v_j\rangle\langle v_j| \otimes |v_j^*\rangle\langle v_j^*|) \right] - \frac{1}{2\Xi} \sum_{(k,l) \in E} \{ \text{Tr} [|\psi\rangle\langle\psi| (|v_k\rangle\langle v_k| \otimes |v_l^*\rangle\langle v_l^*|)] \\
&\quad + \text{Tr} [|\psi\rangle\langle\psi| (|v_l\rangle\langle v_l| \otimes |v_k^*\rangle\langle v_k^*|)] \} \\
&= \frac{1}{\xi} \sum_{j \in V} |\langle v_j|v_j\rangle|^2 - \frac{1}{2\Xi\xi} \sum_{(k,l) \in E} \{ |\langle v_k|v_l\rangle|^2 + |\langle v_l|v_k\rangle|^2 \} = \frac{|V|}{\xi}.
\end{aligned} \tag{4.8}$$

□

The combination of Theorem 22 and Theorem 23 allows for an upper bound of the critical detection efficiency  $\eta_{\text{crit}}$  given in Eq. (4.2) for the quantum violation of the Bell inequality constructed according to Eq. (4.3) from a graph  $G$  produced with the maximally entangled state in Eq. (4.7). This bound can be calculated via invariants of the graph  $G$  that originates the Bell inequality.

**Theorem 24.** *For any Bell inequality of the form in Eq. (4.3) associated to a graph  $G = (V, E)$ , assuming that the number of runs is known, local models simulating the correlations produced by the state in Eq. (4.7) and the measurements described in Theorem 23 are impossible if the detection efficiency fulfills*

$$\eta > \sqrt{\frac{\alpha\xi}{|V|}} \geq \eta_{\text{crit}}, \tag{4.9}$$

where  $\alpha$ ,  $|V|$  and  $\xi$  are the independence number, the number of vertices and the orthogonal rank of  $G$ , respectively.

*Proof.* To prove the claim it is useful to employ the Collins-Gisin parametrization (see Eq. (1.130)) which allows us to write any Bell inequality as a linear combination of joint and marginal probabilities including one less of the outcomes of each measurement. Then, a strategy in case of no-detection is to associate the no-detection event with outcome zero, which is assumed to be the one that does not appear explicitly in the Bell expression. Following this strategy, the probabilities in the Bell expression transform according to

$$p_A(1|j) \mapsto \eta p_A(1|j), \tag{4.10}$$

$$p_B(1|j) \mapsto \eta p_B(1|j), \tag{4.11}$$

$$p(1, 1|i, j) \mapsto \eta^2 p(1, 1|i, j), \tag{4.12}$$

where  $p_A$  and  $p_B$  denote the marginal probability of Alice and Bob, respectively. If the Bell expression contains no marginal items, as it is the case for the inequality in Eq. (4.3), then we have in Eq. (4.2) that  $Q_A = Q_B = 0$ . If the event of no-detection is associated with the outcome zero, then  $\mathcal{X} = 0$ . Consequently, the quantum value in the ideal case becomes  $\eta^2 Q$ . In this case

$$\eta_{\text{crit}} = \sqrt{\frac{\mathcal{C}}{Q}}, \quad (4.13)$$

where  $\mathcal{C}$  is the upper bound of  $I$  for LHV models. Then, using Theorem 22 and Theorem 23 one arrives at

$$\eta_{\text{crit}} = \sqrt{\frac{\mathcal{C}}{Q}} = \sqrt{\frac{\alpha}{Q}} \leq \sqrt{\frac{\alpha \xi}{|V|}}. \quad (4.14)$$

□

## 4.4 Examples of nonlocal correlations with low detection efficiency

In the following we will make use of Theorem 24 in order to identify quantum correlations and Bell inequalities that allow for a low detection efficiency  $\eta_{\text{crit}}$ .

Recall the definition of  $P_n$ , the set of Pauli observables for a system of  $n \geq 2$  qubits. It consists of the nontrivial quantum observables represented by  $n$ -term tensor products of the Pauli matrices  $\sigma_1, \sigma_2, \sigma_3$  and the identity  $\sigma_0 = \mathbb{1}$ . Clearly, the cardinality of  $P_n$  is  $|P_n| = 4^n - 1$ , as  $P_n$  does not contain the identity operator acting on the space  $\mathbb{C}^{2^n}$ .

**Definition 25.** *The set  $\mathcal{P}_n(\mathbb{C})$  of Pauli states for a system of  $n \geq 2$  qubits consists of the common eigenstates of all the maximal subsets of  $P_n$  containing only mutually compatible observables.*

The Pauli states are also called the quantum states arising from the Pauli group [177]. The eigenvectors of each subset of maximal size of  $P_n$  containing only mutually compatible observables provide a unique orthonormal basis of vectors with  $d = 2^n$ . One can show [178] that there are  $L = \prod_{j=1}^n (2^j + 1)$  such subsets and  $\mathcal{P}_n$  is the union of the  $L$  disjoint orthogonal bases. Accordingly,  $|\mathcal{P}_n(\mathbb{C})| = Ld$ . Further, we write  $\mathcal{P}_n(\mathbb{R})$  for the subset of  $\mathcal{P}_n(\mathbb{C})$  represented by vectors with all components in  $\mathbb{R}$ . One has  $|\mathcal{P}_n(\mathbb{R})| = \prod_{j=1}^n (2^j + 2)$ .

**Definition 26.** *The set  $\mathcal{N}_d$  of Newman states for a quantum system of dimension  $d$ , where  $d = 4k$  for  $k \in \mathbb{N}$ , consists of the states represented by  $d$ -dimensional rays with components  $\pm 1$  such that the number of  $-1$  components is even.*

For example, the set of Newman states for  $k = 1$  is given by

$$\mathcal{N}_4 = \{(1, 1, 1, 1), (1, 1, -1, -1), (1, -1, 1, -1), (1, -1, -1, 1)\}. \quad (4.15)$$

It should be noted that the vectors  $(a_1, a_2, a_3, a_4)$  and  $(-1) \cdot (a_1, a_2, a_3, a_4)$  represent the same Newman state, as they belong to the same ray. From this it directly follows that  $|\mathcal{N}_d| = 2^{d-2}$ . To a given set of vectors  $S \subset \mathbb{C}^d$  one can associate its graph of orthogonality. This is the graph in which each vector is represented by a vertex and two vertices are adjacent if and only if their corresponding vectors are mutually orthogonal.

**Definition 27.** *The Newman graph  $Y_d$  is the graph of orthogonality of  $\mathcal{N}_d$ , where  $d = 4k$  with  $k \in \mathbb{N}$ .*

#### 4.4.1 The graph Pauli-4320

The graph of orthogonality of  $\mathcal{P}_4(\mathbb{R})$  has  $\alpha = 72$  and  $\vartheta = \alpha^* = |V|/\xi = 270$ , where  $\alpha$  is the independence number,  $\vartheta$  the Lovász number and  $\alpha^*$  the fractional packing number of the graph. Therefore, by preparing the maximally entangled state in Eq. (4.7) of local dimension  $\xi = 2^4 = 16$  and allowing the parties to choose between the 4320 projective dichotomic measurements represented by  $|v_j\rangle\langle v_j|$  with  $|v_j\rangle \in \mathcal{P}_4(\mathbb{R})$ , they produce a violation of the Bell inequality in Eq. (4.3), which, using Theorem 23, allows us to conclude that

$$\eta_{\text{crit}}(\mathcal{P}_4(\mathbb{R})) \leq 0.516, \quad (4.16)$$

which is an unprecedentedly low upper bound for this dimension. Here it is important to notice that 4320 local choices are not too many for a realistic Bell test. For example, a photonic loophole-free Bell test may have 3502784250 trials [13], which is enough for a Bell test in which each party has to choose between 4320 settings, as there are still 187.7 trials for each possible combination of settings  $(x, y)$ . This is more than three times the number of trials per setting  $(x, y)$  in the first loophole-free Bell test [12]. Recall that the measurements have two outcomes, as in the test of the CHSH Bell inequality [6]. Therefore, only two detectors per party are necessary.

#### 4.4.2 The graph Pauli-36720

We conjecture that the graph of orthogonality of  $\mathcal{P}_4(\mathbb{C})$  has independence number  $\alpha = 396$ . This conjecture is based on the fact that, after months of computation, 396 is the largest value that we have found. In particular, we have found this number many times, which suggests that our search is sufficiently dense. The computation is based on a greedy-type algorithm taking into account the symmetry of the graph as well as known upper bounds for the independence number by means of spectral graph theory. For a given graph  $G = (V, E)$  we proceed as follows:

- (1) Compute the automorphism group of  $G$  and the corresponding orbits. These orbits yield a partition of the vertex set  $\{1, \dots, |V|\}$ . We denote the  $j$ -th orbit by  $O_j$ .
- (2) From each orbit  $O_j$  we select a vertex  $v^j \in O_j$ . For this vertex there are two possibilities:
  - (a)  $v^j$  is part of a maximal independence set, i.e.,  $v^j \in \mathcal{I}(G)$ . In this case, we remove  $v^j$  and all adjacent vertices of  $v^j$ , as the membership of  $v^j$  in  $\mathcal{I}(G)$  excludes the membership of all adjacent vertices. This produces a tuple  $(G^1, 1)$  containing a new graph  $G^1$  and 1, as we have removed one vertex from the independent set of the original graph.
  - (b)  $v^j$  is not part of a maximal independence set, i.e.,  $v^j \notin \mathcal{I}(G)$ . In this case, we can remove the whole orbit  $O_j$  of  $G$  with  $v^j \in O_j$ . Note that  $v^j$  specifies this orbit uniquely as the orbits  $\{O_j\}_j$  form a partition of the vertex set. This procedure yields a new graph  $G^2$ . As we have *not* removed a member of the maximal independent set, we store the tuple  $(G^2, 0)$ .
- (3) This procedure is iterated for each graph that appears within the decomposition, yielding a sequence of graphs with a strictly decreasing number of vertices. Once the number of vertices of all graphs is lower than a predefined threshold  $\kappa$ , for which  $\alpha$  can be computed directly, the decomposition terminates and yields a set  $\{(G^j, H^j)\}$ , where  $G^j$  is a graph and  $H^j \in \{0, 1\}^n$  is the collection of choices made in each step, given that  $n \in \mathbb{N}$  iterations were necessary in order to achieve the threshold  $\kappa$ . The independence number of the graph  $G$  is then the maximum over

$$\tilde{\alpha}_j = \alpha(G^j) + \sum_k (H^j)_k. \quad (4.17)$$

However, the problem with this approach is that the number of graphs in the decomposition grows exponentially with the decomposition depth, revealing the hardness of the problem of computing  $\alpha$ . If one has a sufficiently high lower bound for  $\alpha(G)$  given a priori, one only needs to collect those graphs appearing in the decomposition process whose independence number is larger than this a priori bound. It is important to note that the decomposition in step (2) can be done for *any* of the orbits  $O_j$ , thus one has to choose one particular orbit in each iteration. For each orbit  $O_j$ , we can construct two new graphs according to the cases (a) and (b) having  $n_a^j$  or  $n_b^j$  vertices. We choose the orbit for which the smallest of the numbers  $n_a^j$  and  $n_b^j$  is largest.

In addition,  $\vartheta = \alpha^* = |V|/\xi = 2295$ . Therefore, if the above conjecture is correct, then, by preparing the maximally entangled state in Eq. (4.7) of local dimension  $\xi = 2^4 = 16$  and allowing the parties to choose between the 36720 two-outcome measurements represented by  $|v_j\rangle\langle v_j|$  with  $|v_j\rangle \in \mathcal{P}_4(\mathbb{C})$ , they can provide a violation of



the Bell inequality in Eq. (4.3) which, using Theorem 23, allows us to conclude

$$\eta_{\text{crit}}(\mathcal{P}_4(\mathbb{C})) \leq 0.415. \quad (4.18)$$

## 4.5 Independence number and quantum value of Newman graphs

While the calculation of the independence number of Pauli graphs  $\mathcal{P}_n(\mathbb{C})$  and  $\mathcal{P}_n(\mathbb{R})$  relies on explicit computation, the independence number of Newman graphs  $Y_n$  can be obtained directly from their structure.

**Definition 28.** A Hadamard matrix of order  $n$  is a real  $n \times n$  square matrix  $H_n$  in which all its entries are either  $+1$  or  $-1$  and whose rows are mutually orthogonal.

It directly follows from the definition that the order  $n$  of a Hadamard matrix must be 1, 2 or a multiple of 4. Therefore, if  $n$  is an even number, each pair of rows in a Hadamard matrix represents two mutually orthogonal  $\pm 1$ -vectors in dimension  $n$ . The same is true for the columns of  $H_n$  considered as  $\pm 1$  vectors. Consequently, taking any pair of rows (or any pair of columns), the number of matching entries must be equal to the number of mismatching entries, exactly  $n/2$ .

**Definition 29.** For  $n \in \mathbb{N}$  the Hadamard graph  $\Omega_n = (V, E)$  is the graph with vertex set  $V = \{-1, 1\}^n$  and edge set  $E = \{(\vec{u}, \vec{v}) \in V \times V \mid \langle \vec{u}, \vec{v} \rangle = 0\}$ . More precisely, each vertex is assigned a  $\pm 1$ -vector of length  $n$  and two vertices are adjacent if and only if the corresponding vectors are orthogonal.

Geometrically, the vectors assigned to the vertices of the Hadamard graph  $\Omega_n$  correspond to the directions of the vertices of an  $n$ -dimensional hypercube centered at the origin. Newman states may be seen as a subset of such hypercube directions. Therefore, a Newman graph  $Y_n$  is an induced subgraph of a Hadamard graph  $\Omega_n$ . The graphs  $\Omega_n$  were introduced in Ref. [179, 180] as a tool to provide an algebraic graph theoretic background for Hadamard matrices. Hadamard graphs play an important role in certain quantum communication protocols [181, 182] and some proofs of the Kochen-Specker theorem [183, 184].

**Definition 30.** The lexicographic product of two graphs  $G_1 = (V_1, E_1)$  and  $G_2 = (V_2, E_2)$  is a graph  $G = G_1[G_2]$  such that its vertex set is the cartesian product  $V(G) = V(G_1[G_2]) = V_1 \times V_2$ , and any two vertices  $(v_i^{(1)}, v_k^{(2)})$  and  $(v_j^{(1)}, v_l^{(2)})$  in  $G$  are adjacent if and only if either  $v_i^{(1)}$  is adjacent with  $v_j^{(1)}$  in  $G_1$  or  $v_i^{(1)} = v_j^{(1)}$  and  $v_k^{(2)}$  is adjacent with  $v_l^{(2)}$  in  $G_2$ .

It should be noted that the lexicographic product is associate but not commutative, a fact which is also emphasized by the notation.

The independence number of the Newman graph  $Y_n$  can be obtained by exploiting a connection between  $Y_n$  and the Hadamard graph  $\Omega_n$  [185]. By definition, the graph  $Y_n$  is a subgraph of  $\Omega_n$  induced by a specific subset of its vertices. Here we will focus on those graphs  $\Omega_n$  for which  $n = 4k$  with  $k \in \mathbb{N}$ . Otherwise,  $\Omega_n$  is empty for  $n$  odd, or bipartite for  $n = 2 \pmod{4}$  [185]. Restricting to the first case, the first important observation is that  $\Omega_n$  is the disjoint union of two isomorphic graphs, that is,

$$\Omega_n = \Omega_n^e \sqcup \Omega_n^o, \quad (4.19)$$

where  $\Omega_n^e$  is the graph defined by the vertices corresponding to vectors with an even number of components 1, and  $\Omega_n^o$  is the graph defined by the vertices corresponding to vectors with an odd number of components 1. Therefore, the independence numbers of the graphs are related via

$$\alpha(\Omega_n) = \alpha(\Omega_n^e) + \alpha(\Omega_n^o) = 2\alpha(\Omega_n^e). \quad (4.20)$$

Further, for the orthogonal ranks  $\zeta$  of the graphs we have

$$\zeta(\Omega_n) = \zeta(\Omega_n^e) = \zeta(\Omega_n^o). \quad (4.21)$$

The second step is to notice that  $\Omega_n^e$  is the lexicographic product of  $Y_n$  with the complement of the complete graph on two vertices, that is,  $\Omega_n^e = Y_n[\overline{K}_2]$ . By using Lemma 37 we obtain

$$\alpha(\Omega_n^e) = \alpha(Y_n)\alpha(\overline{K}_2) = 2\alpha(Y_n) \quad \text{and} \quad \zeta(\Omega_n^e) = \zeta(Y_n). \quad (4.22)$$

One can show that the orthogonal rank of  $\Omega_n$  is given by  $\zeta(\Omega_n) = n$  [186, 187]. Consequently, we have  $\zeta(Y_n) = \zeta(\Omega_n^e) = \zeta(\Omega_n) = n$  and the assignment of  $n$ -dimensional rays with components  $-1$  and  $1$  to the vertices of  $Y_n$ , such that adjacent vertices are assigned orthogonal rays, yields an orthogonal representation of  $Y_n$  of minimum dimension.

On the other hand, the independence number of Hadamard graphs  $\Omega_n$  is not always known. The independence number of  $\Omega_n$  is known for  $n = 4p^k$  for  $k \geq 1$  where  $p$  is an odd prime number [188] and also for the case  $n = 2^k$  for  $k \geq 2$  [189]. In both cases one has

$$\alpha(\Omega_n) = 4 \sum_{j=0}^{\frac{n}{4}-1} \binom{n-1}{j}. \quad (4.23)$$

It still remains a conjecture whether Eq. (4.23) is also valid when  $n$  is another multiple of 4 and the first open case is  $n = 40$  [189]. Combining Eq. (4.20), Eq. (4.22) and Eq. (4.23) we obtain that  $\alpha(Y_n) = (1/4)\alpha(\Omega_n)$  and thus

$$\alpha(Y_{28}) = 397594 \quad \text{and} \quad \alpha(Y_{32}) = 3572224. \quad (4.24)$$

In addition,  $Y_n$  has  $|V| = 2^{n-2}$  vertices and  $|E| = 2^{n-4} \binom{n}{n/2}$  edges.

In the following we will show that for the cases  $n = 28$  and  $n = 32$  the orthogonal rank equals their clique number. In order to prove that the Newman graphs  $Y_{28}$  and  $Y_{32}$  contain cliques of size 28 and 32, respectively, note that such cliques correspond to sets of pairwise orthogonal  $\pm 1$ -rays of cardinality 28 in dimension 28 and cardinality 32 in dimension 32, in which the number of  $-1$  components is an even, or alternatively, an odd number. This fact allows us to rephrase this problem in a slightly different and more convenient way, using Hadamard matrices. Then our goal becomes to construct adequate Hadamard matrices  $H_n$  of orders  $n = 28$  and  $n = 32$ . Each row in  $H_n$  is a  $\pm 1$ -vector in dimension  $n$  and, by definition, the  $n$  rows in  $H_n$  constitute a set of  $n$  pairwise orthogonal  $\pm 1$ -vectors in dimension  $n$ . In fact, these vectors are rays since no two rows can have the same entries with opposite signs due to the orthogonality. If necessary, we can transform  $H_n$  into another equivalent  $n \times n$  Hadamard matrix by negating rows or columns, or by interchanging rows and columns, such that the number of  $-1$  components of the row vectors is an even, or alternatively, an odd number. Notice that, in the end, the resulting set of vectors corresponds to a maximum clique of size  $n$  in the Newman graph  $Y_n$ .

According to Hadamard's conjecture, a Hadamard matrix  $H_n$  of order  $n = 4k$  exists for every positive integer  $k \in \mathbb{N}$ . At the present time, after applying the construction methods due to Sylvester, Paley, Williamson and others, the smallest order for which no Hadamard matrix is known is  $n = 668$ . However, there are many orders  $n > 668$  for which  $H_n$  is known. In particular, this implies that all Newman graphs  $\Omega_n$  with  $n = 4k$  with  $n < 668$  satisfy  $\omega(Y_n) = n$ . There exists a well known recursive procedure to construct Hadamard matrices  $H_n$  of order  $n = 2^k$  for  $k \in \mathbb{N}$ , the so-called Sylvester's construction [190]. Applying this procedure,  $H_{32}$  can be obtained. This matrix fulfills the condition that the number of  $-1$  entries in each row is an even number, hence providing a clique of size 32 for the Newman graph  $Y_{32}$ . More precisely, from  $H_{32}$  we can arrive at the following clique of size 32 given by  $\{\vec{u}_i \otimes \vec{u}_j \otimes \vec{v}_k\}$ , where

$$\begin{aligned} \vec{u}_i, \vec{u}_j &\in \{(1, 1, 1, 1), (1, 1, -1, -1), (1, -1, 1, -1), (1, -1, -1, 1)\}, \\ \vec{v}_k &\in \{(1, 1), (1, -1)\}. \end{aligned} \tag{4.25}$$

A Hadamard matrix  $H_{28}$  is of a more convoluted form. It can be obtained through the so-called Paley's construction [191]. There are 487 inequivalent matrices  $H_{28}$ . To exhibit a specific instance of a clique of size 28 induced in  $Y_{28}$ , we look for a matrix  $H_{28}$  such that the number of  $-1$  entries in each row is again an even number. Such a matrix can be found, see Ref. [192]. The set of row vectors, obtained by replacing therein each 0 entry by  $-1$ , constitutes the desired clique.

Finally, we will now prove that for the Newman graphs  $Y_{28}$  and  $Y_{32}$  the quantum value of  $I$  given in Eq. (4.3) can be  $\alpha^*(Y_n) = \vartheta(Y_n) = |V(Y_n)|/\xi(Y_n)$ , where  $V(Y_n)$  denotes the vertex set of the corresponding graph. First, notice that both  $\Omega_n^e$  and  $\bar{K}_2$

are vertex-transitive. Given that  $\Omega_n^e = Y_n[\overline{K}_2]$  we can conclude that  $Y_n$  is also vertex-transitive, since the lexicographical product of two graphs is vertex-transitive if and only if both graph factors are vertex-transitive [193]. Further, it is known that [186]

$$\vartheta(\Omega_n) = \frac{2^n}{n}. \quad (4.26)$$

From Lemma 37 we know that the Lovász number  $\vartheta$  is multiplicative with respect to the lexicographical product and therefore  $\vartheta(\Omega_n^e) = \vartheta(Y_n)\vartheta(\overline{K}_2) = 2\vartheta(Y_n)$ . Notice that  $\vartheta(\Omega_n^e) = (1/2)\vartheta(\Omega_n)$  since  $\Omega_n = \Omega_n^e \sqcup \Omega_n^o$ . As a consequence we obtain

$$\vartheta(Y_n) = \frac{\vartheta(\Omega_n)}{4} = \frac{2^{n-2}}{n}. \quad (4.27)$$

On the other hand, the fractional packing number of a vertex-transitive graph  $G = (V, E)$  satisfies the relation  $\alpha^*(G) = |V|/\omega(G)$ , where  $\omega(G)$  is the clique number of  $G$ . Given that  $|Y_n| = 2^{n-2}$  and knowing from our discussion before that the clique number for  $Y_{28}$  and  $Y_{32}$  is  $\omega(Y_{28}) = 28$  and  $\omega(Y_{32}) = 32$ , using the vertex-transitivity, we obtain that the quantum values of the functional in Eq. (4.3) can be

$$\alpha^*(Y_{28}) = \vartheta(Y_{28}) \quad \text{and} \quad \alpha^*(Y_{32}) = \vartheta(Y_{32}). \quad (4.28)$$

**Definition 31.** A state-independent contextuality (SI-C) set in dimension  $d \geq 3$  is a set of projectors that produces noncontextual correlations, i.e., correlations that violate some noncontextuality inequality, for any initial quantum state of dimension  $d$ .

State-independent contextuality sets play an important role in our method for identifying correlations with a low  $\eta_{\text{crit}}$ , as any SI-C set can be used to produce quantum correlations which violate the graph-based Bell inequality constructed according to Eq.(4.3). It should be noted that a SI-C set is a sufficient but not necessary condition for a violation of the Bell inequality in Eq. (4.3), i.e., there are sets which are *not* SI-C sets and produce a quantum violation of Eq. (4.3).

**Lemma 32.** Let  $\Omega_n$  be a Hadamard graph,  $\Omega_n^e$  the graph defined by the vertices corresponding to vectors with an even number of positive components and  $\Omega_n^o$  the graph defined by the vectors with an odd number of positive components. For  $n \geq 3$  we have

$$\sum_{\vec{v} \in \Omega_n^e} \vec{v}^\top \vec{v} = \sum_{\vec{v} \in \Omega_n^o} \vec{v}^\top \vec{v} = 2^{n-1} \mathbf{1}_n, \quad (4.29)$$

$$\sum_{\vec{v} \in \Omega_n^e} \vec{v} = \sum_{\vec{v} \in \Omega_n^o} \vec{v} = \vec{0}_n, \quad (4.30)$$

where  $\vec{0}_n = (0, \dots, 0) \in \mathbb{R}^n$ .

*Proof.* We will prove the claim by induction over the number of vertices  $n$ . For  $n = 3$  the vertices of the Hadamard graph  $\Omega_3$  are given by

$$\Omega_3 = \{(1, 1, 1), (1, 1, -1), (1, -1, 1), (-1, 1, 1), (1, -1, -1), \\ (-1, 1, -1), (-1, -1, 1), (-1, -1, -1)\}. \quad (4.31)$$

From this it directly follows that  $\Omega_3^o = \{(1, 1, 1), (1, -1, -1), (-1, 1, -1), (-1, -1, 1)\}$  and  $\Omega_3^e = \{(-1, -1, -1), (1, 1, -1), (1, -1, 1), (-1, 1, 1)\}$  and it can be easily verified that Eq. (4.29) and Eq. (4.30) hold. Let us now assume that the claim holds for  $n = k$ . Notice that, by adding an adequate extra component  $\pm 1$  to the vectors of  $\Omega_k^e$  and  $\Omega_k^o$ , we obtain orthogonal representations  $\Omega_{k+1}^e$  and  $\Omega_{k+1}^o$  such that

$$\Omega_{k+1}^e = \{(\vec{v}, 1) \mid \vec{v} \in \Omega_k^e\} \cup \{(\vec{v}, -1) \mid \vec{v} \in \Omega_k^o\}, \quad (4.32)$$

$$\Omega_{k+1}^o = \{(\vec{v}, 1) \mid \vec{v} \in \Omega_k^o\} \cup \{(\vec{v}, -1) \mid \vec{v} \in \Omega_k^e\}. \quad (4.33)$$

This implies that

$$\sum_{\vec{v} \in \Omega_{k+1}^e} \vec{v} = \sum_{\vec{v} \in \Omega_k^e} (\vec{v}, 1) + \sum_{\vec{v} \in \Omega_k^o} (\vec{v}, -1) = (\vec{0}_k, 2^{k-1}) + (\vec{0}_k, -2^{k-1}) = \vec{0}_{k+1}, \quad (4.34)$$

where the next to last equality follows from the fact that  $\Omega_k^e$  and  $\Omega_k^o$  have the same number of elements, i.e.,  $2^{k-1}$ . In a similar manner one can prove that  $\sum_{\vec{v} \in \Omega_{k+1}^o} \vec{v} = \vec{0}_{k+1}$ . Further we have

$$\sum_{\vec{v} \in \Omega_{k+1}^e} \vec{v}^\top \vec{v} = \sum_{\vec{v} \in \Omega_k^e} \begin{pmatrix} \vec{v}^\top \vec{v} & \vec{v} \\ \vec{v}^\top & 1 \end{pmatrix} + \sum_{\vec{v} \in \Omega_k^o} \begin{pmatrix} \vec{v}^\top \vec{v} & -\vec{v} \\ -\vec{v}^\top & 1 \end{pmatrix} \quad (4.35)$$

$$= \begin{pmatrix} \sum_{\vec{v} \in \Omega_k^e} \vec{v}^\top \vec{v} & \vec{0}_k \\ \vec{0}_k^\top & 2^{k-1} \end{pmatrix} + \begin{pmatrix} \sum_{\vec{v} \in \Omega_k^o} \vec{v}^\top \vec{v} & \vec{0}_k \\ \vec{0}_k^\top & 2^{k-1} \end{pmatrix} \quad (4.36)$$

$$= \begin{pmatrix} 2^{k-1} \mathbb{1}_k & \vec{0}_k \\ \vec{0}_k^\top & 2^{k-1} \end{pmatrix} + \begin{pmatrix} 2^{k-1} \mathbb{1}_k & \vec{0}_k \\ \vec{0}_k^\top & 2^{k-1} \end{pmatrix} = 2^k \mathbb{1}_{k+1}. \quad (4.37)$$

In a similar manner one can prove that  $\sum_{\vec{v} \in \Omega_{k+1}^o} \vec{v}^\top \vec{v} = 2^k \mathbb{1}_{k+1}$ . Consequently, the claim holds for all  $n \geq 3$ .  $\square$

**Theorem 33.** *The set of projectors associated to the set of Newman-2<sup>26</sup> states is a SI-C set in dimension  $d = 28$  and the set of projectors associated to the set of Newman-2<sup>30</sup> states is a SI-C set in dimension  $d = 32$ .*

*Proof.* Denote by  $\mathcal{N}_n$  the set of rays constituting an orthonormal representation for the Newman graph  $Y_n$ . It follows from Lemma 32 that

$$\sum_{\vec{v} \in \mathcal{N}_n} |v\rangle \langle v| = \frac{1}{2n} \sum_{\vec{v} \in \Omega_n^e} \vec{v}^\top \vec{v} = \frac{2^{n-2}}{n} \mathbb{1}_n, \quad (4.38)$$

where  $|v\rangle$  denotes a normalized vector while  $\vec{v}$  denotes the unnormalized version. In the case that  $2^{n-2}/n > \alpha(Y_n)$ , the set  $\mathcal{N}_n$  is a SI-C set. In particular, one finds that  $2^{26}/28 > 2396745 > 397594 = \alpha(Y_{28})$  as well as  $2^{30}/32 = 33554432 > 3572224 = \alpha(Y_{32})$ . Using Eq. (4.23) it is also easy to verify that the claim is also true for  $n = 12, 16, 20, 36, 44, 52, 64, 68, 100, 108, 128, 196, 256, 324, 484, 500, 512$ , as these values are either of the form  $4p^k$  for  $p$  prime and  $k \geq 1$  or of the form  $2^k$  for  $k \geq 2$  and also

satisfy the relation  $\alpha(Y_n) < |V(Y_n)|/\omega(Y_n) = 2^{n-2}/n$ . Note that all of the listed values are smaller than 668, which is the smallest value for which no Hadamard matrix is known.  $\square$

One important property of the set of Newman states  $\mathcal{N}_n$  is that  $\eta_{\text{crit}}$  tends to zero as  $n$  grows. For instance, choosing  $n = 512$  yields

$$\eta_{\text{crit}} \leq \sqrt{\frac{\alpha(Y_{512})\xi(Y_{512})}{|V(Y_{512})|}} = \sqrt{\frac{512 * \alpha(Y_{512})}{2^{510}}} \approx 1.6 \times 10^{-14}. \quad (4.39)$$

### The graphs Newman-2<sup>26</sup> and Newman-2<sup>30</sup>

Using the methods presented in Section 4.5, the graph of orthogonality of  $\mathcal{N}_{28}$  has  $\alpha = 397594$  and  $\vartheta = \alpha^* = |V|/\xi = 16777216/7 \approx 2.3967 \times 10^6$ . Therefore, by preparing the maximally entangled state in Eq. (4.7) of local dimension  $\xi = 28$  and allowing the parties to choose between the  $2^{26}$  two-outcome measurements represented by  $|v_j\rangle\langle v_j|$  with  $|v_j\rangle \in \mathcal{N}_{28}$ , they can produce a violation of the Bell inequality in Eq. (4.3) which, using Theorem 23, allows us to conclude that

$$\eta_{\text{crit}}(\mathcal{N}_{28}) \leq 0.407. \quad (4.40)$$

In a similar manner, for  $\mathcal{N}_{30}$  one finds  $\alpha = 3572224$  and  $\vartheta = \alpha^* = 2^{25}$ . Consequently, by preparing the maximally entangled state in Eq. (4.7) of local dimension  $\xi = 32$  and allowing the parties to choose between the  $2^{30}$  two-outcome measurements represented by  $|v_j\rangle\langle v_j|$  for  $|v_j\rangle \in \mathcal{N}_{30}$  they produce a violation of the Bell inequality in Eq. (4.3) such that

$$\eta_{\text{crit}}(\mathcal{N}_{30}) \leq 0.326. \quad (4.41)$$

## 4.6 Incorporating noise in graph-based Bell inequalities

In the context of the critical detection efficiency the effect of noise is typically modeled by assuming that the effective state is of the form

$$\varrho = \mu|\psi\rangle\langle\psi| + (1-\mu)\frac{\mathbb{1}}{d^2}, \quad (4.42)$$

where  $|\psi\rangle \in \mathbb{C}^d$  is the targeted state,  $\mu \in [0, 1]$  is the visibility of the state and  $d$  is the dimension of the local systems. However, it is important to note that there exist situations where the effective state can not be represented in the form of Eq. (4.42).

**Theorem 34.** *Let  $I$  be a Bell inequality which is constructed according to Eq. (4.3) from a graph  $G = (V, E)$ . For states given by Eq. (4.42) the critical visibility  $\mu_{\text{crit}}$ , i.e., the minimal*

value of  $\mu$  in Eq. (4.42) such that one still obtains a violation of Eq. (4.3) is given by

$$\mu_{\text{crit}} \leq \frac{\alpha - \mathcal{Q}_{\text{mix}}}{\frac{|V|}{d} - \mathcal{Q}_{\text{mix}}}, \quad (4.43)$$

where

$$\mathcal{Q}_{\text{mix}} = \frac{1}{d^2} \left( |V| - \frac{|E|}{\Xi} \right). \quad (4.44)$$

*Proof.* As the Bell inequality in Eq. (4.3) is linear in the quantum state, we can compute its value for  $|\psi\rangle\langle\psi|$  and  $(1/d^2)\mathbb{1}$  separately. If  $\{|v_j\rangle\}_{j=1}^{|V|} \subset \mathbb{C}^d$  is an orthonormal representation of  $G$  in dimension  $d$  and we denote  $\Pi_j := |v_j\rangle\langle v_j|$  and  $\Pi_j^* := |v_j^*\rangle\langle v_j^*|$ , we have

$$I\left(\frac{\mathbb{1}}{d^2}\right) = \frac{1}{d^2} \sum_{j \in V} \text{Tr}[\Pi_j \otimes \Pi_j^*] - \frac{1}{2d^2\Xi} \sum_{(k,l) \in E} (\text{Tr}[\Pi_k \otimes \Pi_l^*] + \text{Tr}[\Pi_l \otimes \Pi_k^*]) \quad (4.45)$$

$$= \frac{1}{d^2} \left( |V| - \frac{|E|}{\Xi} \right) =: \mathcal{Q}_{\text{mix}}. \quad (4.46)$$

Further we know from Theorem 23 that the maximally entangled state  $|\psi\rangle$  of local dimension  $d$  yields  $I(|\psi\rangle\langle\psi|) = |V|/d$ . Therefore, for the quantum state  $\rho$  we find

$$I(\rho) = \mu \left( \frac{|V|}{d} - \mathcal{Q}_{\text{mix}} \right) + \mathcal{Q}_{\text{mix}}. \quad (4.47)$$

A violation of  $I$  means  $I(\rho) > \alpha$  which yields the claim.  $\square$

**Theorem 35.** For a Bell inequality of the form given by Eq. (4.3) originating from a graph  $G = (V, E)$  and quantum states as in Eq. (4.42), the critical detection efficiency is given by

$$\eta_{\text{crit}}^2 \leq \frac{\alpha d^2}{|V|(\mu(d-1) + 1) - |E|(1-\mu)/\Xi}. \quad (4.48)$$

*Proof.* The quantum violation of the Bell inequality in Eq. (4.3) with the noisy state  $\rho$  from Eq. (4.42) and perfect detection efficiency is given by

$$\mathcal{Q}' = \mu \mathcal{Q} + (1-\mu) \mathcal{Q}_{\text{mix}}, \quad (4.49)$$

where  $\mathcal{Q} \geq |V|/d$  is the expected quantum violation and  $\mathcal{Q}_{\text{mix}}$  is the value for the maximally mixed state as computed in Eq. (4.46). Thus we obtain by using Eq. (4.13) that

$$\eta_{\text{crit}}^2 = \frac{\mathcal{C}}{\mathcal{Q}} \leq \frac{\alpha}{\mathcal{Q}'} = \frac{\alpha}{\mu \frac{|V|}{d} + (1-\mu) \frac{1}{d^2} (|V| - |E|/\Xi)}, \quad (4.50)$$

which is equivalent to the expression in Eq. (4.48).  $\square$

Theorem 35 implies that, for the graph-based Bell inequalities in Eq. (4.3),  $\eta_{\text{crit}}$  increases rapidly with the number of edges unless  $\mu$  is very close to 1. Therefore, although experimental values of  $\mu$  can be as high as 0.980 for  $d = 3$  and 0.943 for  $d = 17$  [194], it would be desirable to find Bell inequalities for which the same correlations, i.e., the same state and the same measurements, have a value for  $\eta_{\text{crit}}$  and for  $\mu_{\text{crit}}$  that is much less sensitive to noise. We address this problem in Section 4.8.

## 4.7 Bell inequalities with arbitrary low detection efficiency

In the following we describe a method to construct Bell inequalities that offer a critical detection efficiency which is arbitrarily small. More precisely, if there are no restrictions on the number of local settings or the local dimension of the quantum systems, we can identify quantum correlations and a corresponding Bell inequality with respect to which  $\eta_{\text{crit}}$  is as close to zero as desired. These methods are again based on graph-theoretical constructions.

### General theory

**Definition 36.** *The conormal product of two graphs  $G_1 = (V_1, E_1)$  and  $G_2 = (V_2, E_2)$  is a graph  $G = G_1 * G_2 = (V, E)$  such that its vertex set is the cartesian product  $V = V_1 \times V_2$  and any two vertices  $(v_i^{(1)}, v_k^{(2)})$  and  $(v_j^{(1)}, v_l^{(2)})$  in  $G_1 * G_2$  are adjacent if and only if  $v_i^{(1)}$  is adjacent with  $v_j^{(1)}$  in  $G_1$  or  $v_k^{(2)}$  is adjacent with  $v_l^{(2)}$  in  $G_2$ .*

From the definition of the conormal product, it is evident that it is associative and commutative. Further, recall the definition of a spanning subgraph. Given a graph  $G = (V, E)$ , a spanning subgraph  $H = (\tilde{V}, \tilde{E})$  of  $G$  is a subgraph of  $G$  such that  $\tilde{V} = V$ . A spanning subgraph of a graph  $G$  can therefore be seen as a subgraph that results from  $G$  by edge deletions only. For our construction it is crucial that the lexicographical product of graphs as well as the conormal product behave well with respect to the calculation of the independence number and the Lovász number.

**Lemma 37.** *Let  $G$  and  $H$  be graphs. If  $*$  denotes the conormal product and  $\circ$  the lexicographic product, one has*

$$(1) \quad \alpha(G * H) = \alpha(G)\alpha(H) \text{ and } \vartheta(G * H) = \vartheta(G)\vartheta(H).$$

$$(2) \quad \alpha(G \circ H) = \alpha(G)\alpha(H) \text{ and } \vartheta(G \circ H) = \vartheta(G)\vartheta(H).$$

The proof that  $\alpha$  and  $\vartheta$  are both multiplicative with respect to the conormal product can be found in Ref. [108] and Ref. [195]. The same fact with respect to the lexicographical product is proven in Ref. [196] and Ref. [197].



**Theorem 38.** For any Bell inequality that is of the form in Eq. (4.3) and associated to a graph  $G$ , where either  $G = H^{*n}$  or  $G = H^{\circ n}$  for a graph  $H$ , local models are impossible if the detection efficiency is bounded from below by

$$\eta > \sqrt{\frac{\alpha^n \xi^n}{|V|^n}} \geq \eta_{\text{crit}}(G), \quad (4.51)$$

where  $\alpha$ ,  $|V|$  and  $\xi$  are the independence number, the number of vertices and the orthogonal rank of  $H$ , respectively.

*Proof.* The claim follows directly by combining Theorem 24 and Lemma 37.  $\square$

### The case of Pauli-240<sup>n</sup>

The graph of orthogonality of  $\mathcal{P}_3(\mathbb{R})$  has  $\alpha = 16$  and  $\vartheta = \alpha^* = 30$ . Therefore, by preparing the maximally entangled state in Eq. (4.7) of local dimension  $\xi = 2^3 = 8$  and allowing each of the parties to choose between the 240 two-outcome measurements represented by  $|v_j\rangle\langle v_j|$  with  $|v_j\rangle \in \mathcal{P}_3(\mathbb{R})$ , they can produce a violation of the Bell inequality in Eq. (4.3) which, using Theorem 23, allows us to conclude that

$$\eta_{\text{crit}}(\mathcal{P}_3(\mathbb{R})) \leq \frac{\alpha \xi}{|V|} = \frac{2\sqrt{30}}{15} \approx 0.730. \quad (4.52)$$

Therefore, with a system of local dimension  $8^n$  and locally measuring the observables associated to the vertices of the  $n$ -fold lexicographical product of  $\mathcal{P}_3(\mathbb{R})$  with itself, one finds

$$\eta_{\text{crit}}(\mathcal{P}_3^n(\mathbb{R})) \leq \sqrt{\left(\frac{8}{15}\right)^n}. \quad (4.53)$$

In particular, this implies that for  $n = 2$  one has  $\eta_{\text{crit}}(\mathcal{P}_3^2(\mathbb{R})) \leq 0.533$  and for  $n = 3$  one has  $\eta_{\text{crit}}(\mathcal{P}_3^3(\mathbb{R})) \leq 0.389$ . The interest of this method is that it tends faster to  $\eta_{\text{crit}} = 0$  using smaller system dimensions  $d$  than in any previous method. However, the downside is that, at least applied to the examples provided here, it requires too many measurement settings.

### Examples with low $\eta_{\text{crit}}$ and smaller number of settings

As already pointed out, most of the presented examples require too many settings to be tested in actual experiments. This leads to the question of whether we can achieve a low  $\eta_{\text{crit}}$  by using a moderate number of measurement settings, e.g., less than 100.

So far, our strategy for finding examples with a low  $\eta_{\text{crit}}$  was inspired by the graphs of orthogonality of Pauli and Newman states. Here it is important to note that in both cases the graph of orthogonality is a vertex-transitive graph. Therefore, in our search for a systematic method to identify additional examples with low  $\eta_{\text{crit}}$ , we will first

$ V $	$n$	$\min\sqrt{\alpha\omega/ V }$	$ V $	$n$	$\min\sqrt{\alpha\omega/ V }$
18	380	0.816	19	60	0.795
20	1214	0.775	21	240	0.756
22	816	0.739	23	188	0.780
24	15506	0.707	25	464	0.775
26	4236	0.734	27	1434	0.745
28	25850	0.732	29	1182	0.719
30	46308	0.707	31	2192	0.696
32	677402	0.667	33	6768	0.625
34	132580	0.64	35	11150	0.627
36	1963202	0.615	37	14602	0.604
38	814216	0.593	39	48462	0.632
40	13104170	0.571	41	52488	0.561
42	946226	0.6	43	99880	0.635
44	39134640	0.581	45	399420	0.571
46	34333800	0.562	47	364724	0.597

Table 4.1: The minimal value of  $\sqrt{\alpha\omega/|V|}$ , which is a lower bound for  $\eta_{\text{crit}}$ , for all vertex transitive graphs with  $|V| \leq 47$  vertices. Here,  $n$  denotes the number of vertex-transitive graphs with the corresponding number of vertices. The table is taken from Ref. [F].

focus on vertex-transitive graphs. Luckily, vertex-transitive graphs have been investigated for decades [198, 199] and there exist databases with all vertex-transitive graphs with up to 47 vertices [200], all vertex-transitive graphs of degree 3 up to 1280 vertices [201], and all circulant graphs with degrees at most 20 up to 65 vertices, at most 16 up to 70 vertices, and at most 12 up to 100 vertices [202]. Using these databases we can compute  $\eta_{\text{crit}}$  for all these graphs and their complements and then select those which are interesting for our purpose.

For any graph  $G$ , we have the chain of inequalities

$$\omega(G) \leq \vartheta(\overline{G}) \leq \zeta(G) \leq \chi(G), \quad (4.54)$$

where  $\overline{G}$  denotes the complement of  $G$  and  $\omega(G)$ ,  $\vartheta(G)$ ,  $\zeta(G)$ ,  $\chi(G)$  are, respectively, the clique number, the Lovász number, the orthogonal rank and the chromatic number. The clique number  $\omega$  is a trivial lower bound for  $\zeta$ . The problem is that  $\zeta$  cannot be computed efficiently. However, in all the examples with low  $\eta_{\text{crit}}$  that we have identified,  $\zeta = \omega$ . The idea is to use the databases and compute, for each fixed number of vertices  $|V|$ , the minimum of  $\sqrt{\alpha\omega/|V|}$ . This yields a lower bound for  $\eta_{\text{crit}}$  that can be expected, for maximally entangled states and before any optimization, for the corresponding set of graphs. The results of these computations for all vertex-transitive

graphs up to 47 vertices are summarized in Tab. 4.1. There one can see that the aforementioned lower bound for  $\eta_{\text{crit}}$  decreases as the number of vertices increases. Moreover, it suggests that, for maximally entangled states and before any optimization, there are examples with  $\eta_{\text{crit}} < 0.5$  and  $|V| < 100$  vertices. Further, we can use the existing computational tools [203] to estimate the exact  $\zeta$ . To find an orthogonal representation in  $\mathbb{R}^d$  or  $\mathbb{C}^d$  with minimal  $\zeta$  of the promising graphs, we can write each vector in the orthogonal representation as a unit vector using  $d$ , or  $2d$ , real variables and rotate the orthogonal representation into some canonical position to reduce the number of variables. Then, we take into account that the automorphisms of the graph, which can be easily computed, lead to geometric symmetries in the orthogonal representation. Using numerical optimization software, we run the minimization problem where the objective is to minimize the sum of squares of inner products for  $\mathbb{R}^d$ , or the sum of squares of absolute values of inner products for  $\mathbb{C}^d$ . In both cases the sum is taken over those pairs of vectors that are supposed to be orthogonal in the orthogonal representation. Notice that the automorphisms of the graph dramatically reduce the number of variables in the optimization problem because now we need only one vector for each orbit of a symmetry group.

So far we have restricted the search to vertex-transitive graphs, as they are easy to find and admit a lot of symmetry. However, in Ref. [204] it is shown that there are other types of graphs leading to quantum correlations based on maximally entangled states violating a Bell inequality. These graphs are those which admit an orthonormal representation in dimension  $\zeta$  and nonnegative vertex weights  $w = \{w_j\}_{j=1}^{|V|}$  such that

$$\sum_{j=1}^{|V|} \frac{w_j}{\zeta} > \alpha(G, w), \quad (4.55)$$

where  $\alpha(G, w)$  is the independence number of the corresponding weighted graph. Further, it has been shown in Ref. [205] that a condition for these graphs is that the fractional chromatic number  $\chi_f$  satisfies  $\chi_f > \zeta$ . Here it is important to note that this condition, which is not sufficient for the graphs to have associated SI-C sets is, in fact, sufficient for having quantum correlations based on maximally entangled states violating a Bell inequality. Consequently, another strategy to find examples of graphs that allow for a low  $\eta_{\text{crit}}$  would be the following. Find graphs with  $\chi_f > \zeta$ . For each of these cases, find an assignment of weights  $w = \{w_j\}_{j=1}^{|V|}$  such that  $\sum_j w_j / \zeta > \alpha(G, w)$ . Then, we have

$$\eta_{\text{crit}} \leq \sqrt{\frac{\zeta \alpha(G, w)}{\sum_{j=1}^{|V|} w_j}}. \quad (4.56)$$

Interestingly, since these weights are often natural numbers, e.g., for the Yu-Oh set [206] the weights are 2 for four of the vectors and 3 for the other nine vectors [204], one can

see the weighted graphs  $(G, w)$  as non weighted graphs  $\tilde{G}$  with an extended number of vertices. For instance, for the Yu-Oh set, the extended graph  $\tilde{G}$  would have  $4 \times 2 + 9 \times 3 = 35$  vertices. Then, for finding candidates that may have a low  $\eta_{\text{crit}}$ , we can use databases of non weighted graphs of 13 or more vertices, as it is known that the graphs for which  $\chi_f > \zeta$  must have at least 13 vertices [207]. There, we would first identify graphs with  $\chi_f > \omega$ , which is easier to compute than  $\zeta$ . Since  $\omega \leq \zeta$ , this is a necessary condition. Later on, we can use again the existing computational tools [203] to obtain  $\zeta$ .

## 4.8 Optimizing Bell inequalities using symmetries

From the definition of the graph-based Bell inequalities in Eq. (4.3) it is directly clear that only those probabilities  $p(a, b|i, j)$  are taken into account where either  $i = j$  or  $i$  and  $j$  are adjacent in the graph  $G$ . However, in a Bell test Alice and Bob independently choose their measurements in such a way that the choice of one of them is spacelike separated from the recording of the measurement outcome of the other. Therefore, while every observable  $A_i$  of Alice is compatible with every observable  $B_j$  of Bob, the inequality  $I$  in Eq. (4.3) does not use most of the joint probability distributions  $p(a, b|i, j)$ . Consequently, all the not used distributions are wasted. This motivates the following questions. What if we use the same state and the same measurement directions as for the violation of the graph-based Bell inequality in Eq. (4.3) and consider all the joint probability distributions  $p(a, b|i, j)$ ? Is it possible to obtain better Bell inequalities, where better refers to a higher resistance of noise, i.e., a lower  $\mu_{\text{crit}}$ , a lower critical detection efficiency  $\eta_{\text{crit}}$ , or even both? If one aims to design a loophole-free Bell test, one needs that the experimental values for the visibility  $\mu_{\text{exp}}$  and the detection efficiency  $\eta_{\text{exp}}$  are above their respective critical values. That is, we need  $\mu_{\text{exp}} > \mu_{\text{crit}}$  as well as  $\eta_{\text{exp}} > \eta_{\text{crit}}$ . In the following we compute  $\mu_{\text{crit}}$  and  $\eta_{\text{crit}}$  for quantum states of interest. Intentionally, the first example does not offer a low  $\mu_{\text{crit}}$  nor a low  $\eta_{\text{crit}}$ . However, it will guide us to attack more interesting examples.

### The case of Pauli-24

In Fig. 4.1 we have illustrated the graph of orthogonality of the 24 non-normalized states in  $\mathcal{P}_2(\mathbb{R})$ . This graph has  $\alpha = 5$  and  $\vartheta = \alpha^* = 6$ . Therefore, by preparing the maximally entangled state in Eq. (4.7) of local dimension  $\zeta = 2^2 = 4$  and allowing each of the parties to choose between the 24 two-outcome measurements which are represented by  $|v_j\rangle\langle v_j|$  with  $|v_j\rangle \in \mathcal{P}_2(\mathbb{R})$ , Alice and Bob can produce a violation of the Bell inequality in Eq. (4.3). In particular, by using Theorem 23 we can conclude that

$$\eta_{\text{crit}}(\mathcal{P}_2(\mathbb{R})) \leq 0.913 \quad (4.57)$$

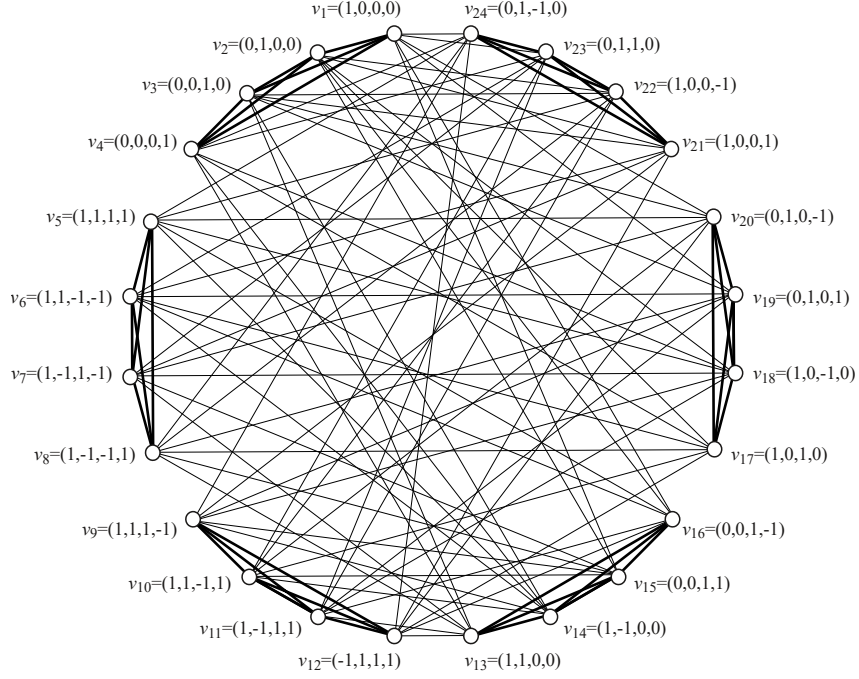


Figure 4.1: Graph of orthogonality of the 24 Pauli states in  $\mathcal{P}_2(\mathbb{R})$ . Vertices (dots) represent states and edges connect those that are orthogonal. The 24 states can be distributed in 6 disjoint orthogonal bases, which are indicated by thicker edges. The figure is taken from Ref. [F].

and by using Theorem 34, we obtain that

$$\mu_{\text{crit}}(\mathcal{P}_2(\mathbb{R})) \leq 0.911. \quad (4.58)$$

#### 4.8.1 Gilbert's algorithm

Although the graph-based Bell inequalities in Eq. (4.3) are neither tight, i.e., they are not facets of the local polytope, nor robust to noise, they can be further improved to offer better detection efficiency and noise robustness. The idea is to take the obtained graph-based Bell inequality as a starting point to obtain a new inequality, which we will parametrise in the Collins-Gisin form, see Eq. (1.130), by employing two different methods. The first method is a linear program which optimizes over the entire local polytope to find the optimal Bell functional [169]. However, this technique requires to enumerate and to store all the local deterministic points which are given by the vertices of the local polytope. Clearly, this becomes an increasingly difficult computational task as the number of measurement settings increases. For instance, the local

polytope corresponding to the Bell inequality with 24 settings per party derived from  $\mathcal{P}_2(\mathbb{R})$  has  $2^{48} \approx 10^{14}$  vertices. This number is too large to be stored on a standard computer. Therefore we have to choose a different approach, which is based on the so called Gilbert's distance algorithm [208] and does not require to store all the vertices of the local polytope. However, it is important to note that for our case of interest, that is,  $\mathcal{P}_2(\mathbb{R})$ , the problem is still intractable if Gilbert's original algorithm is used. The problem only becomes feasible if one incorporates symmetries that are present in the problem.

Gilbert's algorithm is a well known numerical method to detect collisions between convex sets. It has been used for improving the detection efficiency of Bell inequalities [169], deciding whether or not a given correlation is nonlocal [209], and the construction of entanglement witnesses [210, 211]. The algorithm minimizes the distance between local points on facets of the local polytope  $\mathcal{L}$  and a given nonlocal point. The minimization is achieved by iteratively finding a better local point that decreases this distance. The algorithm terminates when the difference of distances between successive iterations falls below a certain threshold value, which is typically chosen to be very small. The resulting Bell functional is then identified as the separating hyperplane between the specified nonlocal point and the local point on the facet found by minimizing the distance.

More precisely, the algorithm is based on having access to an oracle which is capable of maximizing over the local polytope  $\mathcal{L}$  the overlap with a given point, i.e., the inner product between a given point in  $\mathbb{R}^n$ . Initially, one has to specify the local polytope  $\mathcal{L} \subset \mathbb{R}^n$ , presented as the convex hull of its vertices, and a point  $\vec{q} \in \mathbb{R}^n$  associated to the given quantum correlation. Then the algorithm proceeds as follows.

- (1) Choose a point  $\vec{s}_0 \in \mathcal{L}$ .
- (2) Given the input point  $\vec{s}_k$  one uses the oracle to compute

$$\vec{r}_k := \operatorname{argmax}_{\vec{p} \in \mathcal{L}} \langle \vec{q} - \vec{s}_k, \vec{p} - \vec{s}_k \rangle = \operatorname{argmax}_{\vec{p} \in \mathcal{L}} \langle \vec{q} - \vec{s}_k, \vec{p} \rangle. \quad (4.59)$$

- (3) Given  $\vec{s}_k$  and  $\vec{r}_k$ , calculate the convex combination of both which minimizes the distance to the quantum point  $\vec{q}$ , that is,

$$\lambda_k := \operatorname{argmin}_{\lambda \in [0,1]} \|(1 - \lambda)\vec{s}_k + \lambda\vec{r}_k - \vec{q}\|. \quad (4.60)$$

- (4) Define the new starting point  $\vec{s}_{k+1} := (1 - \lambda_k)\vec{s}_k + \lambda_k\vec{r}_k$ .

Since the objective function in Eq. (4.59) is linear and the local polytope is convex, the maximizer will be an extreme point of  $\mathcal{L}$ . More precisely, the maximizer  $\vec{r}_k$  is given by a vertex of  $\mathcal{L}$ . The optimal value for  $\lambda$  in the  $k$ -th iteration of Gilbert's algorithm,

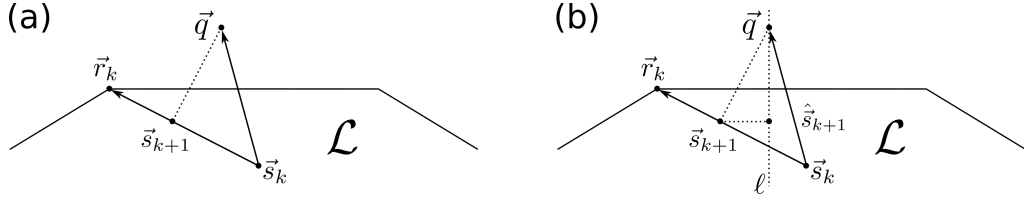


Figure 4.2: (a) Illustration of the standard Gilbert's algorithm. The quantum point  $\vec{q}$  lies outside of the local polytope  $\mathcal{L}$ . Starting with an arbitrary local point  $\vec{s}_k \in \mathcal{L}$ , the oracle yields the point  $\vec{r}_k$  within the local polytope  $\mathcal{L}$ , maximizing the overlap with  $\vec{q} - \vec{s}_k$ . From there, a new starting point  $\vec{s}_{k+1}$  can be calculated. The distance of the new point  $\vec{s}_{k+1}$  to the quantum point  $\vec{q}$  is smaller or equal than the distance between  $\vec{s}_k$  and  $\vec{q}$ . (b) Illustration of Gilbert's algorithm with symmetrization. In this simple example, the quantum point  $\vec{q}$  and the local polytope  $\mathcal{L}$  are invariant under the flip around the line  $\ell$ . After the point  $\vec{s}_{k+1}$  has been computed by means of the standard Gilbert's algorithm, we obtain its symmetrization  $\hat{\vec{s}}_{k+1}$  for the flip around  $\ell$ . The point  $\hat{\vec{s}}_{k+1}$  is used instead of  $\vec{s}_{k+1}$  as the new starting point for the next iteration. The figure is taken from Ref. [F].

denoted by  $\lambda_k$  can be computed directly and is given by

$$\lambda_k = \min \left\{ \frac{\langle \vec{q} - \vec{s}_k, \vec{r}_k - \vec{s}_k \rangle}{\|\vec{r}_k - \vec{s}_k\|}, 1 \right\}. \quad (4.61)$$

In the standard Gilbert's algorithm the oracle is implemented by enumerating all the vertices of the local polytope  $\mathcal{L}$  to compute the inner product in the last equality in Eq. (4.59). For a geometrical interpretation of the iteration, see also Fig. 4.2. Therefore, this algorithm provides a sequence of Bell functionals, which become better with each iteration. However, one does not necessarily obtain a tight Bell inequality as in method based on a linear program. Moreover, calculating the local bound of the resulting Bell functional still remains a hard problem, which again requires enumerating and storing all the local deterministic points of at least one party. This issue is also shared by the oracle in the standard Gilbert's algorithm, since the evaluation of Eq. (4.59) is equivalent to find the local bound of a Bell functional.

#### 4.8.2 Gilbert's algorithm with symmetry

As already mentioned in Section 1.2.3, the vertices of the local polytope are invariant under the following invertible transformations:

- (1) Swapping the outcomes of a measurement setting for either Alice or Bob.
- (2) Simultaneously permuting the measurement settings of Alice and Bob.

(3) Swapping the measurement settings of Alice and Bob.

Here invariant means that the transformations map local correlations to local correlations. The joint probability distributions that can be obtained by performing measurements on a quantum state are also invariant under some of these transformations. We will denote by  $\mathcal{S}$  the subset of transformations which keep the quantum joint probability distributions and the local polytope invariant simultaneously. If the correlations are presented in the Collins-Gisin parametrization, see Eq. (1.130), also the Bell functional can be rephrased in this form

$$I = \left( \begin{array}{c|ccc} & c(a=1|x=1) & \cdots & c(a=1|x=m) \\ \hline c(b=1|y=1) & c(1,1|1,1) & \cdots & c(1,1|m,1) \\ \vdots & \vdots & \ddots & \vdots \\ c(b=1|y=m) & c(1,1|1,m) & \cdots & c(1,1|m,m) \end{array} \right). \quad (4.62)$$

The corresponding Bell inequality can then be calculated as  $\text{Tr}[IP^\top] \leq \lambda$ , where  $P$  denotes the given correlations and  $\lambda$  the local bound of  $I$ . Under the transformations  $S \in \mathcal{S}$  we have

$$\text{Tr}[UP^\top] = \text{Tr}[IS(P)^\top] = \text{Tr}[S^{-1}(I)P^\top] \leq \lambda, \quad (4.63)$$

where  $S(P)$  is the resultant matrix after the transformation  $S$ . Consequently we have

$$\text{Tr}[IP^\top] = \text{Tr}[I\hat{P}^\top] = \text{Tr}[\hat{I}P^\top] = \text{Tr}[\hat{I}\hat{P}^\top], \quad (4.64)$$

where

$$\hat{P} := \frac{1}{|\mathcal{S}|} \sum_{j=1}^{|\mathcal{S}|} S_j(P), \quad \hat{I} := \frac{1}{|\mathcal{S}|} \sum_{j=1}^{|\mathcal{S}|} S_j^{-1}(P), \quad (4.65)$$

with  $|\mathcal{S}|$  to be the cardinality of  $\mathcal{S}$ . Therefore the expressions  $\hat{P}$  and  $\hat{I}$  can be seen as averages of  $P$  and  $I$  with respect to the symmetry group  $\mathcal{S}$ . Using this symmetry  $\mathcal{S}$ , it is sufficient to only consider inequalities that share this symmetry. In particular, given a symmetric inequality, the vertices of the local polytope  $\mathcal{L}$  can be partitioned into different equivalence classes with respect to that symmetry and the local bound can be computed by choosing only *one* representative out of each class. This drastically reduces the total number of required local vertices, allowing for an enumeration of all symmetrized local points and an evaluation of the local bound.

For convenience, here we focus on the symmetrization applied only to Alice's measurement settings. Each equivalence class consists of vertices which can be transformed into each other by using the aforementioned symmetry transformations, while the same is not true for vertices in different classes, that is, the partition generates disjoint sets. This leads to a modified oracle, which can be much more efficiently evaluated than the original one, as the number of equivalence classes could be much smaller than the number of all vertices. As we will see in the following, this is indeed the case for  $\mathcal{P}_2(\mathbb{R})$ .



### 4.8.3 Generating the set of symmetrized vertices

It remains to describe how to obtain the reduced set of vertices on which the optimization in Eq. (4.59) has to run. In the first step, we have to determine the allowed symmetry transformations  $\mathcal{S}$ , which should be shared by the chosen quantum point  $\vec{q}$  and the local polytope  $\mathcal{L}$ . The local polytope  $\mathcal{L}$  is invariant under the permutation of parties, the permutation of measurements for each party and the permutation of the outcomes for each measurement. By construction, the chosen quantum point  $\vec{q}$  is also invariant under the permutation of parties. However, the point  $\vec{q}$  usually changes if the outcomes of the measurements are permuted. In general, determining the permutation symmetries of measurements in the point  $\vec{q}$  can be a difficult task if the number of measurements is large. In the particular case considered here, those symmetry transformations correspond to the ones in the automorphism group of the graph associated to the SI-C sets. Therefore, the symmetry transformations used in the Gilbert's algorithm with symmetrization are the ones in the automorphism group and the permutation of parties. It remains to explain how the vertices are partitioned into different equivalence classes. The first important observation is that we do *not* need to generate, store and classify all vertices, as assignments with a different number of 1's cannot be equivalent to each other, i.e., the number of 1's that appear in the assignment is an invariant with respect to the symmetries considered here. Hence we start with the assignment that only contains zeros. Obviously, this is invariant under all possible permutations. From this, we generate all possible assignments that can be obtained by replacing one 0 by one 1. Within this set, we check whether some of these assignments are equivalent under the given symmetry transformations  $\mathcal{S}$ , which are presented as permutations. Then, we only keep one representative for each class. This procedure is repeated until no 0 is left in the assignment vector.

By selecting only a single vertex from each equivalence class and all the vertices of Bob, we find that the total number of deterministic assignments for Alice is 21564 for the Bell inequality corresponding to  $\mathcal{P}_2(\mathbb{R})$ , while without symmetrization, the number would be  $2^{24}$ . In addition, we also modify Gilbert's algorithm to evaluate the Bell functional according to the symmetrization procedure in Eq. (4.65). Specifically, we symmetrize the local point chosen in each iteration of the program after minimizing its distance from the target nonlocal point, see Fig. 4.2 for a simple illustration. This results in better convergence times of the algorithm since symmetrization does not increase the distance.

### 4.8.4 Application to $\mathcal{P}_2(\mathbb{R})$

Applying Gilbert's algorithm together with the symmetrization technique to the correlations produced by the measurements corresponding to  $\mathcal{P}_2(\mathbb{R})$  and the maximally

entangled state, we obtain the following Bell inequality:

$$I_{\mathcal{P}_2(\mathbb{R})} \leq 0 \quad \text{with} \quad I_{\mathcal{P}_2(\mathbb{R})} = \left( \begin{array}{c|cc} & \vec{v} & \vec{v} \\ \hline \vec{v}^\top & 0_{12} & M_2 \\ \vec{v}^\top & M_2 & 0_{12} \end{array} \right), \quad (4.66)$$

where  $0_{12} \in \mathbb{R}^{12 \times 12}$  denotes the zero matrix,  $\vec{v} \in \mathbb{R}^{12}$  a vector which consists only of  $-6$ 's and

$$M_2 = \begin{pmatrix} 5 & 5 & \bar{4} & \bar{4} & 5 & 5 & \bar{4} & \bar{4} & 5 & 5 & \bar{4} & \bar{4} \\ 5 & 5 & \bar{4} & \bar{4} & \bar{4} & \bar{4} & 5 & 5 & \bar{4} & \bar{4} & 5 & 5 \\ \bar{4} & \bar{4} & 5 & 5 & 5 & 5 & \bar{4} & \bar{4} & \bar{4} & \bar{4} & 5 & 5 \\ \bar{4} & \bar{4} & 5 & 5 & \bar{4} & \bar{4} & 5 & 5 & 5 & 5 & \bar{4} & \bar{4} \\ 5 & \bar{4} & 5 & \bar{4} & 5 & \bar{4} & 5 & \bar{4} & 5 & \bar{4} & 5 & \bar{4} \\ 5 & \bar{4} & 5 & \bar{4} & \bar{4} & 5 & \bar{4} & 5 & \bar{4} & 5 & \bar{4} & 5 \\ \bar{4} & 5 & \bar{4} & 5 & 5 & \bar{4} & 5 & \bar{4} & \bar{4} & 5 & \bar{4} & 5 \\ \bar{4} & 5 & \bar{4} & 5 & \bar{4} & 5 & \bar{4} & 5 & 5 & \bar{4} & 5 & \bar{4} \\ 5 & \bar{4} & \bar{4} & 5 & 5 & \bar{4} & \bar{4} & 5 & \bar{4} & 5 & 5 & \bar{4} \\ 5 & \bar{4} & \bar{4} & 5 & \bar{4} & 5 & 5 & \bar{4} & 5 & \bar{4} & \bar{4} & 5 \\ \bar{4} & 5 & 5 & \bar{4} & 5 & \bar{4} & \bar{4} & 5 & 5 & \bar{4} & \bar{4} & 5 \\ \bar{4} & 5 & 5 & \bar{4} & \bar{4} & 5 & 5 & \bar{4} & \bar{4} & 5 & 5 & \bar{4} \end{pmatrix}, \quad (4.67)$$

with  $\bar{4} = -4$ . Using the maximally entangled state in Eq. (4.7), one finds that the maximal quantum value is given by  $I_{\mathcal{P}_2(\mathbb{R})} = 18$ . In addition, for the correlations produced by  $\mathcal{P}_2(\mathbb{R})$  and the maximally entangled state, we have  $\mu_{\text{crit}} = 7/9 = 0.778$ , which is 14.62% lower than the upper bound in Eq. (4.58) and  $\eta_{\text{crit}} = 4/5 = 0.8$  which is 12.38% lower than the upper bound in Eq. (4.57). We are able to prove that  $\mu_{\text{crit}}$  and  $\eta_{\text{crit}}$  are the smallest possible values for the correlations produced by  $\mathcal{P}_2(\mathbb{R})$  and the maximally entangled state. First, we collect all the 452929 pairs of deterministic assignments for the two parties which achieve the maximal bound for LHV models. Each of the assignments  $P$  can be written in the Collins-Gisin form as in Eq. (1.130). After symmetrization, there are only 132 different matrices  $\hat{P}$ , whose convex combination can lead to the corresponding quantum probability matrix, either with  $\mu = 7/9$  or with  $\eta = 4/5$ , as one can verify by a linear program. It is important to note that the inequality in Eq. (4.66) is not tight. There exists a tight Bell inequality providing the same  $\eta_{\text{crit}}$  and  $\mu_{\text{crit}}$  as the ones for the inequality in Eq. (4.66), but it does not provide the two zero blocks on the off-diagonal as we have in Eq. (4.66). If one wants to keep this block structure, the inequality in Eq. (4.66) is the only solution.

Notice that the symmetries of the initial graph are crucial for finding  $M_2$  in Eq. (4.67). For example, there are only 6 different parameters in the symmetric inequality for the case of  $\mathcal{P}_2(\mathbb{R})$ . In contrast, there are 624 parameters in the non-symmetric inequality for  $\mathcal{P}_2(\mathbb{R})$ . In the general case, there are  $2m + 2$  parameters in the symmetric inequality

for  $\mathcal{P}_m(\mathbb{R})$  and  $\mathcal{P}_m(\mathbb{C})$ . This makes it also possible to find a better inequality without resorting to Gilbert's algorithm. More precisely, we can choose  $t$  different values for each parameter, resulting in  $t^{2m+2}$  different inequalities. For each of the inequalities we can verify whether it separates the target quantum point  $\vec{q}$  and the local polytope  $\mathcal{L}$  or not. This can be done by only considering the deterministic assignments up to symmetry. As discussed before, for the case  $m = 2$ , there are 21564 different deterministic assignments for Alice up to symmetry. Consequently, for a fixed inequality, that is, for a particular choice of the parameters  $t$ , this verification can be done efficiently. In a similar manner we can calculate  $\eta_{\text{crit}}$  and  $\mu_{\text{crit}}$  for each inequality. As one can see in Eq. (4.66), we can set some parameters to 0 for the optimal  $\eta_{\text{crit}}$  and  $\mu_{\text{crit}}$ . This observation can be used to speed up the numerical calculations further.

## 4.9 Conclusion and discussion

The methods introduced in this Chapter allow us to achieve  $\eta_{\text{crit}} < 0.52$  with local dimensions 16 and also show how to obtain Bell inequalities with even lower  $\eta_{\text{crit}}$  and higher resistance to noise. Here it would be interesting to know what the values for the visibility  $\mu$  are that can be achieved experimentally in the required configurations. More precisely, we want to know what pairs  $(\eta_{\text{exp}}, \mu_{\text{exp}})$  can be obtained for the type of states and measurements proposed here, e.g., for  $d = 16$  and Pauli-4320. In addition, we have introduced methods to identify further examples with low  $\eta_{\text{crit}}$  requiring smaller number of settings. As shown in Fig. 4.3, the first step of our method already yields upper bounds for  $\eta_{\text{crit}}$  which are substantially smaller than the lowest previously known values for any dimension  $d \geq 16$ . These values indicate that there are quantum correlations which have sufficiently low  $\eta_{\text{crit}}$  for loophole-free Bell tests with higher-dimensional quantum systems and, eventually, over longer distances.

Our results imply that loophole-free Bell nonlocality can be achieved in carefully designed tests involving pairs of systems of these dimensions, going beyond previous loophole-free Bell tests which are all based on qubits. More interestingly, our results also imply that loophole-free Bell nonlocality can be achieved through longer distances than those in previous loophole-free Bell tests. Indeed, when photons propagate through fibers, they experience propagation losses proportional to the propagation distance which also depend on the optical wavelength. This means that in a Bell test over a long distance, the detection efficiency decreases with the distance. For example, with telecom wavelengths, in 10 km of fiber, we may have losses of 0.2 dB/km, which implies multiplying the detection efficiency that we had before adding the 10km of fiber by a factor of 0.64. Therefore, if we have examples of loophole-free Bell nonlocality requiring  $\eta_{\text{crit}} < 0.5$ , we can achieve loophole-free nonlocality over 10 km if we have  $\eta_{\text{exp}} > 0.785$  before adding the 10 km of fiber, as  $0.785 \times 0.64 = 0.502 > 0.5$ . Such val-

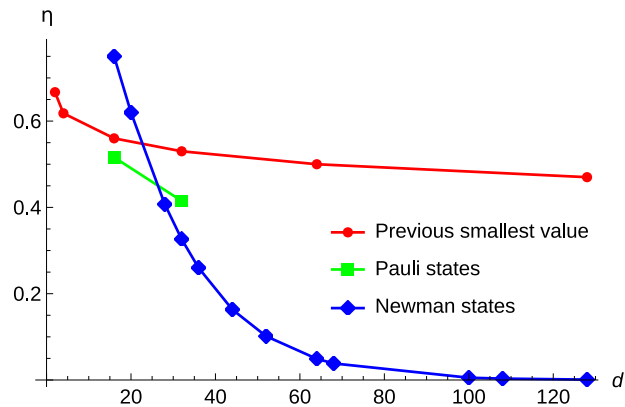


Figure 4.3: The critical detection efficiency  $\eta_{\text{crit}}$  as a function of the dimension  $d$  of the local system. The previous smallest values are those in Ref. [85, 168, 170] and use non maximally entangled pure states. Pauli (Newman) states refer to the case in which the local measurements are projectors on the Pauli (Newman) states and the initial state is maximally entangled, as described in Section 4.3. The figure is taken from Ref. [F].

ues of  $\eta_{\text{exp}}$  have been achieved in previous photonic loophole-free experiments [13], even including fibers and couplings.



# 5 Certifying activation of quantum correlations with finite data

There exist entangled quantum states that do not violate any Bell inequality in a standard Bell test, but their nonlocality can be activated if one allows for an extended experimental setup. However, rigorous statements on the statistical significance of the experimental demonstrations of this activation are not yet available. Behind this there are two difficulties. First, the lack of a method to derive a suitable confidence region from the measured data and second, the lack of an efficient technique to decide locality for every state in the confidence region. In this Chapter we show how both of these problems can be addressed. We introduce a confidence polytope in the form of a so-called hyperoctahedron and provide a computationally efficient method to verify whether a quantum state admits a local hidden state model, thus being unsteerable and, consequently, Bell local. We illustrate how our methods can be used to analyze the activation of quantum correlations by local filtering, specifically, for Bell-nonlocality and quantum steerability. This Chapter is based on Project [C].

## 5.1 Motivation

In a standard Bell test, each of the two parties obtains a particle distributed from a source and can perform measurements on their respective system. After many runs, the correlations of the experiment can be derived from the data. Assuming quantum theory, these correlations can be stratified into classes like quantum entanglement, quantum steering and quantum nonlocality, with each referring to a different level of trust in the local measurement devices [98, 155]. While entanglement can be easier to produce and maintain, it is not enough to relax security assumptions on devices in quantum key distribution, for which Bell-nonlocality or quantum steering is necessary [17, 162, 212]. The latter consist of states entangled strongly enough, to allow for the violation of a Bell inequality or steering inequality.

Interestingly, for pure bipartite quantum states, entanglement turns out to be sufficient in order to reveal Bell-nonlocality in a standard Bell test [213]. Indeed, the simplest Bell inequality, the CHSH inequality [6], can be violated by any such state. Consequently, the notion of entanglement and Bell-nonlocality coincides for this particular

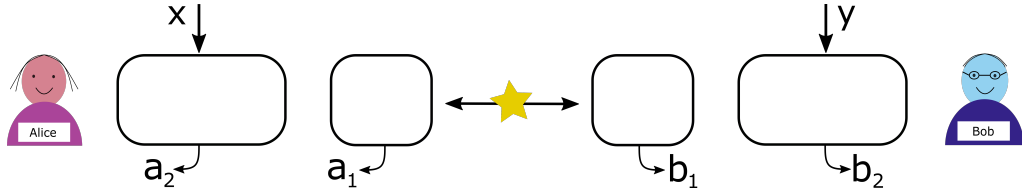


Figure 5.1: Illustration of the activation scenario using local filters. Similar to a standard Bell test there are two parties, Alice and Bob, as well as source (star) distributing a pair of particles. After the parties have received their corresponding system, each of them performs the filtering, yielding outcome  $a_1, b_1 \in \{0, 1\}$  for Alice and Bob, respectively. If both parties observe the correct predefined outcome, they will perform the measurements which can yield a violation of a Bell inequality. This violation will be observed upon observing the correct outcomes  $a_1, b_1$ .

class of states. However, this changes if one also allows for mixed bipartite quantum states. Indeed, Werner showed explicitly [58] that certain highly symmetric entangled states admit a local hidden variable (LHV) model when the measurements at the parties are limited to projective ones. The construction of LHV models for entangled quantum states with all measurements allowed was tackled by Barrett [214]. Consequently, the notion of separability and Bell-local states differ.

This separation of these two notions has an interesting consequence. It turns out that the Bell nonlocality of a local but entangled quantum state can be revealed if one allows for more complex Bell tests [215, 216]. Different extensions have been proposed, for instance, the multi-copy scenario, where  $\rho$  admits an LHV model but  $\rho^{\otimes n}$  can violate a Bell inequality [217, 218]. Another approach to extend the standard Bell test is a scenario where the observers perform a sequence of measurements rather than a single measurement. Such an extension could involve a filter operation for each party before performing a standard Bell test, see also Fig. 5.1. In this context it is important to notice that the enlarged scenario with local filters does not open a loophole for the subsequent Bell test, as the choice of the local settings  $x$  and  $y$  can be made after applying the filters [216, 219].

There already exist experimental demonstrations of the stratification of quantum correlations in these different classes [220] and the interconversion between them by means of local filtering [221–223]. However, a rigorous statistical analysis of the activation of correlations faces difficulties. Indeed, while it is relatively easy to witness the presence of entanglement, quantum steering or Bell nonlocality of the processed state by means of measuring an appropriate inequality, it is in general difficult to demonstrate that the initial state of the experimentally prepared system is contained within a desired set. This is mostly due to two reasons.

Firstly, due to sampling noise and imperfections in the experimental setup, the state can only be determined within a certain confidence region. Current methods for state tomography [224–226] give confidence regions which are either an ellipsoid or a polytope with millions of vertices. Certifying that all states in the confidence region belong to a certain class of correlations, such as Bell-locality, requires the inspection of every extremal point of the confidence region, and thus cannot be carried out in practice.

Secondly, even for a confidence polytope with a small number of vertices, one still faces the problem that for each vertex one has to decide whether it belongs to the targeted class of correlations. The complexity of this task varies with the class of correlations considered. For separable states in low dimensions one can often use criteria like partial transposition [70, 227], which are computationally efficient. For higher-dimensional entanglement, steering, and Bell-local states, this problem is computationally much harder and often cannot be solved in reasonable time even for a relatively low number of states. Recently, this problem has received progress for separability [228] and locality [229–234]. However, for the use case considered here, the algorithms for Bell-local states are still too slow and yield insufficient accuracy.

In this Chapter we will present solutions to both problems and apply our methods to the activation of Bell nonlocality and quantum steerability. More precisely, in Section 5.2 we address the first problem by introducing a simple confidence polytope for quantum state tomography, where the number of vertices scales linearly in the dimension of the state space hence quadratically with the dimension of the underlying Hilbert space. We proceed in Section 5.3 by presenting a solution to the second problem by extending the technique of polytope approximation [231] to give a fast and accurate numerical method to solve the case of qudit-qubit systems. Afterward, we combine in Section 5.4 both techniques and apply them to the activation protocol using local filtering. More precisely, we first present in Section 5.4.1 a family of quantum states that are well suited for the protocol and analyze their nonlocal properties. Then, we discuss the activation of Bell-nonlocality in Section 5.4.2 as well as the activation of quantum steerability in Section 5.4.3. In either case, we provide a rigorous statistical analysis resulting in explicit lower bounds on the number of experimental repetitions.

## 5.2 Construction of a simple confidence polytope

Obviously, for a conclusive demonstration of an activation of Bell-nonlocality or quantum steering one must first demonstrate that the initially prepared quantum state is indeed Bell-local or unsteerable. In particular, this must be achieved with a high level of statistical significance. However, due to noise in the experimental setting, the actual prepared state may deviate from the targeted ideal state, for which locality can be proven. This forces one to learn which high accuracy the effective prepared state,



which can be done by means of quantum state tomography. Still, the tomographic reconstruction has to rely on finite data and if the statistical uncertainty within that data dominates the experimentally introduced errors, one needs in addition a statistical confidence region in state space.

In order to obtain a full reconstruction of the density operator, one has to measure the quantum system with a complete set of measurements. The corresponding outcome probabilities would then allow for an identification of the underlying quantum state, which has generated that data. In the following we write  $E_{a|x}$  for the effect corresponding to the outcome  $a$  of the measurement setting  $x$ . According to the Born rule, if the quantum system is prepared in the state  $\rho$ , the probability of obtaining outcome  $a$  given that the measurement setting was  $x$  is  $p_{a|x}(\rho) := \text{Tr}[E_{a|x}\rho]$ . In an experiment, by implementing the measurement  $x$ , one samples from the distribution  $(p_{a|x}(\rho))_a$ . This yields after  $N$  trials the relative frequencies  $(f_{a|x})_a$  for each setting  $x$ . These frequencies  $f_{a|x}$  can be collected in vector, which we denote by  $\vec{f}$ . In a similar manner, if the measurements  $E = \{E_{a|x}\}$  are fixed and can yield in total  $n$  outcomes, the corresponding probabilities can be organized in a vector as

$$\Phi_E : \text{Mat}_d(\mathbb{C}) \rightarrow \mathbb{R}^n \quad \text{with} \quad \rho \mapsto \Phi_E(\rho) = \left\{ \text{Tr}[E_{a|x}\rho] \right\}_{a,x}. \quad (5.1)$$

Note that one assumes here that the measurement operators  $E_{a|x}$  are known and implemented perfectly. The free least-squares estimator for the underlying quantum state  $\rho$  is defined as the solution of the least-squares problem of minimizing the distance between the given frequencies  $\vec{f}$  and the probabilities induced by the measurement operators  $E_{a|x}$ , that is,

$$\hat{\rho} := \underset{X}{\text{argmin}} \left\{ \sum_{a,x} \left( f_{a|x} - \text{Tr}[XE_{a|x}] \right)^2 \mid X^\dagger = X, \text{Tr}[X] = 1 \right\}. \quad (5.2)$$

Notice that the objective function in Eq. (5.2) can also be written as  $\|\vec{f} - \Phi_E(\rho)\|_2^2$ . The optimization in Eq. (5.2) also involves matrices that are not positive semidefinite, in which case also the closest probability vector  $\Phi_E(X)$  may have negative components. The least-squares estimator has the advantage that it admits a closed form in terms of the map  $\Phi_E$ , that is,

$$\hat{\rho} = (\Phi_E^\dagger \circ \Phi_E)^{-1} \circ \Phi_E^\dagger(\vec{f}). \quad (5.3)$$

In general, if one assumes that the number of experimental runs is sufficiently high, one can make a Gaussian approximation, i.e., one assumes that the empirical frequencies  $\vec{f}$  of the measurement outcomes follow a Gaussian distribution with the mean  $\Phi_E(\rho)$  and covariance matrix  $\Sigma(\rho)$ . Under this assumptions, one can show that [226]

$$\text{Prob}[\|\Phi_E(\hat{\rho}) - \Phi_E(\rho)\|_2 \leq \alpha] \geq F_\ell(2N\alpha^2). \quad (5.4)$$

Here  $N$  denotes the number of repetitions of the tomographic measurements and the parameter  $\alpha$  determines the upper bound on the level of confidence via the cumulative distribution function  $F_\ell$  of the central  $\chi^2(\ell)$  distribution. The parameter  $\ell$  is the linear dimension of the state space, that is,  $\ell = d^2 - 1$  for a  $d$ -dimensional quantum system. The resulting confidence region is an ellipsoid in the state space.

We now explain how to obtain an outer approximation of the confidence ellipsoid in form of a polytope  $\mathcal{P}$  which has only  $2\ell$  vertices. Notice that there already exist methods to obtain confidence regions in form of a polytope [224]. However, their number of vertices is by far too large such that an inspection of all of them with respect to a certain property is impossible. First notice that Eq. (5.4) can be rewritten as

$$\text{Prob}[\rho \in \hat{\rho} + \alpha\mathcal{E}] \geq F_\ell(2N\alpha^2), \quad (5.5)$$

where

$$\mathcal{E} = \{Y : \|\Phi_E(Y)\|_2 \leq 1, Y = Y^\dagger, \text{Tr}[Y] = 0\}. \quad (5.6)$$

The set  $\mathcal{E}$  is an ellipsoid in the set of traceless self-adjoint operators. Therefore, the ellipsoid  $\hat{\rho} + \alpha\mathcal{E}$ , which is the original ellipsoid from Eq. (5.6) but rescaled by a factor of  $\alpha$  and with center  $\hat{\rho}$ , constitutes a confidence region once the parameter  $\alpha$  is chosen according to the intended level of confidence. For example, if one aims a confidence level of  $\gamma = 99\%$ ,  $\alpha$  can be obtained by solving  $F_\ell(2N\alpha^2) = \gamma$ . To obtain an outer approximation of the confidence ellipsoid in the form of a hyperoctahedron, we replace the Euclidean norm  $\|\cdot\|_2$  in the definition of the ellipsoid in Eq. (5.6) by the one-norm  $\|\cdot\|_1$ , which is given by  $\|\vec{v}\|_1 = \sum_j |v_j|$ . Given  $m$  orthonormal vectors  $\{\vec{z}_j\}_{j=1}^m \subset \mathbb{R}^n$  one has for an arbitrary vector  $\vec{v} \in \mathbb{R}^n$

$$\sqrt{\sum_{j=1}^m \langle \vec{b}_j, \vec{v} \rangle^2} \geq \frac{1}{\sqrt{m}} \sum_{j=1}^m |\langle \vec{b}_j, \vec{v} \rangle|. \quad (5.7)$$

Consequently, the set  $\{\vec{v} \in \mathbb{R}^n \mid \sum_j \langle \vec{b}_j, \vec{v} \rangle^2 \leq 1\}$  is contained in the hyperoctahedron  $\{\sqrt{m}\vec{v} \in \mathbb{R}^n \mid \sum_j |\langle \vec{b}_j, \vec{v} \rangle| \leq 1\}$  with extremal points  $\pm\sqrt{m}\vec{b}_j$ . Since we are interested in  $\vec{v} = \Phi_E(Y)$ , we need an orthonormal basis of the range of  $\Phi_E$  over the set of all hermitian operators  $Y$  with zero trace. This yields  $\ell$  orthonormal vectors  $\vec{b}_j = \Phi_E(Y_j)$  with  $Y_j$  an appropriate zero-trace hermitian operator. By rescaling the parameter  $\alpha$  by  $\alpha \mapsto \alpha/\ell$ , one obtains

$$\text{Prob}[\rho \in \hat{\rho} + \alpha\mathcal{P}] \geq F_\ell\left(\frac{2N}{\ell}\alpha^2\right), \quad (5.8)$$

where  $\mathcal{P}$  is the hyperoctahedron spanned by  $\{\pm Y_j\}_j$ . Note that both, the sphere and the polytope, are deformed by the linear map  $\Phi_E$ , see also Fig. 5.2 for an illustration. The deformation of the sphere and the polytope depends on the particular choice of measurement directions  $E$ .

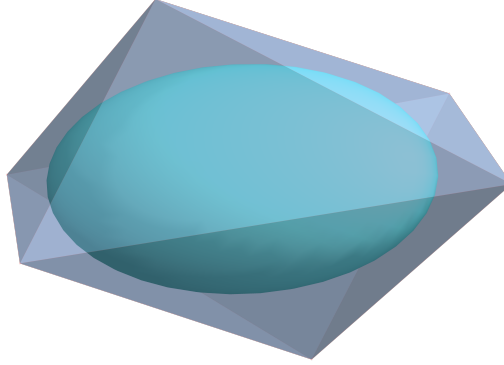


Figure 5.2: Illustration of the outer approximation of a confidence region by a polytope. The inner ellipsoid is the confidence region in state space that has been obtained from the Gaussian approximation in Eq. (5.4). The outer polytope corresponds to the outer approximation by a hyperoctahedron. The figure is taken from Ref. [C].

### 5.3 Steerability with dichotomic measurements

In order to demonstrate the activation of Bell-nonlocality or quantum steerability in a statistical rigorous manner, we have to certify that the initial state is local. This implies the necessity of a method to decide locality for a generic quantum state. Here we show that deciding whether a bipartite quantum state  $\rho$  admits a local hidden state (LHS) model for dichotomic measurements can be solved asymptotically for arbitrary  $\rho \in \mathcal{B}(\mathbb{C}^{d_A} \otimes \mathbb{C}^2)$ , i.e., for qudit-qubit systems. This procedure relies on a reformulation of the problem as a nesting problem of two convex objects [231, 233].

In order to determine whether a LHS model can be constructed for a quantum state  $\rho \in \mathcal{B}(\mathbb{C}^{d_A} \otimes \mathbb{C}^2)$  with respect to dichotomic measurements, denoted by  $\text{LHS}_2$ , one defines the so called critical radius as

$$\mathcal{R}(\rho) := \max \{t \geq 0 \mid \rho_t \text{ admits a } \text{LHS}_2 \text{ model}\}, \quad (5.9)$$

where

$$\rho_t = t\rho + (1-t)\frac{\mathbb{1}}{d_A} \otimes \text{Tr}_A[\rho]. \quad (5.10)$$

It is important to notice that in the definition of the critical radius  $\mathcal{R}$  we only assumed dichotomic measurements. Clearly, we have  $\mathcal{R}(\rho) \geq 1$  if and only if  $\rho$  admits a  $\text{LHS}_2$  model. Recall that the system of Bob is given by a qubit and therefore the set of pure quantum states corresponds to a two-dimensional sphere. We will denote the set of pure states on Bob's side by  $\mathcal{S}$ . Now let  $\mu$  be a probability measure on  $\mathcal{S}$ . We denote by  $\mathcal{K}(\mu)$  the set of all quantum states that Alice can simulate at Bob's side using the

measure  $\mu$ . This set is called the simulability capacity of  $\mu$  and is formally defined as

$$\mathcal{K}(\mu) = \left\{ \int_{\mathcal{S}} g(\sigma) \sigma \, d\mu(\sigma) \mid g : \mathcal{S} \rightarrow [0, 1] \right\}. \quad (5.11)$$

Let  $\mathcal{M}_A$  denote the set of all measurement effects of Alice. Given that the measurement outcome associated with the effect  $E \in \mathcal{M}_A$  was observed, the corresponding conditional state on Bob's side is given by  $\text{Tr}_A[\varrho(E \otimes \mathbf{1})]$ . Consequently, the state  $\varrho_t$  admits a LHS model with respect to dichotomic measurements if the conditional states on Bob's side are a subset of the set of states Alice can simulate for certain probability measures  $\mu$  [233], that is,

$$\{\text{Tr}_A[\varrho(E \otimes \mathbf{1})] \mid E \in \mathcal{M}_A\} \subset \mathcal{K}(\mu). \quad (5.12)$$

The notion of the simulability capacity allows us to rewrite the critical radius as

$$\begin{aligned} \mathcal{R}(\varrho) = \text{maximize} \quad & t \\ \text{such that} \quad & \text{Tr}_A[\varrho_t(E \otimes \mathbf{1})] \in \mathcal{K}(\mu) \quad \forall E \in \mathcal{M}_A \\ & \text{with respect to } t, \mu. \end{aligned} \quad (5.13)$$

Here it is important to note that the nesting condition in Eq. (5.12) can be rephrased by using the dual representation of convex sets.

**Lemma 39** ([233, 235]). *Let  $X$  be a compact convex subset of a finite dimensional Euclidean space. Then, a compact subset  $Y$  is contained in  $X$  if and only if*

$$\max_{\vec{x} \in X} \langle \vec{z}, \vec{x} \rangle \geq \max_{\vec{y} \in Y} \langle \vec{z}, \vec{y} \rangle \quad (5.14)$$

for all vectors  $\vec{z}$  in the Euclidean space.

Therefore, one can equivalently rewrite Eq. (5.12) as

$$\max_{\tau \in \mathcal{K}(\mu)} \langle F, \tau \rangle \geq \max_{E \in \mathcal{M}_A} \text{Tr}[\varrho_t E \otimes F] \quad (5.15)$$

and this condition has to hold for all hermitian operators  $F$  acting on Bob's system. Moreover, using the definition of the simulability capacity  $\mathcal{K}(\mu)$ , one can solve the maximization of the left-hand side of Eq. (5.15) explicitly, which gives

$$\max_{\tau \in \mathcal{K}(\mu)} \langle F, \tau \rangle = \max_{g: \mathcal{S} \rightarrow [0, 1]} \langle F, \int_{\mathcal{S}} g(\sigma) \sigma \, d\mu(\sigma) \rangle = \max_{g: \mathcal{S} \rightarrow [0, 1]} \int_{\mathcal{S}} \langle F, \sigma \rangle g(\sigma) \, d\mu(\sigma). \quad (5.16)$$

Clearly, as the function  $g : \mathcal{S} \rightarrow [0, 1]$  is arbitrary, the optimal choice is simply given by the indicator function with respect to the set of states for which  $\langle F, \sigma \rangle > 0$ . Therefore, one arrives at

$$\max_{\tau \in \mathcal{K}(\mu)} \langle F, \tau \rangle = \int_{\mathcal{S}} \max\{\langle F, \sigma \rangle, 0\} \, d\mu(\sigma). \quad (5.17)$$

Combining Eq. (5.13) and Eq. (5.17) allows us to rewrite the critical radius as

$$\begin{aligned} \mathcal{R}(\varrho) = & \text{maximize} && t \\ & \text{such that} && \int_{\mathcal{S}} \max \{ \langle F, \sigma \rangle, 0 \} d\mu(\sigma) \geq \max_{E \in \mathcal{M}_A} \text{Tr}[Q_t E \otimes F] \forall F \\ & \text{with respect to} && t, \mu. \end{aligned} \tag{5.18}$$

This is an optimization problem over the set of probability measures supported on the Bloch sphere  $\mathcal{S}$  of Bob. Further, the optimization involves an infinite number of constraints, as the whole set of hermitian operators  $F$  has to be considered. The crucial step to solve this problem is to introduce a polytope  $\mathfrak{P} \subset \mathbb{R}^3$  which is the convex hull of  $v \in \mathbb{N}$  vertices,  $\mathfrak{P} = \text{conv}(\{\sigma_j\}_{j=1}^v)$ , approximating Bob's Bloch sphere  $\mathcal{S}$  from inside or from the outside, which gives a lower or an upper bound on the critical radius  $\mathcal{R}(\varrho)$ , respectively. Upon approximating the Bloch sphere by a polytope  $\mathfrak{P}$ , the probability measure  $\mu$ , originally supported on  $\mathcal{S}$ , turns into a probability distribution supported only on the vertices of  $\mathfrak{P}$ . We will write  $\omega = \{\omega_j\}_{j=1}^v$  for this distribution, where  $\omega_j = \mu(\sigma_j)$  denotes the mass of the measure  $\mu$  in the point  $\sigma_j$ . Given such a polytope approximation  $\mathfrak{P}$ , the capacity  $\mathcal{K}(\mu)$  of  $\mu$  also turns into a polytope, denoted by  $\mathcal{K}_{\mathfrak{P}}(\omega)$ , in the operator space of Bob's system. As a consequence, the constraints in the optimization problem in Eq. (5.18) simplify as one only has to consider those hermitian operators  $F$  corresponding to normal vectors of the facets of  $\mathcal{K}_{\mathfrak{P}}(\omega)$ , which are of finite number. This is in stark contrast to the original (exact) optimization problem in Eq. (5.18), where an infinite number of constraints has to be considered. Crucially, these normal vectors of  $\mathcal{K}_{\mathfrak{P}}(\omega)$  are only dependent on the polytope approximation  $\mathfrak{P}$  of the Bloch sphere and independent of the particular probability weights  $\omega$  on the vertices of the polytope. In order to find the normal vectors of the facets of the capacity  $\mathcal{K}_{\mathfrak{P}}(\omega)$  for the given polytope  $\mathfrak{P}$ , one can proceed as follows. For a certain operator  $F$  and a probability measure of the form  $\mu = \sum_{j=1}^v \omega_j \delta_{\sigma_j}$  one finds

$$\max \{ \langle F, \tau \rangle \mid \tau \in \mathcal{K}_{\mathfrak{P}}(\omega) \} = \sum_{j=1}^v \omega_j \max \{ \langle F, \sigma_j \rangle, 0 \}, \tag{5.19}$$

where  $\delta_{\sigma_j}$  denotes the point measure in the vertex  $\sigma_j$ . The operator  $F$  corresponds to a facet of  $\mathcal{K}_{\mathfrak{P}}(\omega)$  if the maximizers on the left-hand side in Eq. (5.19) form a hyperplane in the four dimensional vector space over Bob's system. The solution of the left-hand side gives the maximizers  $\tau^*$  as

$$\tau^* = \sum_{j \in A} \omega_j \sigma_j + \sum_{j \in B} \omega_j \xi_j \sigma_j, \tag{5.20}$$

where  $A := \{j \mid \langle F, \sigma_j \rangle > 0\}$  and  $B = \{j \mid \langle F, \sigma_j \rangle = 0\}$  with any  $\xi_j \in [0, 1]$ . One can see that this forms a hyperplane if there are at least three points in  $B$ , that is, there are

three  $\sigma_j \in \mathfrak{F}$  such that  $\langle F, \sigma_j \rangle = 0$ . This means that  $F$  defines a plane that goes through at least three points of  $\mathfrak{F}$ . In the following we will denote the set of all operators  $F$  that define a plane going through at least three points of  $\mathfrak{F}$  by  $\mathfrak{F}(\mathfrak{F})$ . Here it is important to notice that  $\mathfrak{F}(\mathfrak{F})$  is independent of the probability weights  $\omega$ . However, unlike to the case where only a two-qubit system is considered, in the qudit-qubit case, one has to face additional difficulties. The problem is that even under this polytope approximation, the optimization problem in Eq. (5.18) is not yet a linear program. Indeed, the right-hand side of the constraint in Eq. (5.18) still depends on the parameter  $t$  in a complicated way. However, this complication can be overcome as we will explain in the following. As the objective function is linear in  $E$ , it follows that the maximum is attained at an extreme point of the set  $\mathcal{M}_A$ . The extreme points of  $\mathcal{M}_A$  are exactly projections of rank  $\ell$  with  $\ell = 0, 1, \dots, d_A$  where  $d_A$  denotes the dimension of Alice's system. Consequently one has

$$\max_{E \in \mathcal{M}_A} \text{Tr}[q_t E \otimes F] = \max_{0 \leq \ell \leq d_A} \max_{\substack{E \in \mathcal{M}_A \\ \text{Tr}[E] = \ell}} \text{Tr}[q_t E \otimes F]. \quad (5.21)$$

Therefore, the right-hand side of Eq. (5.21) can be rewritten as

$$t \max_{E \in \mathcal{M}_A} \max_{\text{Tr}[E] = \ell} \text{Tr}[qE \otimes F] + \frac{(1-t)\ell}{d_A} \text{Tr}[q_B F]. \quad (5.22)$$

In total, we end up with a linear program which is given by

$$\begin{aligned} \mathcal{R}_{\mathfrak{F}}(q) = \text{maximize } & t \\ \text{subject to } & \sum_{j=1}^{\nu} \omega_j \max\{\langle F, \sigma_j \rangle, 0\} \geq t\eta_{\ell}(F) + (1-t)\frac{\ell}{d_A} \text{Tr}[q_B F] \\ & \sum_{j=1}^{\nu} \omega_j = 1, \end{aligned} \quad (5.23)$$

where the optimization is with respect to  $t$  and the probability weights  $\{\omega_j\}$ . Further, the constraint in Eq. (5.23) has to hold for all ranks  $0 \leq \ell \leq d_A$  and all operators  $F \in \mathfrak{F}(\mathfrak{F})$ . In addition we have introduced the function

$$\eta_{\ell}(F) := \max \{ \text{Tr}[q(E \otimes F)] \mid E \in \mathcal{M}_A, \text{Tr}[E] = \ell \}. \quad (5.24)$$

Here it is important to note that  $\eta_{\ell}(F)$  is simply the sum over the  $\ell$  maximal eigenvalues of the operator  $\text{Tr}_B[q\mathbb{1} \otimes F]$ . Therefore, for a polytope  $\mathfrak{F}$  with  $\nu$  vertices, the linear program consists of  $\mathcal{O}(\nu)$  variables and  $\mathcal{O}(d_A \nu^3)$  constraints, which can be efficiently solved. Consequently, the size of the program is only linear in Alice's dimension, rendering the analysis of general qudit-qubit states possible. Notice that previous methods are based on the approximation of the set of measurements by a polytope [229, 230, 232, 234], which yields a semidefinite program with an exponential

size in the approximation polytope. These demand much higher computational resources and cannot be practically applied to states that are of interest for the activation of nonlocality.

## 5.4 Activation of nonlocality by local filtering

With the tools provided in Section 5.2 and Section 5.3 we can answer the question of how many samples are needed in order to statistically conclusive verify the activation of Bell nonlocality or quantum steerability. For both types of correlations, we consider in the activation protocol a family of qutrit-qubit states for the system of Alice and Bob given by

$$\rho_{\mu,q} := qW_\mu + (1-q)|2\rangle\langle 2| \otimes \frac{\mathbb{1}}{2}, \quad (5.25)$$

where  $0 \leq \mu \leq 1$  and  $0 \leq q \leq 1$ . Here  $W_\mu$  is a qutrit-qubit embedded Werner state, that is,

$$W_\mu = \mu|\psi^-\rangle\langle\psi^-| + (1-\mu)(|0\rangle\langle 0| + |1\rangle\langle 1|) \otimes \frac{\mathbb{1}}{4} \quad (5.26)$$

and  $|\psi^-\rangle$  refers to a qutrit-qubit state that is only supported on the first two levels  $|0\rangle, |1\rangle$  of the qutrit system, i.e.,  $|\psi^-\rangle = (1/\sqrt{2})(|01\rangle - |10\rangle)$ . Such a state arises naturally for photonic qubits, when there is a loss of photons with probability  $q$  on Alice's side and the vacuum state is explicitly considered [236,237].

In order to demonstrate the activation of nonlocality, Alice applies on her side the local filter  $F = |2\rangle\langle 2|$ . Upon implementing the measurement  $(F, \mathbb{1} - F)$  she sends the measurement outcome to Bob. If the first outcome occurs, the resulting state will be a product state and is discarded by Alice and Bob. If the second outcome is observed, one finds that the post-measured state is given by

$$\text{Tr}[(\mathbb{1} - F)\rho_{\mu,q}(\mathbb{1} - F)]^{-1} (\mathbb{1} - F)\rho_{\mu,q}(\mathbb{1} - F) = W_\mu. \quad (5.27)$$

In this case, Alice and Bob know that they share the Werner state  $W_\mu$  and they are ready to perform a Bell test using this state. If now  $\mu > 1/\sqrt{2}$ , the state can violate the CHSH inequality [6]. However, it should be noticed that the projection of the state onto  $|0\rangle\langle 0| + |1\rangle\langle 1|$  is in general very challenging to implement on a photonic platform.

As already pointed out, even before filtering and before performing a standard Bell test, one has to perform tomography of the initially prepared system to certify its locality. For the specific scenario of one qutrit and one qubit as in Eq. (5.25) we choose for the tomographic measurement the Pauli operators  $\sigma_1, \sigma_2, \sigma_3$  for the qubit and the operators  $\sigma_{0,1}, \sigma_{0,2}, \sigma_{0,3}, \sigma_{1,1}, \sigma_{1,2}, \sigma_{1,3}, \sigma_{2,1}, \sigma_{2,2}$  for the qutrit, where  $\sigma_{j,\mu}|j\rangle = |j\rangle$  and  $\sigma_{j,\mu}|k\rangle = \sigma_\mu|\ell\rangle$  with  $\ell = k$  for  $k < j$  and  $\ell = k - 1$  for  $k > j$ . We only include 3 of the four outcomes for each measurement setting, which yields a smaller vector  $\Phi_E(\rho)$

while still the analysis in Section 5.2 is applied. Notice that there is a freedom in the orientation of the hyperoctahedron since the operators  $Y_j$  only need to satisfy the orthogonality conditions. We choose the operators at random, such that the orientation of the polytope does not prefer any specific direction.

### 5.4.1 Locality of the targeted state family

Let us now determine for which values of  $0 \leq \mu \leq 1$  and  $0 \leq q \leq 1$  the states in Eq. (5.25) are local with respect to generalized measurements. First observe that the existence of a LHS model implies the existence of a LHV model. The idea is to reduce the construction of a LHS model for all generalized measurements to that for dichotomic measurements. We will make use of the following lemma, which was first introduced in Ref. [238].

**Lemma 40** ([238, 239]). *Let  $\rho \in \mathcal{B}(\mathbb{C}^{d_A} \otimes \mathbb{C}^{d_B})$  and suppose that  $\rho$  has a LHS<sub>2</sub> model. Then, the state*

$$\tilde{\rho} := \frac{1}{d_A} \rho + \frac{d_A - 1}{d_A} \sigma_A \otimes \text{Tr}_B[\rho] \quad (5.28)$$

with  $\sigma_A \in \mathcal{B}(\mathbb{C}^{d_A})$  any arbitrary state admits a LHS model for generalized measurements.

Here it is important to notice that the proof of Lemma 40 in Ref. [239] remains valid even when the operator  $\rho$  is not positive semidefinite, as long as the conditional states on Bob's side remain positive. It has been shown that  $\rho_{\mu,q}$  admits a LHS<sub>2</sub> model if  $q \leq 2(1 - \mu)$  for  $1/2 \leq \mu \leq 1$  or  $1/2 \leq \mu \leq 1$ . Using Lemma 40, one can directly see that the state  $\rho_{\mu,q}$  admits a LHS model for all generalized measurements if  $q \leq \frac{2}{3}(1 - \mu)$  with  $1/2 \leq \mu \leq 1$  or  $q \leq 1/3$  with  $0 \leq \mu \leq 1/2$ . This forms the area framed by the dotted boundary in Fig. 5.3 (left).

However, it is known that the Werner state  $W_\mu$  admits a LHV model with respect to arbitrary generalized measurements for  $\mu \leq 5/12$  [214], which is not included in the area framed by the dotted boundary in Fig. 5.3 (left) we derived above. The idea is to consider the convex hull of this area and the point at  $(\mu = 5/12, q = 1)$ . Notice that  $q$  and  $\mu$  parametrize the state space non-linearly. In order to carry out convex geometry operations, we observe that  $\rho_{\mu,q}$  is the convex combination of three points in the state space, namely  $\Pi \otimes \mathbb{1}$  with  $\Pi = |0\rangle\langle 0| + |1\rangle\langle 1|$ ,  $|\psi^-\rangle\langle \psi^-|$  and  $|2\rangle\langle 2| \otimes \mathbb{1}$ , see Fig. 5.3 (right). A point in this triangle represents a valid quantum state  $\rho_{\mu,q}$  and the relation to the parameters  $\mu$  and  $q$  is illustrated in Fig. 5.3 (right). Here the polygonal area in Fig. 5.3 (left) is now no longer a polygon. The convex hull of this area is computed by finding a line going through the point corresponding to  $\mu = 5/12$  and  $q = 1$  that is also a tangent of this area. A detailed calculation gives the touching point at  $\mu_0 = \sqrt{22}/4 - \frac{1}{2}$ ,  $q_0 = 1 - \sqrt{22}/6$ . The linear line connecting  $(\mu = \mu_0, q = q_0)$  and



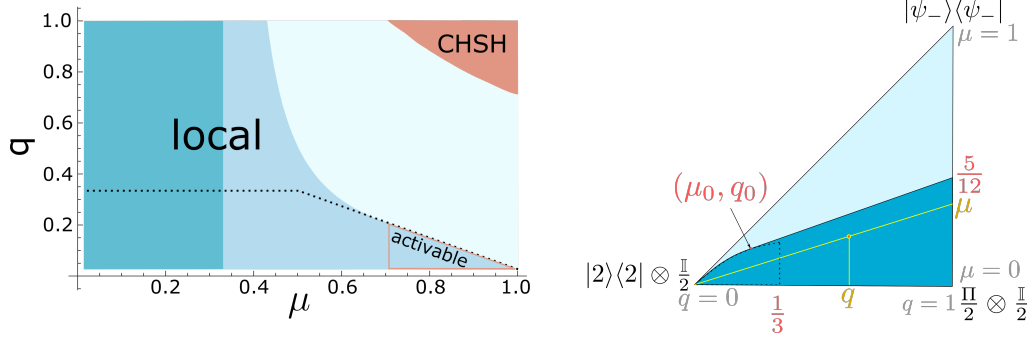


Figure 5.3: (left) Local properties of the family of quantum states  $\rho_{\mu,q}$  defined in Eq. (5.25), depending on the parameters  $\mu$  and  $q$ . The region colored in dark cyan indicates the range of the parameter  $\mu$  for which the Werner state  $W_\mu$  is separable,  $\mu \leq 1/3$ , and hence so is  $\rho_{\mu,q}$ . The curved, blue region represents a set of parameters for which the state can be proved to be local. If  $\mu, q$  are chosen in the activable region (red triangle), that is  $\mu \geq 1/\sqrt{2}$  and  $q \leq \frac{2}{3}(1 - \mu)$ ,  $W_\mu$  violates the CHSH inequality while the state  $\rho_{\mu,q}$  remains local. If the parameter are in that region, they can be used for activation of Bell nonlocality. The locality of the area framed by dotted boundary can be derived from Refs. [236–238] (right) The convex hull of  $\Pi \otimes \mathbb{1}$ ,  $|\psi^-\rangle\langle\psi^-|$  and  $|2\rangle\langle 2| \otimes \mathbb{1}$  is a triangle in the state space. The yellow lines demonstrate how the parameters  $q$  and  $\mu$  for an arbitrary state (orange) in the triangle can be computed. The dotted boundary represents the corresponding dotted boundary in the left figure. The convex hull of this area with the point  $q = 1, \mu = \frac{5}{12}$  can be computed by finding the touching point  $(\mu_0, q_0)$  (red). The figures are taken from Ref. [C].

( $\mu = 5/12, q = 1$ ) in Fig. 5.3 (right) is translated back to

$$q \leq \frac{-29 + 6\sqrt{22}}{-24 + 6\sqrt{22} - 12\mu}. \quad (5.29)$$

### 5.4.2 Application to Bell-nonlocality

The experimental feasibility for the activation of Bell nonlocality depends critically on the number  $N'$  of state preparations that are needed to certify with high confidence that the initial state  $\rho_{\mu,q}$  is local. For this aim we compute the largest scaling factor of the confidence polytope around  $\rho$  such that the polytope is still fully contained in the set of local states. A lower bound on this scaling factor is given by

$$\epsilon_{\mu,q}^* := \min_{\substack{s=\pm 1 \\ 1 \leq j \leq 35}} \max \{ \epsilon \mid \mathcal{R}(\tilde{\rho}_{\mu,q} + s\epsilon Y_j) \geq 1 \}, \quad (5.30)$$

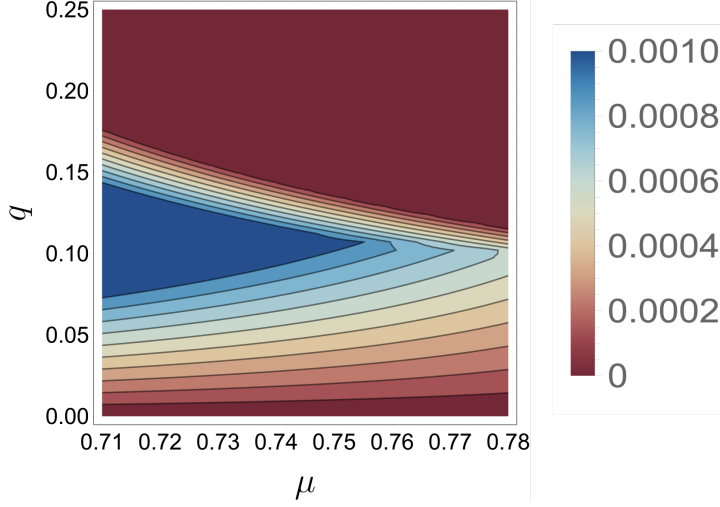


Figure 5.4: Maximal scaling factor  $\epsilon_{\mu,q}^*$  as defined in Eq. (5.30) depending on  $\mu$  and  $q$ . All parameter pairs yielding  $\epsilon_{\mu,q}^* = 0$  (red area) cannot be used for a conclusive activation. As an inner approximation of Bob's Bloch sphere, we have used a icosidodecahedron which has 30 vertices. The figure is taken from Ref. [C].

where  $\{\pm Y_j\}_j$  span the hyperoctahedron  $\mathcal{P}$  with 70 vertices in total introduced in Section 5.2 and  $\tilde{\rho}_{\mu,q}$  is the state defined in Eq. (5.28). However, it is important to notice that this lower bound  $\epsilon_{\mu,q}^*$  is in general not tight. Indeed, there exist quantum states that do not admit a LHS model but are Bell local. In Fig. 5.4 we display  $\epsilon_{\mu,q}^*$  as a function of the parameters  $\mu$  and  $q$  for the region relevant for the activation for Bell nonlocality using the CHSH inequality. One observes that a large area in the parameter space yields roughly the same maximal value  $\epsilon^* \approx 0.001$ , rendering the target parameter  $q$  and  $\mu$  robust to experimental imperfections.

It now follows from our statistical analysis that a sufficient number of state preparations is given by

$$N' = \frac{24\ell F_\ell^{-1}(\gamma)}{2(\epsilon_{\mu,q}^*)^2}, \quad (5.31)$$

where  $\gamma$  denotes the desired level of confidence, the factor 24 reflects the number of measurement settings per tomography and  $\ell = 35$  is the dimension of the state space. We obtain that  $N' = 7.5 \times 10^8$  state preparations are sufficient for a confidence level of  $3\sigma$ , that is,  $\gamma = 99.7\%$ . Interestingly, if one lowers the confidence level to  $1\sigma$ , that is,  $\gamma = 68.3\%$ , the number of samples does not decrease substantially and one finds  $N' = 4.6 \times 10^8$ .

### 5.4.3 Application to quantum steering

Clearly, the applicability of the methods introduced in Section 5.2 and Section 5.3 are not restricted to the activation of Bell-nonlocality. Indeed, one could also imagine a scenario where Bob can characterize the states he obtains, yielding more information than just the output statistics of the black-box measurements which are then used to obtain a possible violation of the CHSH inequality. This situation corresponds to the activation of steerability by means of local filters. The crucial point is that the Werner state  $W_\mu$  in Eq. (5.26) becomes steerable for smaller parameters  $\mu$ , i.e.,  $W_\mu$  can be steerable but Bell-local.

The possibility of choosing a smaller parameter  $\mu$  yields the existence of larger confidence polytopes in the sense of Eq. (5.30). This implies that a smaller number of samples is needed in order to achieve a predefined confidence level. After the filter operation, instead of demonstrating the violation of the CHSH inequality with the obtained black-box outputs, one shows the violation of a steering inequality, i.e., one witnesses the non-existence of a LHS model for the filtered state. Motivated by the work in Ref. [220], we consider the steering inequality of the form

$$S_M := \frac{1}{M} \sum_{j=1}^M \langle A_j \sigma_{\vec{n}_j}^B \rangle \leq C_M, \quad (5.32)$$

where  $A_j \in \{-1, +1\}$  is a random variable describing the outcome of Alice's measurement, which is still described by a black-box as in the CHSH setting, while Bob explains his outcomes quantum mechanically via the measurement of projective measurements along the directions  $\vec{n}_j \in \mathbb{R}^3$ . Further,  $C_M$  denotes the largest value that  $S_M$  can take when the correlations are explained by means of a LHS model. For the Werner state  $W_\mu$  it is known that in the limit of  $M \rightarrow \infty$  there exist measurement settings for Bob such that  $W_\mu$  is steerable if and only if  $\mu > 1/2$ . However, for the minimal case of  $M = 2$  measurements,  $W_\mu$  is steerable if and only if  $\mu \geq 1/\sqrt{2}$  where Bob's measurement directions are given by the eigenvectors of  $\sigma_1$  and  $\sigma_2$ , forming a square on the Bloch sphere. Hence in this setting we do not obtain a decrease of the allowed parameter space of  $\mu$ . However, if the number of measurements is increased and the directions are chosen in an appropriate manner, the value of  $C_M$  can be lowered significantly. In fact, it is easy to show that if the state  $W_\mu$  is prepared, the value of  $S_M$  in Eq. (5.32) will be given by  $\mu$ . If  $M = 6$  and the measurement directions on Bob's side are given by an octahedron, then one will find  $C_6 \approx 0.5393$ . Consequently, the parameter regime can be chosen larger, i.e., one allows for  $\mu \in [0.5393, 1]$ , which will offer the possibility for larger confidence polytopes, see also Fig. 5.5. Indeed, one finds that for  $\mu \approx 0.5410$  and  $q \approx 0.1836$  the maximal size of the polytope is given by  $\epsilon \approx 0.048$ , which turns out to be optimal for the allowed parameter region. More generally, increasing the number of measurement directions in Eq. (5.32) does not yield a significant smaller number of required samples. For instance, choosing  $M = 10$  measurement directions,

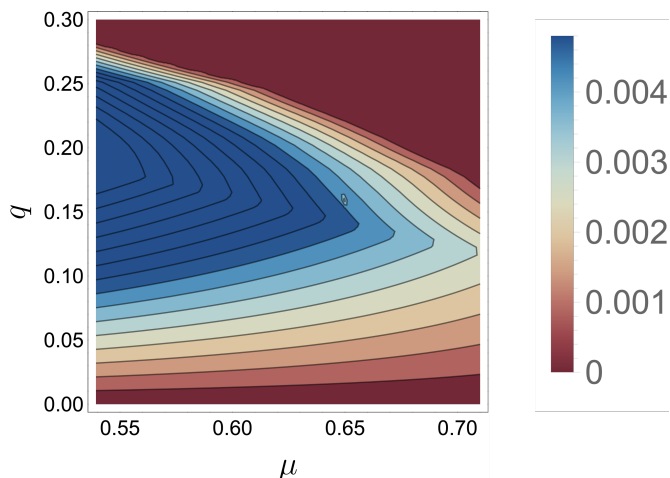


Figure 5.5: Maximal scaling factor  $\epsilon_{\mu,q}^*$  as defined in Eq. (5.30) depending on  $\mu$  and  $q$ . Here one allows for smaller parameters  $\mu$  as in Section 5.4.2, as  $W_\mu$  become steerable before it become Bell-nonlocal. All parameter pairs yielding  $\epsilon_{\mu,q} = 0$  (red area) cannot be used for conclusive activation. As an inner approximation of Bob's Bloch sphere, we have used a icosidodecahedron which has 30 vertices. The figure is taken from Ref. [C].

the steerability of  $W_\mu$  can be revealed by  $S_{10}$  for  $\mu \geq 0.5236$  [220], which is a marginal improvement compared to  $S_6$ .

Applying our statistical analysis to the case of activation of quantum steering for  $M = 6$  measurements for the prepared state with parameters ( $\mu \approx 0.5410, q \approx 0.1836$ ) yielding  $\epsilon^* \approx 0.048$  we find that for a confidence level of  $3\sigma$   $N' = 3.2 \times 10^7$  samples are sufficient. Similar to the activation of Bell nonlocality, lowering the confidence level does not yield a substantial improvement in the number of samples. For instance, to achieve a confidence level of  $1\sigma$  one still needs  $N' = 2.0 \times 10^7$  samples.

## 5.5 Conclusion and discussion

In this Chapter, we developed methods to tackle the two major theoretical problems hindering a conclusive activation of Bell-nonlocality and quantum steerability. First, we introduce a confidence polytope with only  $\mathcal{O}(d^2)$  vertices with  $d$  the dimension of the quantum system which can in addition be efficiently computed. Second, we provide an efficient method to verify whether a quantum state admits a local hidden state model and thus being Bell local. In particular, our method only scales linearly in the dimension  $d_A$  of Alice's system. The combination of both methods allow us to obtain a sufficient number of state preparations that are needed to demonstrate that the initially prepared state is indeed Bell local. This number is likely to be in

the reach of near-future experimental setups. However, it is important to notice that the developed methods are more general and their applicability is going beyond these particular scenarios and can be used for the general certification of constraint quantum correlations. For instance, such scenarios could involve the certification of one-way steerability or the activation of bound entanglement via local filters. Further, also for demonstrating the activation of correlations in the multi-copy scenarios, one has to prove that the initial states are local.

# 6 Finding resourceful multipartite quantum states

The access to multipartite quantum states is an indispensable requirement for many applications in quantum information processing. Although the usefulness of a given quantum state depends on the particular task, entanglement often appears as a resource. Here we design an iterative method for finding maximally resourceful multipartite quantum states. Choosing initially a generic state, we show that in each step of the algorithm the resourcefulness increases. We illustrate the universality of our method by applying it to various different resource quantifiers and present a detailed analysis for the geometric measure of entanglement. Finally, we identify for moderate sizes the corresponding states, revealing an interesting connection to AME states as well as novel correlations for states preparable in the triangle network. This Chapter is based on Project [D].

## 6.1 Motivation

Multipartite quantum states are ubiquitous in quantum information science. Certain of them appear as an important resource for quantum information processing and can be used in certain tasks to outperform their classical counterparts [15, 17, 240]. Indeed, so-called magic states turn out to be a resource for fault-tolerant quantum computation [241, 242] while cluster states are resourceful for measurement-based quantum computation [243, 244]. Additionally, the power of quantum metrology heavily relies on the ability to prepare certain multipartite quantum states [245–247]. However, for a particular given task, it is in general very challenging to identify those multipartite states which can yield the largest advantage.

For many applications entanglement of the quantum state has been proven to be a powerful resource [29]. It offers a complex and rich structure resulting in the impossibility of a quantification by means of a single number. This results in a variety of different quantifiers, each emphasizing on a different property which makes a state a valuable resource [248–250]. A prominent example is the notion of absolutely maximally entangled (AME) states, which turn out to be notoriously difficult to characterize [106, 251–256]. Still, the analysis of AME states is important for understanding

quantum error correction and is regarded as one of the central problems in the field of quantum information theory [257, 258].

The *geometric measure of entanglement* [259–262], quantifying the proximity of a quantum state to the set of product states, has an intuitive meaning and also offers multiple operational interpretations. For instance, it relates to multipartite state discrimination using LOCC [76], the additivity of channel capacities [263], quantum state estimation [264] and was also used to describe quantum phase transitions [265–268]. Further, it has been realized that generic quantum states are highly entangled [269]. In complexity theory, identifying maximally entangled states and computing their geometric measures allows for the identification of cases where the MAX- $N$ -local Hamiltonian problem and its product state approximation deviate maximally [270, 271]. So, although high entanglement does not guarantee that a quantum state is useful for all tasks [272–274], finding maximally entangled states has been recognised as a natural and important problem [271]. So far, however, maximally entangled states have only been identified within the low-dimensional family of symmetric qubit states, where their computation is related to the problem of distributing charges on the unit sphere [250, 275, 276], or within the family of graph states that stem from bipartite graphs [277].

Mathematically, the complexity of the task reflects the fact that pure multipartite states are described by tensors. In contrast to the matrix case, notions like ranks and eigenvalues are for tensors much less understood and their computation turns out to be a hard problem [278, 279]. Interestingly, the geometric measure is closely related to the recently introduced concept of tensor eigenvalues [261, 280–283], offering a much more complex structure as the matrix case [284] as well as to the notion of injective tensor norms [285, 286] and matrix permanents [287]. Here, maximally entangled states offer maximal tensor eigenvalues [288] and it was conjectured that the overlap of a multipartite qubit state with the set of product states decreases exponentially in the number of particles [286]. So, the identification of maximally entangled states provides valuable intuition to decide this conjecture.

In this Chapter, we design an iterative method for finding maximally resourceful multipartite quantum states. As we illustrate this method in detail for the geometric measure, we introduce it in Section 6.2 and explain how it can be approximated. In Section 6.3, we present our algorithm for the simple case of three qubits and provide in Section 6.4 a proof of its monotonicity. Then, we present our numerical findings in Section 6.5 and analyse how the algorithm behaves for random starting points in Section 6.6. Afterwards, we generalize our algorithm to entangled subspaces in Section 6.7. In addition, we demonstrate the universality of our method by applying it to the stabilizer rank in Section 6.8, to the Schmidt rank and states with fixed bond dimension in Section 6.9 as well as to states that are preparable in the triangle scenario in Section 6.10. Finally, we discuss in Section 6.11 how our results relate to known

upper bounds on the maximal entanglement that can be present in a quantum system.

## 6.2 Concepts and notation

Any  $n$ -particle qudit state  $|\psi\rangle \in (\mathbb{C}^d)^{\otimes n}$  can be expanded in the local computational bases as

$$|\psi\rangle = \sum_{j_1, \dots, j_n} \psi_{j_1, \dots, j_n} |j_1\rangle \cdots |j_n\rangle, \quad (6.1)$$

where  $\psi_{j_1, \dots, j_n} \in \mathbb{C}$  are the coefficients of the state. Similarly, the set of all product states is given by states  $|\pi\rangle$  that are of the form  $|\pi\rangle = |\pi_1\rangle \otimes \cdots \otimes |\pi_n\rangle$ , where  $|\pi_j\rangle \in \mathbb{C}^d$  is the description of the state hold by the  $j$ th party.<sup>1</sup> Therefore, we can view product states as a submanifold of all states. This motivates a *geometric definition* of entanglement, which is given by measuring the proximity of  $|\psi\rangle$  to the set of product states

$$\min_{|\pi\rangle} \|\ |\psi\rangle - |\pi\rangle \|^2 \quad (6.2)$$

Indeed, intuitively one would assume that a state is more entangled when it is further away from all product states. For a given state  $|\psi\rangle$  the quantity in Eq. (6.2) yields an optimization problem subject to the constraint that  $\langle \pi_j | \pi_j \rangle = 1$ . By considering the Lagrangian dual problem [261] one finds that a necessary condition for optimality of a product state  $|\pi\rangle$  in Eq. (6.2) is given by

$$\langle \psi | \left( \bigotimes_{\substack{l=1 \\ l \neq j}}^n |\pi_l\rangle \right) = \lambda \langle \pi_j |, \quad \left( \bigotimes_{\substack{l=1 \\ l \neq j}}^n \langle \pi_l | \right) |\psi\rangle = \lambda |\pi_j\rangle, \quad (6.3)$$

where  $\lambda$  corresponds to the Lagrange multiplier enforcing the normalization constraint  $\langle \pi | \pi \rangle = 1$ . The system of polynomial equations in Eq. (6.3) is called nonlinear eigenproblem and the value  $\lambda \in [-1, 1]$  is called a tensor eigenvalue of  $|\psi\rangle$ . The largest possible  $\lambda$  for which a solution of Eq. (6.3) exists is called the *entanglement eigenvalue* and the optimizer  $|\pi\rangle$  corresponds to the closest separable state such that

$$\lambda = \max_{|\pi\rangle} |\langle \psi | \pi \rangle|. \quad (6.4)$$

As an entanglement measure should be zero for all separable states, one defines the *geometric measure of entanglement* of a quantum state  $|\psi\rangle$  via [259–262]

$$G(|\psi\rangle) := 1 - \lambda^2(|\psi\rangle), \quad (6.5)$$

<sup>1</sup>More precisely, the Segre embedding refers to the map which allows one to identify the cartesian product of two projective spaces as a projective variety. If  $\mathbb{C}P^m$  and  $\mathbb{C}P^m$  are projective spaces then they can be seen as an embedded submanifold of  $\mathbb{C}P^{(n+1)(m+1)-1}$ .



where  $\lambda$  denotes the largest entanglement eigenvalue of  $|\psi\rangle$  as given in Eq. (6.4). This quantity for pure states can be extended to mixed states via the convex roof construction and is then a proper entanglement monotone [261].

While computing the entanglement eigenvalue  $\lambda$  for a generic pure state  $|\psi\rangle \in (\mathbb{C}^d)^{\otimes n}$  is, in principle, an NP-hard problem in the local dimension  $d$  if  $n > 2$  [279,289], there is a simple seesaw iteration that can be used to find good approximations [290–292]. For a three-partite state  $|\psi\rangle \in (\mathbb{C}^d)^{\otimes 3}$ , the algorithm starts with a random product state  $|a_0b_0c_0\rangle$ . From this we can compute the non-normalized state  $|\tilde{a}\rangle = \langle b_0c_0|\psi\rangle$ , and make the update

$$|a_0\rangle \mapsto |a_1\rangle = \frac{1}{\sqrt{\langle \tilde{a}|\tilde{a}\rangle}}|\tilde{a}\rangle. \quad (6.6)$$

The procedure is repeated for the second qubit  $|b_0\rangle$ , starting in the product state  $|a_1b_0c_0\rangle$ . This is then iterated until one reaches a fixed point. Of course, this fixed point is not guaranteed to be the global optimum, in practice, however, this method works very well.

### 6.3 A simple algorithm for maximizing the geometric measure

We are now going to discuss the algorithm for the case of three qubits as the generalization to arbitrary multiparticle systems is straightforward. As initial state  $|\varphi\rangle$  we choose a random pure three qubit state. Then, we compute its closest product state  $|\pi\rangle$  via the see-saw algorithm described above. We can assume without loss of generality that  $|\pi\rangle = |000\rangle$ . Following Eq. (6.4), we write  $\lambda = |\langle \varphi|\pi\rangle|$  for the maximal overlap of  $|\varphi\rangle$  with the set of all product states. Note that for a generic quantum state the closest product state is unique. The key idea is now to perturb the state  $|\varphi\rangle$  in a way that the overlap with  $|\pi\rangle$  decreases. If  $|\pi\rangle$  is the unique closest product state and the perturbation is small, one can then expect that the overlap with *all* product states decreases. So, we consider the orthocomplement of  $|\pi\rangle = |000\rangle$ , that is, the complex subspace spanned by  $|001\rangle, |010\rangle, |100\rangle, |011\rangle, |101\rangle, |110\rangle, |111\rangle$ . This subspace gives rise to a projection operator  $\Pi = \mathbb{1} - |\pi\rangle\langle\pi|$  and we compute the best approximation of the state  $|\varphi\rangle$  within this subspace, given by  $|\eta\rangle = \Pi|\varphi\rangle/\mathcal{M}$ , where  $\mathcal{M}$  denotes the normalization. Then, we shift the state  $|\varphi\rangle$  in the direction of  $|\eta\rangle$  by some small amount  $\theta > 0$ . Hence the state update rule is given by

$$|\varphi\rangle \mapsto |\tilde{\varphi}\rangle := \frac{1}{\mathcal{N}}(|\varphi\rangle + \theta|\eta\rangle) \quad (6.7)$$

where  $\mathcal{N}$  is a normalization factor. In the next step, we calculate the best rank one approximation to  $|\tilde{\varphi}\rangle$ . This process is iterated until the geometric measure is not increasing under the update rule in Eq. (6.7). In this case, one can reduce the step size or

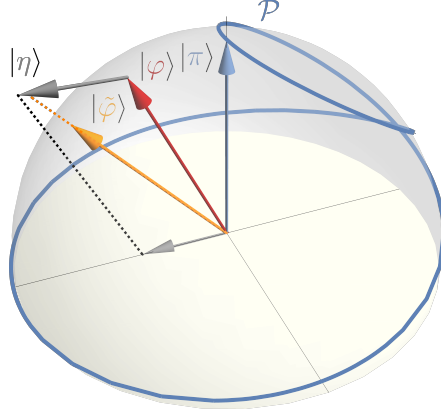


Figure 6.1: Schematic illustration of the iteration step of the algorithm. The set of all states is represented by the half sphere and the set of product states by the lower dimensional manifold  $\mathcal{P}$ . If the algorithm is initialized in state  $|\varphi\rangle$  (blue arrow), we first compute the best approximation within  $\mathcal{P}$ , denoted by  $|\pi\rangle$  (light blue arrow). Then, we compute the projector into the orthocomplement of  $|\pi\rangle$ , which is here given by the  $xy$ -plane. The portion of  $|\varphi\rangle$  within the  $xy$ -plane is given by  $|\eta\rangle$  (gray arrow). The new state  $|\tilde{\varphi}\rangle$  is then the normalized version of  $|\varphi\rangle + |\eta\rangle$ . Here we have set the step size  $\epsilon = 1$ . The figure is taken from Ref. [D].

the algorithm terminates. One can directly check that the overlap with  $|\pi\rangle$  is smaller for  $|\tilde{\varphi}\rangle$  than for  $|\varphi\rangle$ . Indeed, one has

$$|\langle\pi|\tilde{\varphi}\rangle|^2 = \frac{1}{\mathcal{N}^2} |\langle\pi|\varphi\rangle + \theta\langle\pi|\eta\rangle|^2 = \frac{\lambda^2}{\mathcal{N}^2} < \lambda^2 \quad (6.8)$$

since  $\mathcal{N} > 1$  if  $\theta > 0$ . In fact, a much stronger statement holds, which is also one of the main results of this chapter.

**Theorem 41.** *For a generic quantum state  $|\psi\rangle$  there always exists a  $\Theta > 0$  such that the updated state  $|\tilde{\psi}\rangle$  according to Eq. (6.7) with step size  $\theta < \Theta$  fulfills  $G(|\psi\rangle) < G(|\tilde{\psi}\rangle)$ .*

## 6.4 The proof of monotonicity

As pointed out above, the algorithm relies on having access to subroutine able to compute an approximate version of

$$|\pi\rangle = \operatorname{argmin}\{|\pi\rangle : 1 - |\langle\varphi|\pi\rangle|^2 : |\pi\rangle \text{ product state}\}. \quad (6.9)$$

In the following, we will frequently make statements about generic states. In these cases, we require the assumptions that a state is not a product state and that the best

product state approximation  $|\pi\rangle$  is unique up to a phase. The first step in order to prove monotonicity of the algorithm is to show that small variations in the initial state can only lead to small variations in the best product state approximation. As we will see, this follows from the more general observation that under certain conditions the value  $y_0$ , where a function  $f(x_0, y)$  assumes its minimum (for a given  $x_0$ ), depends continuously on  $x_0$ . Further, by virtue of the canonical embedding, we can identify any  $|\varphi\rangle \in \mathbb{C}^n$  with a  $|\tilde{\varphi}\rangle \in \mathbb{R}^{2n}$  and consequently we can omit the absolute in Eq. (6.9).

**Lemma 42.** *Let  $X, Y$  be compact and  $f : X \times Y \rightarrow \mathbb{R}$  be uniformly continuous. Further, suppose that for  $x_0 \in X$  the value  $y_0 := \operatorname{argmin}_{y \in Y} f(x_0, y)$  is unique. Then for all  $\varepsilon > 0$  there exists  $\delta > 0$  such that for all  $x \in U_\delta(x_0)$  we have  $\operatorname{argmin}_{y \in Y} f(x, y) \subset U_\varepsilon(y_0)$ , where  $U_\delta(x_0)$  and  $U_\varepsilon(y_0)$  denote vicinities of  $x_0$  and  $y_0$ , respectively. In other words, the function  $\operatorname{argmin}$  is continuous in  $x_0$ .*

*Proof.* For the given  $\varepsilon$  we can split the set  $Y$  in the vicinity  $U_\varepsilon(y_0)$  and its complement  $\bar{U}_\varepsilon(y_0)$ . In particular we have

$$\begin{aligned} f(x_0, y_0) &= \min_{y \in U_\varepsilon(y_0)} f(x_0, y) \\ &< \min_{y \in \bar{U}_\varepsilon(y_0)} f(x_0, y) =: f(x_0, \tilde{y}_0), \end{aligned} \quad (6.10)$$

that is,  $\tilde{y}_0$  denotes the value where the minimum in  $\bar{U}_\varepsilon(y_0)$  is assumed. Let us denote the difference between the function values as

$$\xi = f(x_0, \tilde{y}_0) - f(x_0, y_0) > 0. \quad (6.11)$$

By the uniform continuity we can choose  $\delta > 0$  such that for all  $\tilde{x} \in \mathbb{R}^n$  with  $\|\tilde{x} - x_0\| < \delta$  and for all  $y$  we have

$$|f(\tilde{x}, y) - f(x_0, y)| < \frac{\xi}{2}. \quad (6.12)$$

Then we have

$$f(\tilde{x}, y_0) < f(x_0, y_0) + \frac{\xi}{2}, \quad (6.13)$$

but for all  $y \in \bar{U}_\varepsilon(y_0)$

$$f(\tilde{x}, y) > f(x_0, \tilde{y}_0) - \frac{\xi}{2} > f(x_0, y_0) + \frac{\xi}{2}, \quad (6.14)$$

which implies that the minimum of  $f(\tilde{x}, y)$  lies in the vicinity  $U_\varepsilon(y_0)$ .  $\square$

**Corollary 43.** *Let  $|\varphi\rangle$  be a pure quantum state and suppose that its best product state approximation  $|\pi\rangle$  is unique. Then, for all  $\tau > 0$ , there exists a  $\xi > 0$  such that the best product state approximation  $|\tilde{\pi}\rangle$  of  $|\tilde{\varphi}\rangle \in \mathcal{U}_\xi(|\varphi\rangle)$  lies in  $\mathcal{U}_\tau(|\pi\rangle)$ .*

*Proof.* The function  $f(x, y) := |\langle x, y \rangle|^2$  is continuous on  $\mathbb{R}^{2n} \times \mathbb{R}^{2n}$ . Further, the space  $B_1 := \{x \in \mathbb{R}^{2n} : \|x\| = 1\}$  is compact and thus also  $M := B_1 \times B_1$ . Then, by the Heine-Cantor theorem [293],  $f$  is uniformly continuous on  $M$ . Since we assume  $|\pi\rangle$  to be unique, we can apply Lemma 42, which guarantees for all  $\tau > 0$  the existence of  $\xi > 0$  such that  $|\tilde{\pi}\rangle \in U_\tau(|\pi\rangle)$  for  $|\tilde{\psi}\rangle \in U_\xi(|\psi\rangle)$ .  $\square$

According to Eq. (6.7), the updated state needs a renormalization given by  $\mathcal{N} = \mathcal{N}(|\psi\rangle, \theta)$ . The next ingredient for the proof of the main result is a Lemma that gives later an upper approximation of the function  $1/\mathcal{N}$ .

**Lemma 44.** *There exists  $C > 0$  such that for all  $q \in [0, 1]$  and  $x > 0$  we have*

$$\frac{1}{\sqrt{1 + 2qx + x^2}} < 1 - qx + Cx^2. \quad (6.15)$$

*More precisely, the above inequality holds for all  $C \geq 3$ .*

*Proof.* As both sides of Eq. (6.15) are positive, we can square them such that the inequality remains true. This yields the equivalent inequality

$$\begin{aligned} 0 &< [2C - 3q^2 + 1]x^2 + 2q[C - 1 + q^2]x^3 \\ &+ [C^2 - C(4q^2 - 2) + q^2]x^4 + 2q[C^2 - C]x^5 + C^2x^6 \\ &=: f_2x^2 + f_3x^3 + f_4x^4 + f_5x^5 + f_6x^6. \end{aligned} \quad (6.16)$$

Now observe that for each of the constants  $f_k = f_k(C)$ , there is a  $C_k > 0$  such that  $f_k(C) \geq 0$  for all  $C \geq C_k$ . Indeed, we have  $C_2 := \max\{(1/2)(3q^2 - 1), 0\}$ ,  $C_3 := 1 - q^2$ ,  $C_5 = C_6 := 1$ . The choice of  $C_4$  depends on whether  $q^2 \geq 1/2$  or not. If we denote  $\alpha = |4q^2 - 2|$ , we obtain for the case  $q^2 < 1/2$  that  $C^2 + \alpha C + q^2 > 0$ , what is trivially fulfilled for any  $C \geq 1$ . If  $q^2 > 1/2$ , we need  $C^2 - \alpha C + q^2 > 0$ . But  $C^2 - \alpha C + q^2 \geq C^2 - \alpha C = C(C - \alpha) > 0$ , we obtain  $C > \alpha$ . In general, we have  $C_4 := \max\{1, |4q^2 - 2|\}$ . This implies that for  $C \geq \tilde{C} := \max\{C_k | k = 2, \dots, 6\}$  all coefficients are positive. Hence  $0 < f_2x^2$  implies  $0 < f_2x^2 + f_3x^3 + f_4x^4 + f_5x^5 + f_6x^6$  if  $x > 0$ . Consequently, it is sufficient to only consider the problem  $0 < f_2x^2$ , what is true for  $C \geq C_7 := (1/2)(3q^2 - 1)$ . Hence, choosing  $C \geq \max\{\tilde{C}, C_7\}$  yields the claim. Taking the maximum over all  $C_k$  with respect to  $q \in [0, 1]$  yields that  $C > 2$ .  $\square$

*Proof of Theorem 41.* Let us start with some step-size  $\theta_0 > 0$  that we will choose in the end appropriately and consider

$$|\tilde{\psi}\rangle = \frac{1}{\mathcal{N}}(|\psi\rangle + \theta_0|\eta\rangle). \quad (6.17)$$

It is important to note that  $|\eta\rangle$  is a normalized state, that is,  $|\eta\rangle = 1/(\sqrt{1 - \lambda^2})(\mathbb{1} - |\pi\rangle\langle\pi|)|\psi\rangle$ . This yields  $\langle\psi|\eta\rangle = \sqrt{1 - \lambda^2}$ .

The best product state approximation  $|\tilde{\pi}\rangle$  of  $|\tilde{\psi}\rangle$  can be parameterized using the old product state, i.e.,  $|\tilde{\pi}\rangle = \sqrt{1-\delta^2}|\pi\rangle + \delta|\chi\rangle$ , for a normalized, appropriately chosen  $|\chi\rangle$  and  $\delta > 0$ . Using  $\langle\pi|\eta\rangle = 0$  and  $|\langle\chi|\eta\rangle| \leq 1$  we obtain

$$\tilde{\lambda} := |\langle\tilde{\pi}|\tilde{\psi}\rangle| = \frac{1}{\mathcal{N}} |\langle\tilde{\pi}|\psi\rangle + \theta_0\delta\langle\chi|\eta\rangle| \leq \frac{1}{\mathcal{N}} (|\langle\tilde{\pi}|\psi\rangle| + |\theta_0\delta\langle\chi|\eta\rangle|) \quad (6.18)$$

$$\leq \frac{1}{\mathcal{N}}(\lambda + \delta\theta_0). \quad (6.19)$$

Using that  $\mathcal{N} = \sqrt{1 + 2\theta_0\sqrt{1-\lambda^2} + \theta_0^2}$  and Lemma 44, there exists  $C > 0$  such that

$$\begin{aligned} \tilde{\lambda} &< (1 - \theta_0\sqrt{1-\lambda^2} + C\theta_0^2)(\lambda + \delta\theta_0) \\ &= \lambda + \delta\theta_0 - \lambda\sqrt{1-\lambda^2}\theta_0 + \theta_0^2[C(\lambda + \delta\theta_0) - \delta\sqrt{1-\lambda^2}] \\ &= \lambda + \theta_0(\delta - \lambda\sqrt{1-\lambda^2}) + \mathcal{O}(\theta_0^2). \end{aligned} \quad (6.20)$$

Note that  $\lambda\sqrt{1-\lambda^2} > 0$ , since  $0 \neq \lambda \neq 1$ . This comes from the fact that a generic state is not a product state with  $\lambda = 1$  and any state has at least some overlap with some product state. So, if  $\delta < \lambda\sqrt{1-\lambda^2}$  we have  $\tilde{\lambda} < \lambda$  for suitably small  $\theta_0$ .

It remains to show that we can guarantee that  $\delta$  obeys this condition. We start with a given value of  $\lambda$  and consider a number  $0 < \delta_1 < \lambda\sqrt{1-\lambda^2}$ . According to Corollary 43, we can find a  $\theta_1 > 0$  such that  $|\tilde{\pi}\rangle \in U_{\delta_1}(|\pi\rangle)$  if  $|\tilde{\psi}\rangle \in U_{\theta_1}(|\psi\rangle)$ . Then, this gives us an upper bound on  $\theta_0$  for Eq. (6.17), so that the resulting  $\delta < \delta_1$  in Eq. (6.20) is small enough to guarantee a negative slope for the linear term. Still,  $\tilde{\lambda} < \lambda$  is not guaranteed, due to the  $\mathcal{O}(\theta_0^2)$  term in Eq. (6.20). But, for the given values of  $\lambda, \delta_1$  and  $C$ , we can also compute from Eq. (6.20) a second threshold  $\theta_2$ , which guarantees  $\tilde{\lambda} < \lambda$ . Then we can finally take in the statement of Theorem 41  $\theta = \min\{\theta_1, \theta_2\}$  and the proof is complete.  $\square$

It is remarkable that in this proof the fact that  $|\pi\rangle$  is a product state was never used. So, the algorithm can also be used if the overlap with states from some other arbitrarily chosen subset of the (pure) state space shall be minimized. In particular, the given subset must not necessarily be presented in form of a smooth submanifold, but could also be discrete. In addition, this allows the extension to arbitrary distance-like measures defined with respect to pure states. Further, the algorithm also applies to the case where each particle may have a different degree of freedom, e.g.,  $\mathbb{C}^2 \otimes \mathbb{C}^3 \otimes \mathbb{C}^5$ . For this, only the subroutine computing the best rank-1 approximation has to be modified.

## 6.5 Maximally entangled states of the geometric measure

After a sufficient number of iterations, the algorithm yields a highly entangled state given in coordinates with respect to a random basis. Hence, generically each component of the tensor is nonzero. However, in order to understand the structure of the

state, we seek for a concise representation in which most of the coefficients vanish. As we consider two states to be equal if there is a LU transformation connecting them, this requires a parametrization of the set of unitary matrices.

### 6.5.1 Finding a concise representation of quantum states

First notice, that  $U(d)$  is the semidirect product of  $U(1)$  with  $SU(d)$  and hence we can restrict to parametrizations of  $SU(d)$ . For qubits, one can make use of the fact that  $SU(2)$  is diffeomorphic to the 3-sphere  $S^3$ . In particular, an arbitrary  $SU(2)$  matrix can be written as

$$U = \begin{pmatrix} \alpha & -\beta^* \\ \beta & \alpha^* \end{pmatrix}, \quad \alpha, \beta \in \mathbb{C} \text{ with } |\alpha|^2 + |\beta|^2 = 1. \quad (6.21)$$

Consequently, the parametrization involves four real parameters and one quadratic constraint. However, for  $d \geq 3$  the  $SU(d)$  is not homeomorphic to a sphere anymore, or a product of them, e.g.,  $SU(3)$  is *not* homeomorphic to  $S^8$ . Consequently, for systems of higher dimensions different approaches exist [294–296]. Here we have used the Jarlskog parametrization [295], what is a simple recursive scheme for parametrization, which can be easily implemented numerically. First, notice that any  $X \in U(d)$  can be written as  $X = \Phi_\alpha Y \Phi_\beta$  where  $\Phi_\alpha = \text{diag}(e^{i\alpha_1}, \dots, e^{i\alpha_d})$ ,  $\Phi_\beta$  similar and  $Y$  a unitary  $d \times d$  matrix. Now,  $Y$  is decomposed into a product of unitaries, that is,  $Y = \prod_{k=2}^d A_{d,k}$  with

$$A_{d,k} = \begin{pmatrix} A^{(k)} & 0 \\ 0 & \mathbb{1}_{d-k} \end{pmatrix}, \quad (6.22)$$

$$U(d) \ni A^{(k)} = \begin{pmatrix} \mathbb{1}_{d-1} - (1 - \cos(\theta_k))|a_k\rangle\langle a_k| & \sin(\theta_k)|a_k\rangle \\ -\sin(\theta_k)\langle a_k| & \cos(\theta_k) \end{pmatrix}, \quad (6.23)$$

where  $|a_k\rangle \in \mathbb{C}^{d-1}$  normalized to one, i.e.,  $\langle a_k|a_k\rangle = 1$  and  $\theta_k \in [0, 2\pi)$  an arbitrary angle.

We now describe how this parametrization can be used to bring the numerically found states into a concise form. Here, two different cases can be considered. If one has a guess for the possible state, e.g., the marginals are all maximally mixed so one expects an AME or  $k$ -uniform state, one could compute the fidelity between the numerical state  $|\psi\rangle$  and the guess  $|\varphi_{\text{guess}}\rangle$ , i.e.,  $\sup |\langle \psi | U_1 \otimes \dots \otimes U_n | \varphi_{\text{guess}} \rangle|$ . If there is no possible candidate, the idea is to minimize a function  $f : U(d) \times \dots \times U(d) \rightarrow \mathbb{R}$  depending on the state, which becomes minimal if many entries of the state vanish. For instance, given the state  $|\psi\rangle$  a natural candidate would be  $f(U_1, \dots, U_n) = \sum |\langle U_1 \otimes \dots \otimes U_n | \psi \rangle|_{i_1, \dots, i_n}|$ . Given two states  $|\phi\rangle, |\psi\rangle$  we regard them as equal, if  $\mathcal{F}(|\psi\rangle, |\phi\rangle) \geq 1 - \epsilon$  with  $\epsilon < 10^{-6}$ , where  $\mathcal{F}$  denotes the fidelity.

### 6.5.2 Results for qubit systems

Here we present the numerical findings of our algorithm for the case of multi-qubit systems. A concise summary of the entanglement properties of the states found is given in Tab. 6.1. For the minimal case of a two-qubit system, the maximally entangled state with respect to the geometric measure is the Bell state, and the algorithm directly converges to its maximum, see Fig. 6.3. For the larger case of three qubits, there are two different classes of genuine multipartite entanglement with respect to SLOCC, namely  $|W\rangle$  and  $|\text{GHZ}\rangle$ , given by

$$|W\rangle = \frac{1}{\sqrt{3}}(|001\rangle + |010\rangle + |100\rangle), \quad |\text{GHZ}\rangle = \frac{1}{\sqrt{2}}(|000\rangle + |111\rangle). \quad (6.24)$$

While  $G(|\text{GHZ}\rangle) = 1/2$ , one has  $|W\rangle = 5/9$ , what turns out to be the maximizer among all tripartite qubit states [297]. It should be noted, that in this case the maximizer belongs to the family of symmetric states. Operationally, the  $W$  state is the state with the maximal possible bipartite entanglement in the reduced two-qubit states [67]. For four qubits, the algorithm yields after 300 iterations the state

$$|\tilde{M}\rangle = \frac{1}{\sqrt{3}}(|\text{GHZ}\rangle + e^{\frac{2\pi}{3}i}|\text{GHZ}_{34}\rangle + e^{\frac{4\pi}{3}i}|\text{GHZ}_{24}\rangle), \quad (6.25)$$

where  $|\text{GHZ}_{ij}\rangle$  means a four-qubit GHZ state where a bit flip is applied at party  $i$  and  $j$ , that is,

$$|\text{GHZ}_{ij}\rangle = X_i \otimes X_j |\text{GHZ}\rangle. \quad (6.26)$$

Note that the phases that appear in the superposition in Eq. (6.25) form a trine in the complex plane and that the state is a phased Dicke state [298]. This state can be shown to be LU equivalent to the so called Higuchi-Sudbery or  $M$  state [254, 299], which appears as maximizer of the Tsallis  $\alpha$ -entropy in the reduced two-particle states for  $0 < \alpha < 2$ . Note that this state is not symmetric with respect to permutations of the parties. Similar to the  $W$  state, the entanglement of the state in Eq. (6.25) appears to be robust, that is, uncontrolled decoherence of one qubit does not completely destroy the entanglement of the remaining qubits [299]. Further, as a four-qubit AME state does not exist, the  $M$  state can be viewed as the best possible replacement, since the one-body marginals are maximally mixed and all two-body marginals, albeit not maximally mixed, have the same spectrum.

For five qubits the algorithm converges to a state  $|G_5\rangle$ , that can be identified with the ring cluster state, which is a 5-cycle graph state, yielding a geometric measure of  $0.86855 \approx (1/36)(33 - \sqrt{3})$ . The state  $|G_5\rangle$  appears also in the context of the five-qubit error correcting code [300]. In a similar manner, for six qubits we obtain a graph state  $|G_6\rangle$  with a measure of  $0.9166 \approx 11/12$ . This is again connected to quantum error correction. Indeed, both states  $|G_5\rangle$  and  $|G_6\rangle$  are AME states. For seven qubits we find

$n$	$G_{\max}^{\text{symm}}$	$G_{\max}$	$ \varphi\rangle_{\max}$
2	1/2	1/2	$ \psi^-\rangle$
3	$0.5555 \approx 5/9$	$0.5555 \approx 5/9$	$ \text{W}\rangle$
4	$0.6666 \approx 2/3$	$0.7777 \approx 7/9$	$ \text{M}\rangle$
5	$\approx 0.7006$	$0.8686 \approx (1/36)(33 - \sqrt{3})$	$ \text{G}_5\rangle$
6	$0.7777 \approx 7/9$	$0.9166 \approx 11/12$	$ \text{G}_6\rangle$
7	$\approx 0.7967$	$\geq 0.941$	$\text{MMS}(7, 2)$

Table 6.1: Maximally entangled states found by the algorithm for systems between two and seven qubits. Here  $|\varphi\rangle_{\max}$  refers to the state found by algorithm and  $G_{\max}$  denotes the geometric measure of the corresponding state.  $G_{\max}^{\text{symm}}$  denotes the maximal entanglement among symmetric states, as shown in [275]. The table is taken from Ref. [D].

a numerical state with maximally mixed two-body marginals, where the spectra of the three-body marginals are all the same. This gives rise to the notion of maximally marginal symmetric states.

**Definition 45.** Let  $|\psi\rangle \in (\mathbb{C}^d)^{\otimes n}$  be a pure  $n$ -partite qudit state. We call  $|\psi\rangle$  maximally marginal symmetric (MMS) if for all  $m < \lfloor \frac{n}{2} \rfloor$  the  $m$ -body marginals are proportional to the identity and the spectra of all  $\lfloor \frac{n}{2} \rfloor$ -body marginals are equal.

Obviously, a state is MMS if and only if it  $\lfloor \frac{n}{2} \rfloor - 1$ -uniform and the spectra of all  $\lfloor \frac{n}{2} \rfloor$ -body marginals are equal.

### 6.5.3 Higher-dimensional systems

#### Bipartite systems

For the bipartite case the generalized Bell states  $|\psi_d\rangle = (1/\sqrt{d}) \sum_j |jj\rangle$  are maximally entangled with  $G(|\psi_d\rangle) = 1 - (1/d)$ . We find that for  $2 \leq d \leq 10$  the algorithm yields the corresponding state  $|\psi_d\rangle$  with high fidelity and that the number of iterations needed until convergence appears to be independent of  $d$ , see also Fig. 6.3.

#### Tripartite systems

In the three-qutrit case we obtain the total antisymmetric state  $|\text{AS}_3\rangle$  given by

$$|\text{AS}_3\rangle = \frac{1}{\sqrt{6}}(|012\rangle + |201\rangle + |120\rangle - |210\rangle - |102\rangle - |021\rangle). \quad (6.27)$$



In general, antisymmetric states  $|\text{AS}_n\rangle$  can be constructed for all  $n$ -partite  $n$ -level systems via

$$|\text{AS}_n\rangle = \frac{1}{n!} \sum_{j_1, \dots, j_n} \epsilon_{j_1, \dots, j_n} |j_1 \cdots j_n\rangle, \quad (6.28)$$

and their geometric measure can be computed analytically as  $(n-1)/n!$  [301]. In the particular case of  $n = 3$ , we obtain  $G = \frac{5}{6} \approx 0.8333$ . Note that  $|\Psi_3\rangle$  is an AME state. More generally, in the tripartite case a procedure is known to construct AME states for arbitrary  $d$  [302]. This leads to

$$\text{AME}(3, d) \sim \sum_{i, j=0}^{d-1} |i\rangle|j\rangle|i+j\rangle \quad (6.29)$$

where  $i+j$  is computed modulo  $d$ . Interestingly, the state  $\text{AME}(3, 3)$  that is computed according to Eq. (6.29) only yields a measure of  $2/3$ , thus smaller than the measure of  $|\text{AS}_3\rangle$ .

#### Four-partite systems

For the case of four ququads, that is, four four-level systems, the algorithm gives interesting insights into the AME problem. First, the  $\text{AME}(3, 4)$  state corresponding to Eq. (6.29) has a geometric measure of  $G(\text{AME}(3, 4)) = 0.75$ . However, our algorithm yields a state given by

$$|\psi_{3,4}\rangle = \frac{1}{2\sqrt{2}}(|022\rangle + |033\rangle + |120\rangle + |131\rangle + |212\rangle + |203\rangle + |310\rangle + |301\rangle) \quad (6.30)$$

with  $G(|\psi_{3,4}\rangle) = 7/8 = 0.875$ . After applying local unitaries, this state can be seen as arising from three Bell pairs distributed between three parties in a triangle-like configuration. In addition,  $|\psi_{3,4}\rangle$  is an AME state, i.e., all one-party marginals are maximally mixed. As the geometric measure of the states  $|\psi_{3,4}\rangle$  and  $\text{AME}(3, 4)$  differs, they belong to different SLOCC classes.

In the case of four qutrits the algorithm converges to a state with a geometric measure of  $0.888 \approx 8/9$ . This state can be identified to be the  $\text{AME}(4, 3)$  state given by [302]

$$\begin{aligned} \text{AME}(4, 3) = \frac{1}{3}(&|0000\rangle + |0112\rangle + |0221\rangle + |1011\rangle + |1120\rangle \\ &+ |1202\rangle + |2022\rangle + |2101\rangle + |2210\rangle) \end{aligned} \quad (6.31)$$

For four ququads the algorithm converges to the antisymmetric state  $|\text{AS}_4\rangle$  defined in Eq. (6.28), yielding a measure of  $23/24 \approx 0.9583$ . Interestingly, while being 1-uniform, this state is not AME. The so far only known  $\text{AME}(4, 4)$  state is a graph state [303, 304], see also Fig. 6.2, and yields a measure of  $15/16 \approx 0.9375$ . Finally, the recently found

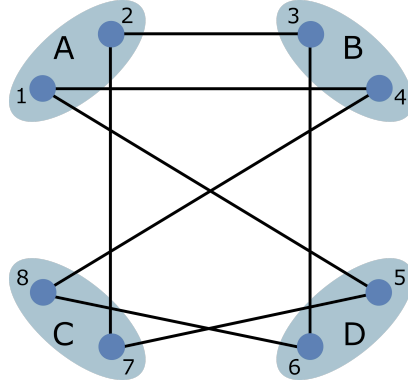


Figure 6.2: Graph of the known AME(4,4) state [303]. The figure is taken from Ref. [D].

AME(4,6) [258] is not maximally entangled with respect to the geometric measure. While  $G(\text{AME}(4,6)) \approx 0.9781$ , our algorithm converges to a state with a measure of  $G \approx 0.9806$ . Further, by inspecting the marginals of the found state one finds that they are very close to be maximally mixed. This gives strong evidence that in the case of four quhex there exists another SLOCC inequivalent AME state.

### Five-partite systems

Similar to the construction procedure of AME states for the tripartite case given by Eq. (6.29), there exists a general method to construct AME(5, $d$ ) states [302, 305]

$$\text{AME}(5,d) \sim \sum_{j,k,l=0}^{d-1} \omega^{jk} |j\rangle |k\rangle |j+k\rangle |l+k\rangle |l\rangle \quad (6.32)$$

where  $\omega = e^{\frac{2\pi i}{d}}$ . In the case of a three-dimensional system  $d = 3$ , the algorithm converges to the AME(5,3) state yielding a measure of approximately 0.96122. For the case of five ququads, this prescription yields the state AME(5,4) with  $G(\text{AME}(5,4)) = 31/32 = 0.96875$ . Here the algorithm converges to a state  $|\psi_{5,4}\rangle$  with a larger geometric measure, in particular  $G(|\psi_{5,4}\rangle) > 0.975$ . However, here we cannot identify a closed expression of the state. The numerical result suggests that the maximizer is again an AME state.

### The implementation and performance of the algorithm

Thanks to the update rule in Eq. (6.7), we can make use of advanced descent optimization algorithms in order to obtain faster convergence and higher robustness against local optima [117, 306]. We have implemented a descent algorithm with momentum as well as the Nesterov accelerated gradient (NAG) [307]. The idea behind the momentum version is to keep track of the direction of the updates. More precisely, the

update direction  $|\eta_n\rangle$  in the  $n$ -th iteration will be a running average of the previously encountered updates  $|\eta_1\rangle, \dots, |\eta_{n-1}\rangle$ . For the NAG method, the update vector is, contrary to Eq. (6.7), evaluated at a point estimated from previous accumulated updates, and not at  $|\varphi\rangle$ . In general it should be noted that the seesaw algorithm for computing the best rank one approximation is prone to local maxima. Therefore a randomization should be used, i.e., we run the iteration for many different initial states. The number of iterations as well as the number of initial states depends on the number of parties and the local dimension. The iteration typically converges fast, e.g., for three qubits 10 iterations are sufficient and for five ququads, 30 iterations. For small systems the number of initial states can be chosen small, e.g., for three qubits 10 different initial points make the largest overlap robust while for larger systems more initial states are necessary, e.g., for five ququads 100 points were taken. The step size  $\theta$  used in the update rule in Eq. (6.7) depends on the size of the system and on the variation of the measure of the iterates. For systems of small and moderate size, we initially choose  $\theta = 0.01$ . After a certain number of iterations (mostly around 400), the measure of the iterates is not increasing anymore, but fluctuates around a certain value where the amount of fluctuation depend on the step size. In this case, the step size is reduced according to  $\theta \mapsto \theta/2$  and one proceeds with the new step size. However, if  $\theta$  becomes small it is also useful to improve the precision in the computation of the best product state approximation.

## 6.6 The algorithm for typical states

We have seen that the presented algorithm can yield maximally entangled states with respect to the geometric measure. This is important as they can be seen as highly complex quantum states that on the other hand still offer a lot of structure. From this, two questions arise: First, how large is the gain of entanglement if one compares the maximizer to randomly chosen states and second, how does the performance of the algorithm behaves for such random starting points?

### 6.6.1 The entanglement of typical states

In order to sample from the set of pure states, i.e.,  $|\psi\rangle \in (\mathbb{C})^{\otimes n}$ , we can identify the state space with the unit sphere of  $\mathbb{C}^{d^n}$ . In the following we will denote the  $m$ -dimensional complex unit sphere by  $S^{m-1}$ . Further, we say that a complex-valued random variable  $X$  is standard normal distributed, denoted by  $X \sim \mathcal{N}(0,1)$ , if the real-valued random variables  $\Re(X)$  and  $\Im(X)$  are independent and standard normal distributed. In order to sample a point uniformly at random according to the Haar measure on  $S^n$ , see also Ref. [308] for a more detailed explanation, one might consider a sequence  $(X_1, \dots, X_n)$  of normal distributed independent random variables

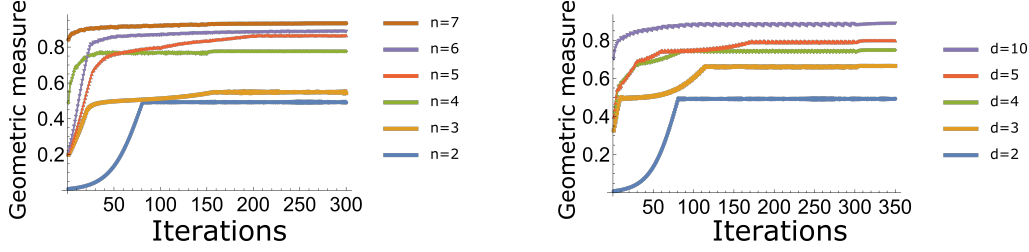


Figure 6.3: Performance of the algorithm for multi-qubit systems (left). Initializing with a random state, for each iteration the geometric measure  $G$  is computed. The step-size  $\theta$  depends on the size of the system and we have  $\theta_{n=2} = \theta_{n=3} = \theta_{n=4} = 0.01$  and  $\theta_{n=5} = \theta_{n=6} = 0.06$ . For the case  $n = 3$ , the slope starts to decrease when a measure of  $G \approx 0.5$  is reached, resulting from the fact that the GHZ state is an exceptional point of the function and yields  $G(|\text{GHZ}\rangle) = 0.5$ . A similar behavior can be observed in the case  $n = 4$  for the  $|M\rangle$  and  $|L\rangle$  state. Convergence of the algorithm for bipartite systems of different local dimension  $d \in \{2, 3, 4, 5, 10\}$  (right). For local dimension  $d \leq 5$  we have chosen the step-size as  $\theta = 0.01$ . For  $d > 5$  the step-size was chosen as  $\theta = 0.1$ . After 350 iterations, the iterates had a very high fidelity with the  $d$ -dimensional maximally entangled state. The figure is taken from Ref. [D].

$X_j \sim \mathcal{N}(0, 1)$ . Then, by the property of Gaussians, the vector  $(X_1, \dots, X_n) \in \mathbb{C}^n$  is a rotationally invariant  $n$ -dimensional Gaussian. Normalizing the vector  $(X_1, \dots, X_n)$  yields a uniform random point on  $\mathbb{S}^{n-1}$ .

We now numerically approximate the distribution of entanglement in multi-qubit systems with respect to the Haar measure. The distribution for different number of particles is presented in Fig. 6.5. To obtain the corresponding data, we have randomly sampled  $n$ -qubit states and computed the corresponding geometric measure. For the cases  $n = 3, 4$  we choose  $10^5$  states. To compute the geometric measure for each of those states, 20 iterations and 20 random starting points are sufficient in order to obtain a robust output. Since for the larger cases  $n = 5, 6, 7$  the dimension of the underlying space increases, we choose  $10^6$  states and calculate the geometric measure using 80 iterations and 100 random starting points, yielding a robust estimate for the geometric measure. Here we regard a computation as robust, if the variance of the outputs is smaller than  $10^{-7}$ . It should be noted that no analytical expression is known neither for the exact distribution of  $G$  with respect to the Haar measure, nor for its moments. However, using the geometric measure of the sampled quantum states we can obtain estimates for the moments of the distribution. In order to get optimal estimates for higher moments, i.e., with small error probabilities, one can

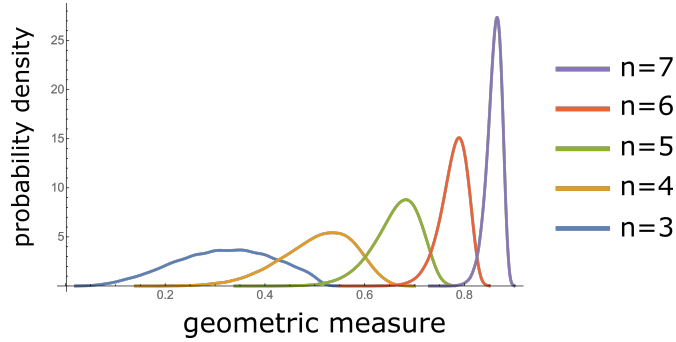


Figure 6.4: The probability distribution of the geometric measure for different multi-qubit systems. It can be easily seen that the maximum of the distribution is shifted to higher values when the number of parties increases. By having an approximation to the distribution, we can also calculate the first and second moments, see Tab. 6.2. The figure is taken from Ref. [D].

use the method of  $U$ -statistics [309]. For the mean value, Hoeffding's inequality yields that the estimated expectation value  $\mathbb{E}[G]$  is very close to the true value with very high success probability. The exact numerical values can be found in Tab. 6.2. The geometric measure of typical tensors was also recently numerically investigated in [310].

System	$\mathbb{E}[G]$	$\text{Var}[G]$
3 qubits	0.3089	0.0094
4 qubits	0.4950	0.0054
5 qubits	0.6534	0.0023
6 qubits	0.7731	0.0008
7 qubits	0.8570	0.0002

Table 6.2: The first two moments of the geometric measure  $G$  for small multi-qubit systems. The expected amount of entanglement increases with the number of parties and concentrates around its means, which is indicated by the decrease of the variance. The table is taken from Ref. [D].

More generally, there exist bounds which limit the measure of sets of states having a small amount of entanglement [273], see also Section 6.11. Particularly, for multi-qubit system composed of more than 11 constituents, one has

$$\mu_{Haar}[\{|\varphi\rangle \in (\mathbb{C}^2)^{\otimes n} : G(|\varphi\rangle) < 1 - \frac{8n^2}{2^n}\}] \leq e^{-n^2}. \quad (6.33)$$

Therefore, if  $n$  is large enough, almost all states will have a geometric measure close to 1.

### 6.6.2 Performance for random starting points

To analyze the sensitivity to local maxima of the algorithm, we consider the behavior of the iterates for different starting points for multi-qubit systems. Although in the concrete implementation we have used 100 different starting points and selected among those the state with the largest measure, here we will illustrate different trajectories of the iterates for five different Haar random starting points using the same step size  $\epsilon = 0.05$ . Within the computation of the geometric measure in each step, we used 50 iterations and 50 random product states. The exact behaviors can be found in Fig. 6.5. Here it is important to note that after the iterations each of the states has approximately the same geometric measure. This indicates that at least for the case of a small number of parties, the algorithm is stable with respect to local optima.

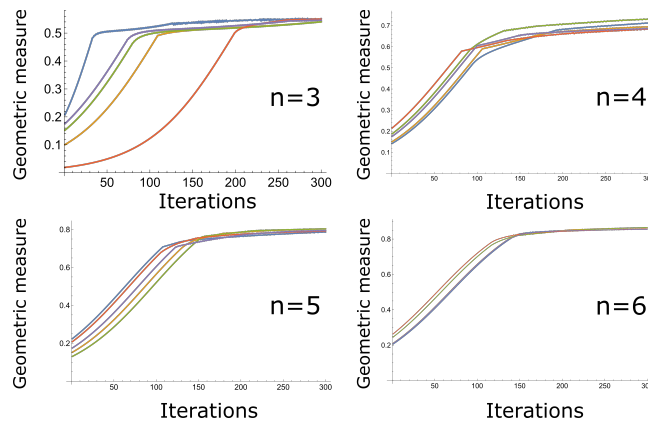


Figure 6.5: Trajectory of the geometric measure of the iterates of the algorithm for different starting points and different numbers of qubits. While for the case  $n = 3$  the different trajectories differ, this deviation becomes smaller if the number of parties increase. However, it should be noted that after 300 iterations each state displays roughly the same amount of entanglement, showing that at least for the case of  $n = 3, 4, 5, 6$  the algorithm is robust w.r.t. to local optima. It can be seen that for increasing  $n$  also the geometric measure of the initial state increases, as the expected geometric measure increases. Further, the different trajectories are getting more narrow with increasing  $n$ , reflecting the concentration property of the geometric measure. The figure is taken from Ref. [D].

## 6.7 The algorithm for entangled subspaces

One can extend the algorithm such that it also applies to subspaces. More precisely, we want to construct an orthonormal basis for a subspace  $V \subset (\mathbb{C}^d)^{\otimes n}$  such that the least entangled state contained in  $V$  is as entangled as possible in comparison with all other subspaces, that is,

$$V_{\max} := \operatorname{argmax}_V \min_{|\psi\rangle \in V} G(|\psi\rangle). \quad (6.34)$$

Note that the notion introduced in Eq. (6.34) differs from the concept of genuine entangled subspaces (GES) [311], where all states within the subspace have to be genuine entangled. However, our algorithm can readily be modified to this situation and thus search for GES with maximal genuine multipartite entanglement.

To keep the notation simple, we will explain the idea of the algorithm for the case of a two-dimensional subspace of three qubits. First, we choose a two-dimensional subspace randomly, which can be described by a projection operator  $P$  that is of the form

$$P = |v\rangle\langle v| + |w\rangle\langle w|, \quad (6.35)$$

where  $\langle v|w\rangle = 0$  and  $|v\rangle, |w\rangle \in (\mathbb{C}^2)^{\otimes 3}$ . Next, we compute the best rank-one approximation to  $P$  which is given by

$$|\pi\rangle := \operatorname{argmax}_{|abc\rangle} \operatorname{Tr}[P|abc\rangle\langle abc|]. \quad (6.36)$$

This can be done with the iterative method already described in the context of finding upper approximations of the geometric measure. More generally, the optimization yields the best product state approximation to the state in the range of  $P$  which is least entangled. In particular, this implies that if  $\operatorname{im}(P)$  contains a product state, the assigned geometric measure to  $P$  will be zero. For a given step size  $\theta > 0$  the first part of the update rule is then given by

$$P \mapsto \tilde{P} := P - \theta|\pi\rangle\langle\pi|. \quad (6.37)$$

However, the operator  $\tilde{P}$  in Eq. (6.37) is in general not a projector and thus does not correspond to a subspace. Therefore we compute the eigenvectors corresponding to the two largest eigenvalues of the operator  $\tilde{P}$  that we will call  $|v_1\rangle$  and  $|v_2\rangle$ . The new projection is then given by  $|v_1\rangle\langle v_1| + |v_2\rangle\langle v_2|$ . Clearly, this algorithm reduces to the one for the maximization of the geometric measure if we choose the rank of the projector to be one.<sup>2</sup> It is known that the maximal dimension of a subspace  $V$  of an  $n$ -partite

<sup>2</sup>This is particularly clear from the viewpoint of projective geometry. Here one identifies a one-dimensional subspaces, a so-called ray, with points (vectors of unit length) on the unit sphere.

qudit system which contains no product state is given by

$$\dim(V) \leq d^n - dn + n - 1. \quad (6.38)$$

Consequently, for a two qubit system, the maximal entangled subspace is of dimension one and spanned by a Bell state, see Eq. (1.6). Indeed, if we apply in this case our algorithm to larger subspaces, the measure we assign to these subspaces stay zero. For three qubits, the algorithm converges to the W state as one basis vector and to the state

$$|V\rangle = \frac{1}{\sqrt{3}}(|001\rangle + e^{\frac{2\pi}{3}i}|110\rangle + e^{\frac{4\pi}{3}i}|101\rangle) \quad (6.39)$$

as the other. The subspace spanned by  $|W\rangle$  and  $|V\rangle$  then has a remarkable property: All states within this subspace are maximally entangled, yielding the same geometric measure as the W state. This has potential applications in information processing. Indeed, in this subspace qubit states may be encoded and then any set of states is difficult to discriminate by local means [76].

Concerning higher dimensions, we also compute the maximally entangled subspace of dimension two for two qutrits. Here we obtain an embedded Bell state  $|\chi_1\rangle = (1/\sqrt{2})(|01\rangle - |10\rangle)$  and the state

$$|\chi_2\rangle = \frac{1}{\sqrt{14}}(|20\rangle + |02\rangle + \sqrt{6}(|21\rangle + |12\rangle)), \quad (6.40)$$

which can also be seen as the superposition of two Bell pairs. It turns out that the least entangled state within this subspace has a geometric measure of  $1/2$ . Again applications can be envisaged, as any state in this subspace has at least one ebit of entanglement.

## 6.8 Application to states with a fixed stabilizer rank

In the following, we explain how the algorithm can be used to find states which can not be approximated well by states of a fixed stabilizer rank. For this purpose, we will first introduce the concept of stabilizer rank and explain why it is an important tool for the classical simulation of quantum computations. Afterwards, we introduce the modified algorithm and present numerical results for a slight variant of the stabilizer rank which is numerically much more tractable.

### 6.8.1 The stabilizer rank

The stabilizer rank  $\chi$  of a multi-qubit system was first introduced in the context of simulation of Pauli-based quantum computation (PBC) [312]. While the main aim of Ref. [312] is to quantify how many classical resources (time) are needed in order to simulate a PBC on  $(n+k)$  qubits by a PBC using only  $n$  qubits, their introduced



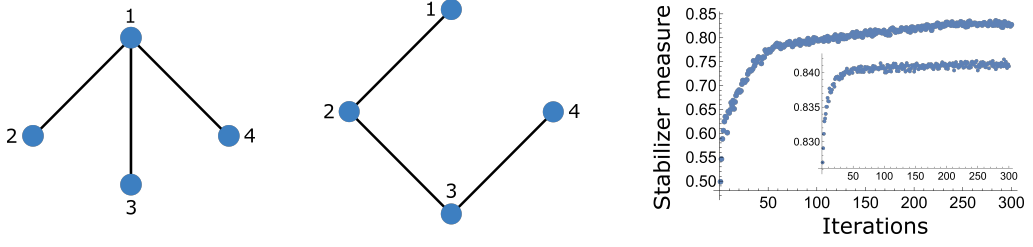


Figure 6.6: The graph corresponding to the four particle GHZ state (left) and the graph of the four particle cluster state (middle). Performance of the algorithm for the case of four particles (right). The optimization is implemented for stabilizer rank  $\chi = 2$  in the restricted model of 32 stabilizer states, corresponding to GHZ and cluster graph state basis. We start with a random initial state and a step size of  $\epsilon = 0.1$  (outset). Then, we change step size to  $\epsilon = 0.01$  (inset). The figure is taken from Ref. [D].

method turned out to also improve the cost of a brute force full classical simulation. It can be shown that a brute force classical simulation of a  $n$ -qubit PBC, which includes state operations in the computational basis, has a simulation cost (time) of  $\mathcal{O}(n2^n)$ , where  $n$  denotes the number of qubits involved in the computation. However, one can devise an algorithm which classically simulates a  $n$  qubit PBC in only  $2^{\alpha n} \text{poly}(n)$ , where  $\alpha \approx 0.94$  [312]. The idea of the proof is to expand certain  $n$ -qubit states as a linear combination of stabilizer states, as  $n$ -qubit stabilizer states only require  $\mathcal{O}(n^2)$  classical bits to store [313]. The minimal number of stabilizer states needed is then called the stabilizer rank. More precisely, for an arbitrary  $n$ -qubit state  $|\psi\rangle$ , its stabilizer rank  $\chi$  is defined as the minimum number  $r \geq 1$  of stabilizer states needed to represent this state, that is,

$$|\varphi\rangle = \sum_{k=1}^r \alpha_k |s_k\rangle, \quad (6.41)$$

where  $\alpha_k \in \mathbb{C}$  and  $|s_k\rangle$  is a stabilizer state, i.e., one has  $|s_k\rangle = U_k |0\rangle^{\otimes n}$  for some  $n$ -qubit Clifford operation  $U_k$ . It is therefore highly desirable, from the practical viewpoint of classical simulation as well as from a theoretical perspective, to know which states can not be approximated well by stabilizer rank  $\chi$  states. For this, we define the stabilizer measure  $\mathcal{S}_k : (\mathbb{C}^2)^{\otimes n} \rightarrow \mathbb{R}$  with

$$|\psi\rangle \mapsto \mathcal{S}_k(|\psi\rangle) := 1 - \sup\{|\langle\varphi|\omega\rangle|^2 : \chi(|\omega\rangle) = k\}. \quad (6.42)$$

### 6.8.2 The modified algorithm and results

For a fixed  $k$ , the stabilizer measure in Eq. (6.42) directly allows for an application of the algorithm. First, generate the set of all stabilizer states and form all subsets of

size  $k$ . Second, draw an initial state  $|\psi\rangle$  at random. Then, compute the closest stabilizer rank  $k$  state, i.e., for given set of stabilizers  $|s_1\rangle, \dots, |s_k\rangle$ , one has to maximize the overlap of the given state  $|\psi\rangle$  with  $|\omega\rangle = \sum_{j=1}^k \alpha_j |s_j\rangle$ . This is a constraint optimization over  $k$  complex parameters  $\alpha_1, \dots, \alpha_k$  subject to the normalization of the stabilizer rank  $k$  state. If the best approximation  $|\omega\rangle$  is found, one updates  $|\psi\rangle \mapsto (1/\mathcal{N})(|\psi\rangle + \theta|\eta\rangle)$ , where  $|\eta\rangle = (\mathbb{1} - |\omega\rangle\langle\omega|)|\psi\rangle$ . This procedure is then iterated.

We have implemented the algorithm for the restricted toy model, see also Fig. 6.6 (right). We initialized with a random state and iterated 300 times with a step size of  $\theta = 0.1$  until  $\mathcal{S}_k^R$  does not increase anymore. Then we change the step size to  $\theta = 0.01$  and iterate again 300 times. The algorithm converges to a state  $|\psi_{\max}\rangle$  yielding  $\mathcal{S}_2^R(|\psi_{\max}\rangle) \approx 0.8415$ . A comparison of  $\mathcal{S}_2^R$  for different four-qubit states can be found in Tab. 6.3.

However, the number of  $n$  qubit stabilizers  $S(n)$  can be computed explicitly [178] and is given by  $S(n) = 2^n \prod_{j=1}^n (2^j + 1)$ , which scales exponentially in the number of qubits. For instance, one has  $S(3) = 1080$ ,  $S(4) = 36720$  and  $S(5) = 2423520$ . Further, in order to evaluate the overlap with stabilizer rank  $k$  states, one has to inspect all  $\binom{S(n)}{k}$  different combination and optimize over the  $k^2 - 1$  real parameter which becomes infeasible even for small  $n, k$ . We therefore consider a non-trivial toy model, where we restrict to a specific subset of stabilizer states. In particular, we consider the set consisting of the graph state bases corresponding to four qubit GHZ-graph and the four-qubit cluster-graph, see Fig. 6.6, as for four qubits there are only two inequivalent graphs. For each graph, the set of all eigenstates with a different eigenvalue signature is a basis for  $(\mathbb{C}^2)^{\otimes 4}$  and thus our toy model comprises 32 stabilizer states. Note that for one graph, those states form an orthonormal basis, while the set of eigenstates of two graphs displays a highly nontrivial geometric structure. If the optimization in Eq. (6.42) only runs over this restricted set, we write  $\mathcal{S}_k^R$  for the corresponding stabilizer measure. Note that contrary to the geometric measures  $G_m$ , two LU equivalent states do not need to have the same measure  $\mathcal{S}_k$ , as we try to maximize the overlap w.r.t. a set that is not invariant under LU operations. Of course, the measure  $\mathcal{S}_k$  of two states which are Clifford equivalent is the same.

State $ \psi\rangle$	$ M\rangle$	$ L\rangle$	$ H\rangle^{\otimes 4}$	$ R\rangle^{\otimes 4}$	$ \psi_{\max}\rangle$
Measure $\mathcal{S}_2^R$	$1/3 \approx 0.333$	$1/2$	$\approx 0.4457$	$0.4604$	$0.8415$

Table 6.3: Results of the optimization with respect to the restricted stabilizer measure  $\mathcal{S}_2^R$  for rank two. The states  $|H\rangle$  and  $|R\rangle$  are magic states for which it is conjectured to have the smallest possible stabilizer rank among all non stabilizer single-qubit states [312]. The state  $|\psi_{\max}\rangle$  refers to the state found by the algorithm for which  $\mathcal{S}_2^R$  appears to be maximal. The table is taken from Ref. [D].

## 6.9 Application to states with a fixed Schmidt rank

In the following we explain how the algorithm can be used to find states which can not be well approximated by states with a fixed Schmidt-rank or, for the particular case of matrix product states, with a fixed bond dimension.

### 6.9.1 Schmidt rank, tensor rank and border rank

Apart from the geometric measure, there are also other quantifier of entanglement offering different operational interpretations [72,73]. One such possibility is to quantify the amount of entanglement in a system by its *dimensionality*. High-dimensional quantum entanglement can be exploited to tolerate larger amounts of noise in quantum communication protocols, making them a valuable resource [314].

Let  $|\varphi\rangle$  be an  $n$ -particle quantum state where each constituent is a  $d$ -level quantum system. Any such  $|\varphi\rangle$  can be written as

$$|\varphi\rangle = \sum_{k=1}^R \mu_k |\pi_k^{(1)}\rangle \otimes \cdots \otimes |\pi_k^{(n)}\rangle, \quad (6.43)$$

where  $|\pi_k^{(j)}\rangle \in \mathbb{C}^d$  and  $\mu_k \in \mathbb{C}$ . Note that the different summands  $|\pi_k^{(1)}\rangle \otimes \cdots \otimes |\pi_k^{(n)}\rangle$  do not have to be orthogonal as in the case of the Schmidt decomposition of bipartite pure states. The minimal number of terms  $R$  needed in order to decompose  $|\varphi\rangle$  into the form in Eq. (6.43), i.e., into a sum of rank-1 tensors, is called the tensor rank of  $|\varphi\rangle$  and is denoted by  $\text{rk}(|\varphi\rangle)$ . The decomposition of  $|\varphi\rangle$  in Eq. (6.43) is also called rank decomposition or minimal CP decomposition [315].

The Schmidt measure  $P$  of a state  $|\psi\rangle$  is defined as  $P(|\varphi\rangle) := \log_2 \text{rk}(|\varphi\rangle)$  [316]. Clearly, in the case of a bipartite system, the minimal number of rank-1 terms needed is given by the Schmidt rank of the state. Once  $P$  is defined for pure states, one can extend it to the full state space of mixed quantum states in a natural way, i.e., via a convex roof construction [317]. The Schmidt measure gives also rise to an approximation version similar to the geometric measure. For this, define the generalized geometric measure  $G_k : (\mathbb{C}^d)^{\otimes n} \rightarrow \mathbb{R}$  with

$$|\varphi\rangle \mapsto G_k(|\varphi\rangle) := 1 - \sup\{|\langle\varphi|\omega\rangle|^2 : \text{rk}(|\omega\rangle) = k\}, \quad (6.44)$$

which measures how well a given quantum state  $|\varphi\rangle$  can be approximated by a state of tensor rank  $k$ .

Therefore, the generalized geometric measure  $G_k$  gives more detailed information about the entanglement structure present in the state  $|\varphi\rangle$ , and reproduces the geometric measure  $G$  for the special case  $k = 1$ . Clearly, the generalized geometric measure is a decreasing function in the rank  $k$ , that is, for given  $|\varphi\rangle$  we have  $G(|\varphi\rangle) = G_1(|\varphi\rangle) \leq G(|\varphi\rangle) \leq \cdots \leq G_\mu(|\varphi\rangle)$ , where  $\mu$  is the maximal rank that a tensor in  $(\mathbb{C}^d)^{\otimes n}$  can

have. For instance, for the bipartite case  $\mathbb{C}^{d_A} \otimes \mathbb{C}^{d_B}$ , the Schmidt decomposition directly leads to  $\mu = \min\{d_A, d_B\}$ . For the case of three qubits, it is known [318] that  $\mu = 3$ , but for the general case only not tight upper bounds exist. Obviously, if we find  $G_k(|\varphi\rangle) = 0$ , then  $G_{\tilde{k}}(|\varphi\rangle) = 0$  for all  $\tilde{k} \geq k$ . The algorithm for maximizing the measure  $G_k$  then proceeds as follows. First, draw the initial state  $|\varphi\rangle$  at random and compute its best rank- $k$  approximation, that is,

$$|\omega\rangle := \operatorname{argmin} \{|\omega\rangle : 1 - |\langle\varphi|\omega\rangle|^2 : \operatorname{rk}(|\omega\rangle) = k\}. \quad (6.45)$$

Second, use the update rule in Eq. (6.7) and the direction  $|\omega\rangle$  to obtain iterates that have a larger generalized geometric measure  $G_k$ .

### 6.9.2 Computing rank- $k$ approximations

There are different methods to compute low rank approximations to a given tensor, most prominently the CP-ALS algorithm, which updates each mode individually by a least-squares optimization while keeping all other modes fixed. For a concise summary of those methods, see Ref. [315]. However, to keep notation simple, we will generalize the algorithm mentioned in Section 6.2 for higher rank approximations, which we will illustrate for the three-particle case. For a given state  $|\psi\rangle \in (\mathbb{C}^d)^{\otimes 3}$ , consider the Schmidt decomposition  $|\psi\rangle_{A|BC} = \sum_{k=1}^D \lambda_k |\psi_k^A\rangle |\psi_k^{BC}\rangle$ . For the overlap with a rank  $k$  state given by Eq. (6.43) one finds

$$\langle\psi|_{A|BC}\pi\rangle = \sum_{j=1}^D \sum_{k=1}^r \lambda_j \alpha_k \langle\psi_j^A|\pi_k^{(A)}\rangle \langle\psi_j^{AB}|\pi_k^{(B)}\rangle |\pi_k^{(C)}\rangle = \sum_{k=1}^r \langle\tilde{\gamma}_k|\pi_k^{(A)}\rangle, \quad (6.46)$$

where we have defined  $\langle\tilde{\gamma}_k| = \sum_{j=1}^D \lambda_j \alpha_k \langle\psi_j^A|\pi_k^{(A)}\rangle \langle\psi_j^{AB}|\pi_k^{(B)}\rangle |\pi_k^{(C)}\rangle$ . Note that  $|\tilde{\gamma}_k\rangle$  is not normalized and we denote by  $|\gamma_k\rangle$  its normalized version. The norm is given by  $\alpha'_k = \sqrt{\langle\tilde{\gamma}_k|\tilde{\gamma}_k\rangle}$ . We now update the states of system  $A$  as well as the coefficients  $\alpha_k$  that appear in the decomposition as  $|\pi_k^{(A)}\rangle \mapsto |\gamma_k\rangle$  and  $\alpha_k \mapsto \alpha'_k$ . This procedure is iterated with respect to each mode and repeated many times. In order to reduce the sensitivity for local maxima, different randomly chosen starting points  $|\pi\rangle$  should be used.

To test the reliability of the algorithm, we apply it to quantum states for which the tensor rank is already known. For instance, one has  $\operatorname{rk}(|\text{GHZ}\rangle) = 2$ ,  $\operatorname{rk}(|W\rangle) = 3$  and  $\operatorname{rk}(|W\rangle^{\otimes 2}) = 7$  [319]. However, one should notice that for tensor approximations of higher rank ( $k \geq 2$ ) the problem of degeneracy exists, reflecting that the set of tensors of rank  $k$  is open which is in contrast to the matrix case. In particular, for the bipartite (matrix) case the Eckart-Young theorem ensures that the best rank  $k$  approximation is simply given by the eigenvectors corresponding to the  $k$  largest singular values. Further it is known that deflation techniques, i.e., compute best rank-1 approximation, subtract it and then iterate, are not fruitful for tensors.

A tensor is degenerate if it may be approximated *arbitrarily well* by tensors of lower rank. Also well known [289, 315], we will discuss this phenomenon for the case of the W state. The W state  $|W\rangle = (1/\sqrt{3})(|001\rangle + |010\rangle + |100\rangle)$  has tensor rank 3, but can be approximated arbitrarily well by tensors of tensor rank-2. Consider the family of rank 2 states

$$|\pi(\alpha)\rangle = \frac{\alpha}{N_\alpha}(|0\rangle + \alpha^{-1}|1\rangle)^{\otimes 3} - \frac{\alpha}{N_\alpha}|0\rangle^{\otimes 3} \text{ with } N_\alpha = \sqrt{3 + 3\alpha^{-2} + \alpha^{-4}}, \quad (6.47)$$

where  $N_\alpha$  assures that  $\langle \pi(\alpha) | \pi(\alpha) \rangle = 1$  for all  $\alpha > 0$ . It directly follows that  $\langle W | \pi(\alpha) \rangle = \sqrt{3}/\sqrt{3 + 3\alpha^{-2} + \alpha^{-4}}$  which tends to 1 for  $x \rightarrow \infty$ . To emphasize that problem, one also says that the W state has border rank 2. Therefore, it is not surprising that our approximation algorithm yields that  $G_2(|W\rangle)$  is numerically 0. As it turns out, the generalized geometric measure  $G_k$  has also the advantage that it distinguishes more clearly between different forms of entanglement, e.g., the L state and the M state. In particular, one finds that  $G_3(|L\rangle) = 0$  while  $G_3(|M\rangle) \approx 0.2626$ .

### 6.9.3 The modified algorithm and results

The smallest case where a maximization of  $G_2$  could be considered is in principle  $(\mathbb{C}^2)^{\otimes 3}$ . Here it is known that there are only two classes of genuine multipartite entanglement with respect to SLOCC, namely the GHZ and the W class [67]. Further, the tensor rank is monotonically decreasing under SLOCC operations [316]. In particular, one can show that the set of all tensors of rank three is the closure of the SLOCC orbit of the W state. However, the W state has border rank two and thus all states within his orbit. This implies the nonexistence of border rank three tensors in  $(\mathbb{C}^2)^{\otimes 3}$ . As a consequence, the algorithm cannot be applied to this case, as in each step the given state can be approximated arbitrarily well by states of rank 2. Therefore the first nontrivial case is a system of four qubits with respect to rank two and three. Here, states of border rank three and four exist and are subsets of non-vanishing measure. We implement the algorithm with 20 different random initial points for each case. For the case of rank 2, we identify the state to be in 7 of the 20 runs the M state and in the other cases the four qubit cluster state. Further, the numerical computation of  $G_2$  for  $|C_4\rangle$  is stable with respect to a randomization of the starting point and one finds  $G_2(|C_4\rangle) = 1/2$ .

In this case, the rank two approximation found by the algorithm yielding the value  $G_2(|C_4\rangle) = 1/2$  can proven to be optimal. The cluster state is a graph state and we can consider the bipartition 23|14 (see also Fig. 6.6). This effectively yields the four-dimensional maximally entangled state  $|00\rangle + |11\rangle + |22\rangle + |33\rangle$ . Because all singular values coincide, we can choose the best rank two approximation as  $|00\rangle + |11\rangle$  which yields a squared overlap of 1/2. As  $G_m(|\psi\rangle_{1234}) \geq G_m(|\psi\rangle_{12|34})$ , this proves the optimality. However, as the M state is not a graph state, the situation is different. For the computation we choose  $10^6$  random starting points and make 300 iteration for each.

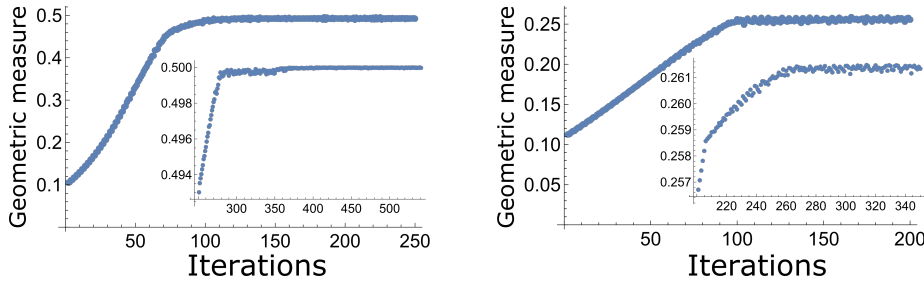


Figure 6.7: Performance of the algorithm for the case of four qubits and different generalized geometric measures. The left figure shows the maximization of the distance to the set of tensor rank 2 states, that is, the measure  $G_2$ . We start with a random state and choose a step size of  $\epsilon = 0.05$  (main). We compute the rank 2 best approximation with 100 iterations and 100 random initial points. After 150 iterations, there is no increase in the measure and the values start to fluctuate. We then change the step size to  $\epsilon = 0.001$  and compute the best rank 2 approximation with 200 iterations and 200 random initial points (inset). The right figure shows the maximization of the distance to the set of tensor rank 3 states, that is, the measure  $G_3$ . Similar, we start with a random state and choose a step size of  $\epsilon = 0.05$  (main) and compute the best rank 3 approximation with 100 iterations and 100 random initial points. Then, we change the step size to  $\epsilon = 0.001$  and compute the best rank 3 approximation with 200 iterations and 200 random initial points (inset). The figure is taken from Ref. [D].

The total computation was running for 27 hours. The optimal approximation found by the algorithm yields  $G_2(|M\rangle) = 5.000003(9)$ . This suggests that  $G_2(|M\rangle) = 1/2$  which would imply that the maximizer of  $G_2$  is not unique anymore in contrast to  $G_1$ . For the case of rank 3 approximations, the algorithm converges for all 20 starting points to the  $M$  state. With a similar computation as for  $G_2$  we find that  $G_3(|M\rangle) = 0.2626050$ . However, the algorithm can not be applied to maximize  $G_k$  for  $\geq 4$ . This is due to the fact that randomly drawn tensors in  $(\mathbb{C}^2)^{\otimes 4}$  can be approximated well by rank 4 tensors.

### Application to matrix product states

In general, it is an open question which forms of nontrivial quantum dynamics can be simulated efficiently by classical means. However, it is a celebrated result that if the computation only involves pure states containing a restricted amount of entanglement, such an efficient simulation is possible [320]. The power of those simulation algorithms rely on the fact that if the entanglement present in the multi-qudit state  $|\psi\rangle$  is bounded,

a low-dimensional sparse representation of  $|\psi\rangle$  can be derived. This is the so-called matrix product state (MPS) representation. A MPS of bond dimension  $\chi$  can be written as

$$|\psi\rangle = \sum_{i_1, \dots, i_n=1}^d \text{Tr} \left[ A^{[1]i_1} \cdot \dots \cdot A^{[n]i_n} \right] |i_1 \cdots i_n\rangle, \quad (6.48)$$

where  $A^{[l]i_l} \in \text{Mat}_\chi(\mathbb{C})$  for  $l = 1, \dots, n$ . Any state  $|\varphi\rangle$  can be represented as a MPS if the bond dimension is sufficiently large [320]. For a fixed number of parties  $n$  with local dimension  $d$  we write  $\text{MPS}(\chi)$  for the set of all MPS with bond dimension  $\chi$ . Notice that  $\text{MPS}(\chi)$  is a manifold in state space, which coincides with the set of product states for  $\chi = 1$  and has a nested structure, i.e.,  $\text{MPS}(\chi) \subset \text{MPS}(\chi + 1)$ . In particular, this shows that computing the best MPS approximation of given bond dimension  $\chi$  to a given state  $|\varphi\rangle$  is a hard problem, which includes for  $\chi = 1$  a NP-hard problem [279, 289].

Recently, a variant of the geometric measure was introduced, where the distance is measured with respect to the set of matrix product states of a given bond dimension  $\chi$  [321]. How well a generic multi-particle quantum state can be approximated by matrix product states of bond dimension  $\chi$  is then quantified by  $\mathcal{E}_k : (\mathbb{C}^d)^{\otimes n} \rightarrow \mathbb{R}$  with

$$|\psi\rangle \mapsto \mathcal{E}_k(|\psi\rangle) := 1 - \sup\{|\langle\psi|\omega\rangle|^2 : |\omega\rangle \in \text{MPS}(\chi)\}. \quad (6.49)$$

As already pointed out in Ref. [320], the tensor rank is not a continuous function. Consequently, there exist states  $|\psi\rangle$  that need a high bond dimension to be represented exactly, but very good approximations w.r.t. some norm exist where the optimizer has a significantly lower bond dimension. This property could then be used to obtain efficient simulations with small errors, even for *seemingly* high entangled states. Therefore, finding states which are difficult to approximate by states with fixed bond dimension is important as they do not allow for an efficient classical approximate simulation and can thus be regarded as genuine quantum resources.

Further, upper bounds on the distance between an arbitrary  $n$ -particle state  $|\psi\rangle$  and the set of  $\text{MPS}(\chi)$  can be derived [322]. Indeed, one can show that for any  $|\omega\rangle \in \text{MPS}(\chi)$  one has

$$\| |\psi\rangle - |\omega\rangle \|^2 \leq 2 \sum_{k=1}^{n-1} \epsilon_k(\chi), \quad (6.50)$$

with  $\epsilon_k(\chi) := \sum_{j=k+1}^{d_k} \lambda_j^{[k]}$ , where  $\lambda_j^{[k]}$  are the Schmidt coefficients of the bipartition  $[1 \dots k] | [k+1 \dots n]$  and  $d_k$  is the local dimension of the smaller of the two subsystems for that bipartition. Here our algorithm can find states for which the bound in Eq. (6.50) will be maximal and thus gives information about its tightness.

## 6.10 Application to independent triangle preparable states

A class of quantum states with importance for quantum information processing is the set of states which can be prepared in a quantum network [323,324]. For simplicity, in the following we will restrict to the triangle network and assume that the independent sources distribute pairs of qubit systems. Note that our algorithm can also be applied to more advanced network topologies as well as to higher local dimensions. Following Ref. [325], we call this network the independent triangle network, abbreviated ITN. From the structure of the network, see also Fig. 6.8 (middle), we obtain that a state  $|\psi\rangle$  can be prepared in the ITN if and only if there exist unitaries  $U_A, U_B$  and  $U_C$  and bipartite states  $|a\rangle, |b\rangle$  and  $|c\rangle$  such that

$$|\psi\rangle = U_A \otimes U_B \otimes U_C |abc\rangle. \quad (6.51)$$

Here it is important to note that the order of the subsystems is different for the unitaries and the states. For instance,  $U_A$  acts on the joint system  $A_1A_2$  while  $|a\rangle$  defines the joint state between  $B_2C_1$ . We denote the set of all states that can be prepared by means of two-qubit sources by  $\Delta_I$ . However, it turns out that  $\Delta_I$  admits a highly non-trivial structure [325]. Indeed,  $\Delta_I$  is not convex, certain separable (product) states are not contained and it is a subset of measure zero within the entire set of quantum states. This renders its analysis and characterisation difficult and not much is known about the structure of this set. As triangle states constitute an important resource, it is an interesting question how well a given quantum state can be approximated by states from  $\Delta_I$ . In order to quantify this property, we define the triangle measure  $\mathcal{T} : (\mathbb{C}^d)^{\otimes 6} \rightarrow \mathbb{R}$  as

$$|\psi\rangle \mapsto \mathcal{T}(|\psi\rangle) := 1 - \sup \{ |\langle \psi | \omega \rangle|^2 : |\omega\rangle \in \Delta_I \}. \quad (6.52)$$

Clearly, for any ITN preparable state we have  $\mathcal{T} = 0$ .

### 6.10.1 Network state approximations

Similar to the problem of finding the best rank-1 or best rank- $k$  approximation to a given state in the computation of the generalized geometric measures  $G_k$ , one can also devise a seesaw type algorithm for computing the ITN state which maximizes the overlap [325]. We will shortly recapitulate this algorithm. For simplicity, let us assume that the given target state  $|\psi\rangle$  is composed out of six qubits. We initialize the algorithm with a unitary for each party  $U_A, U_B, U_C$  and a state for each source  $|a\rangle, |b\rangle, |c\rangle$ . Now, if we keep the unitaries fixed, finding a better choice for the source states can be seen as finding the best rank one approximation with respect to a  $4 \times 4 \times 4$  system. After updating the source states, we have to optimize for the unitaries. To compute



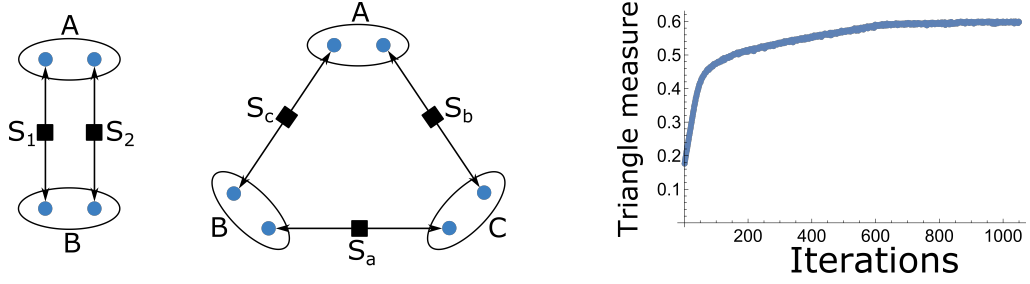


Figure 6.8: The genuine multilevel scenario (left). Each of the independent sources  $S_1$  and  $S_2$  generates a bipartite entangled system and sends one part to A and the other to B. Consequently, each of the parties receives two physical systems on which a joint unitary transformation can be performed. The set of all states that can be obtained in this scenario is abbreviated by  $\Delta_{GM}$ . The independent triangle scenario (middle). Three statistically independent sources  $S_a, S_b, S_c$ , distributing physical systems to the parties A,B,C. Upon receiving the two independent systems, each party can perform a joined unitary on the hold particles. The set of all states that can be obtained in this scenario is abbreviated by  $\Delta_I$ . The performance of the algorithm for the maximization of the triangle measure  $\mathcal{T}$  (right). The figure is taken from Ref. [D].

the optimal choice for  $U_A$ , define the new state  $|\tilde{\psi}\rangle = \mathbb{1}_A \otimes U_B \otimes U_C |\psi\rangle$ . Then one has [325]

$$\max_{U_A} |\text{Tr}[U_A \otimes \mathbb{1}_B \otimes \mathbb{1}_C |\tilde{\psi}\rangle \langle abc|]| = \max_{U_A} |\text{Tr}_A[U_A \rho_A]|, \quad (6.53)$$

with  $\rho_A = \text{Tr}_{BC}[|\tilde{\psi}\rangle \langle abc|]$ . From the singular value decomposition of  $\rho_A = UDV^\dagger$  one can then derive the optimal form of  $U_A$  as  $U_A = VU^\dagger$ . Alternating these two optimizations multiple times gives a good approximation to the optimal state. However, it should be noted that similar to the rank one approximation routine, this algorithm is prone to local maxima. In order to stabilize the optimal solution one should use multiple random initial choices.

### 6.10.2 The modified algorithm and results

In order to maximize the measure  $\mathcal{T}$  in Eq. (6.52) we initially choose a random starting point  $|\psi\rangle$ . We use the seesaw algorithm to compute the best network state approximation, called  $|\omega\rangle$ . Then we update the state according to  $|\psi\rangle \mapsto (1/\mathcal{N})(|\psi\rangle + \theta|\eta\rangle)$  with  $|\eta\rangle = (\mathbb{1} - |\omega\rangle \langle \omega|)|\psi\rangle$  and  $\mathcal{N} \geq 1$  a normalization. This procedure is iterated.

In the simplified case of only two parties and two independent sources, the problem of triangle preparability reduces to the multilevel entanglement problem [326], see

Fig. 6.8 (left). In that case, the states which have the smallest overlap among the set of all preparable states are known and can be analytically derived. In the case of four qubits, the state

$$|\xi\rangle = \sqrt{\frac{3}{4}}|00\rangle + \frac{1}{2\sqrt{3}}(|11\rangle + |22\rangle + |33\rangle) \quad (6.54)$$

has the largest distance to the set of decomposable states. Indeed, this state is found by the algorithm with very high fidelity.

For the ITN scenario, we choose 20 random starting points and a step size of  $\theta = 0.1$ . Further, to compute the best network state approximation, i.e, the optimal unitaries and source states, we choose 100 iterations in the seesaw routine and 500 different initial points. This maximization procedure is iterated 200 times. Then, the step size is changed to  $\theta = 0.01, 0.001, 0.0001$  while we also increase the precision of the computation of the approximation. While the seesaw routine always makes 100 iterations, the number of random starting points was 1000, 1500, 3000 respectively. For each of the choices, the maximization was done for 200 steps. After 800 optimization steps in total there was no increase in the measure  $\mathcal{T}$  and the fluctuations were of the same size as those arising in the computation of the best network state approximation. The convergence of the algorithm is illustrated in Fig. 6.8 (right). The state found by the algorithm, denoted by  $|\psi_{\max}\rangle$ , yields a triangle measure  $\mathcal{T}$  very close to 0.6, see Tab. 6.4. This is very interesting as this state outperforms all states that have a high triangle measure known so far [325]. Further, the state shares an interesting entanglement structure. One finds that while the one-body marginals onto the first five systems have all the same spectrum (0.4, 0.6), the marginal onto system six is approximately pure, hence unentangled with the remaining particles.

State $ \psi\rangle$	GHZ <sub>2</sub>	GHZ <sub>3</sub>	GHZ <sub>4</sub>	W	AME(3,4)	$ \psi_{3,4}\rangle$	AS <sub>3</sub>	$ \psi\rangle_{\max}$
Measure $\mathcal{T}$	$\frac{1}{2}$	$\frac{5}{9}$	$\frac{1}{2}$	$\frac{1}{3}$	$\frac{1}{2}$	0	$\approx 0.46$	0.6

Table 6.4: Results of the seesaw optimization yielding upper bounds on the triangle measure  $\mathcal{T}$ . The states AME(3,4) and  $|\psi_{3,4}\rangle$  are defined in Eq. (6.29) and Eq. (6.30) respectively. AS<sub>3</sub> is the totally antisymmetric state on three qutrits as given by Eq. (6.28), embedded into the ququad system. The state  $|\psi_{\max}\rangle$  refers to the state found by the algorithm. The table is taken from Ref. [D].

## 6.11 Relation to upper bounds on the geometric measure

Our algorithm yields highly entangled states with respect to the geometric measure. This raises the question whether their entanglement saturates fundamental upper

bounds of this quantity. Interestingly, similar to the problem of computing the geometric measure of a quantum state, finding nontrivial upper bounds turns out to be a difficult task.

### 6.11.1 Bounds on the maximal entanglement

For a given physical system  $\mathcal{H} = \otimes_k \mathcal{H}_k$  with  $\mathcal{H}_k \cong \mathbb{C}^{d_k}$  one can assign the following value

$$\mathcal{G}(\mathcal{H}) := \sup\{G(|\varphi\rangle) : |\varphi\rangle \in \mathcal{H}\}, \quad (6.55)$$

which is characteristic for the space  $\mathcal{H}$ . Indeed, it can be shown that  $\mathcal{G}(\mathcal{H})$  is directly related to the inradius of  $B_{\mathcal{H}_1} \hat{\otimes} \cdots \hat{\otimes} B_{\mathcal{H}_n}$ , where  $B_{\mathcal{H}_j}$  denotes the unit ball in  $\mathcal{H}_j$  and  $\hat{\otimes}$  is the projective tensor product [286], which is for two closed convex sets  $K_1, K_2$  defined as  $K_1 \hat{\otimes} K_2 = \text{conv}\{x \otimes y \mid x \in K_1, y \in K_2\}$ . Finding upper bounds on the geometric measure means finding upper bounds on  $\mathcal{G}(\mathcal{H})$ , bounding the maximal amount of entanglement that can be present in the system. However, similar to the question of maximal tensor rank, this turns out to be a hard question with only partial answers so far [327]. A trivial upper bound can be obtained directly from the normalization of the state. Indeed, given an arbitrary state  $|\psi\rangle \in (\mathbb{C}^d)^{\otimes n}$  there must exist a tensor product of computational basis states  $|j_1 \cdots j_n\rangle$  such that  $|\langle \psi | j_1 \cdots j_n \rangle|^2 \geq d^{-n}$ . Consequently we have that  $\mathcal{G}((\mathbb{C}^d)^{\otimes n}) \leq 1 - d^{-n}$ . Although the derivation of this bound is trivial, it is difficult to obtain improved bounds. For instance, one could also consider upper bounds from a convex relaxation

$$\inf_{|\psi\rangle} \sup_{|\pi\rangle} |\langle \psi | \pi \rangle|^2 = \inf_{\rho \in \mathcal{S}} \sup_{\sigma \in \text{SEP}} \text{Tr}[\rho\sigma] \geq \sup_{\sigma \in \text{SEP}} \inf_{\rho \in \mathcal{S}} \text{Tr}[\rho\sigma] = \sup_{\sigma \in \text{SEP}} \lambda_{\min}(\sigma) = \frac{1}{d^n}, \quad (6.56)$$

where  $\mathcal{S}$  is the set of mixed states associated to  $\mathcal{H}$  and  $\text{SEP} \subset \mathcal{S}$  the set of separable states. The first equality is true due to the convexity of the objective function and the fact, that pure states (product states) are the extreme points of  $\mathcal{S}$  (SEP). Further, for any bounded function  $f(x, y)$  one has  $\sup_x f(x, y) \geq f(x, y) \Rightarrow \inf_y \sup_x f(x, y) \geq \inf_y f(x, y)$  and thus we obtain  $\inf_y \sup_x f(x, y) \geq \sup_x \inf_y f(x, y)$ . The last equality is attained for the maximally mixed state. However, it should be noted that this coincides with the trivial upper bound. This bound is not tight, even not for two qubits.

Also for the special case of symmetric states an upper bound on the geometric measure can be derived [328]. Here it should be noted that the space of symmetric quantum states is much smaller than the space of all states. Indeed, while the state space for  $n$  qubits has dimension  $2^n$ , the space of symmetric states only has dimension  $n + 1$ . In particular, one can show that for a  $n$  qudit system, the geometric measure is upper bounded by  $G \leq 1 - 1/c$  where  $c = \binom{n+d-1}{n}$ , which follows from a similar

normalization argument as the upper bound for generic states, but restricting to the Dicke basis. For  $n$  qubits this yields the bound  $G \leq 1 - 1/(n+1)$ , which is not tight for at least a small number of parties. However, it is interesting to see that for  $n \geq 5$ , the maximal entangled state found by the algorithm violates that bound.

### 6.11.2 Asymptotic scaling

Another open problem regards the asymptotic scaling of the geometric measure of multi-qubit systems (see Problem 8.27 in [286]). Here the question is whether there exists a constant  $C > 0$  and for any  $n \geq 1$  a quantum state  $|\psi\rangle \in (\mathbb{C}^2)^{\otimes n}$  such that  $G(|\psi\rangle) \geq 1 - 2^{-n}C$ . If  $|\psi\rangle \in (\mathbb{C}^2)^{\otimes n}$  is randomly chosen according to the Haar measure, with high probability it fulfills  $G(|\psi\rangle) \geq 1 - 2^{-n}Cn \log(n)$ , i.e., there is a parasitic factor  $n \log(n)$  [286]. Note that this bound is an improved version of Eq. (6.33) if the number of particles is large. Although we cannot solve that problem here, we can put bounds on the constant  $C$  under the assumption that the optimal states found by the algorithm are indeed the global optimizers of the geometric measure. If  $G_{\max,n}$  denotes the maximal measure of a  $n$  qubits system, we have  $C_n \geq C_{\min,n} = 2^n(1 - G_{\max,n})$ . One can see in Tab. 6.5 that the lower bound  $C_{\min}$  of the constant  $C$  grows with the number of qubits. In particular, the increase of  $C_{\max}$  is not decreasing with the number of qubits  $n$  as one would expect if there would exist a constant  $C$  which is independent of  $n$ . Indeed, in this case one would have  $C_{\min} \leq \tilde{C} < \infty$  for all  $n$  and as optimal constant  $C$  one can simply take the maximum (that exists by assumption) over all the  $C_{\min}$ .

### 6.11.3 The maximal entangled state and optimal norm constants

The maximal entangled state with respect to the geometric measure in a certain tensor space  $\mathcal{H} = \otimes_k \mathcal{H}_k$  has also an interpretation from the viewpoint of normed vector spaces. First note that  $\|\psi\|_\sigma = \sup_{|\pi\rangle} |\langle \pi | \psi \rangle|$  is a norm on  $\mathcal{H}$ . Further, because  $\mathcal{H}$  is finite dimensional, all norms on  $\mathcal{H}$  are equivalent, i.e., if  $\|\cdot\|$  and  $\|\cdot\|'$  are two norms on  $\mathcal{H}$ , then there exist constants  $C_1, C_2 > 0$  such that  $C_1 \|\psi\| \leq \|\psi\|' \leq C_2 \|\psi\|$  for all (possible un-normalized) vectors  $\psi \in \mathcal{H}$ . In our case, the two norms are given by  $\|\cdot\|_\sigma$  and  $\|\cdot\|_2$ . From this perspective, the algorithm solves the problem of finding

$$C := \min \left\{ \frac{\|\psi\|_\sigma}{\|\psi\|_2} : \psi \in \mathcal{H} \right\}. \quad (6.57)$$

Therefore, knowing  $C > 0$  implies that  $\|\psi\|_\sigma \|\psi\|_2^{-1} \geq C$  for all  $\psi$ , thus  $C \|\cdot\|_2 \leq \|\psi\|_\sigma$ . Because the algorithm explicitly yields the minimizing state, the inequality is tight, hence the found constant  $C$  is optimal.

System	$G_{\max}$	$C_{\min}$
2 qubits	1/2	2
3 qubits	5/9	32/9 = 3.555
4 qubits	5/9	32/9 = 3.555
5 qubits	$(1/36)(33 - \sqrt{3})$	$(8/9)(3 + \sqrt{3}) \approx 4.206$
6 qubits	16/3	16/3 = 5.333
7 qubits	0.941	$\approx 7.552$
8 qubits	0.961	$\approx 9.984$

Table 6.5: The maximal entanglement  $G_{\max}$  that can be present in a certain system according to the algorithm. This can be used to put lower bounds  $C_{\min}$  on the constant  $C$  in the asymptotic scaling of the geometric measure in multi-qubit systems. For the cases where no analytical expression of the quantum state and thus of  $G_{\max}$  could be derived, i.e., 7 and 8 qubits, we have taken the value for  $G_{\max}$  for the state found by the algorithm with the largest geometric measure. For each case, the algorithm was run for 100 random initial states while the geometric measure of the final states coincided with very small deviation. The table is taken from Ref. [D].

## 6.12 Conclusion and discussion

We have presented an iterative method for the computation of maximally resourceful quantum states. We provided a convergence analysis and showed that in each step the resourcefulness of the iterates increases. We illustrated our approach for the special case of the geometric measure, allowing us to identify interesting quantum states, discover novel AME states, and characterize highly entangled subspaces which may be useful for information processing. We further demonstrated the universality of the algorithm for various other quantifiers, yielding novel forms of correlations in the triangle network. Concerning further research, our results also suggest a variety of avenues for further theoretical exploration. Can the algorithm be used to find new AME states for cases where the existence is still open, e.g., for systems consisting of more than five quhex, or to find new SLOCC inequivalent AME states? In particular, we have strong numerical evidence that there exists a second AME state for the four quhex case. From a mathematical perspective, the algorithm can give insights into the structure of tensor spaces and could offer intuition to solve open problems concerning the asymptotics of tensor norms [286].

# 7 Real eigenstructure of regular simplex tensors

Characterizing the eigenvectors of a given tensor is an important task for many applications involving large data arrays, such as high-dimensional quantum states and data science. However, this turns out to be in general a computationally hard problem. Here we provide a full characterization of the real eigenstructure of regular simplex tensors. This is supplemented by the robustness analysis of all normalized eigenvectors. The robustness of a tensor eigenvector is of particular importance if it should be computed by the tensor power method. Finally, we discuss the relationship between the obtained eigenvectors and the generators from the symmetric tensor decomposition. This Chapter is based on Project [H]. Each of the three authors contributed in equal parts to this work.

## 7.1 Motivation

A prominent difficulty in quantum simulation is the exponentially growing dimension of the underlying Hilbert space, rendering an efficient treatment impossible. The same problem appears in the more general context of multivariate data arrays with a large number of modes [329, 330]. Apart from quantum physics, they play an important role in neuroscience [331], algebraic statistics [332], computer vision [333] as well as in the algorithmic knowledge retrieval from large datasets. Clearly, a multimodal dataset can only be handled efficiently after imposing a certain structural representation, which typically also encodes many of its geometric properties, like symmetries or other correlations between the tensor entries. In this context, one often focuses on a symmetric tensor  $\mathcal{T}$  whose entries are invariant under permutation of the indices and are typically stored in the symmetric decomposition format,

$$\mathcal{T} = \sum_{j=1}^r \lambda_j \vec{v}_j^{\otimes d}, \quad (7.1)$$

which is a linear combination of the  $d$ -fold tensor product of certain  $n$ -dimensional, normalized vectors  $\vec{v}_1, \dots, \vec{v}_r \in \mathbb{R}^n$ , with real weights  $\lambda_1, \dots, \lambda_r \in \mathbb{R}$ . The set of all symmetric tensors forms a vector space and is denoted by  $S^d(\mathbb{R}^n)$ . Generalizing concepts

from linear algebra [280, 334], a real eigenpair  $(\lambda, \vec{v})$  with  $\lambda \in \mathbb{R}$  and  $\vec{v} \in \mathbb{R}^n$  of a symmetric tensor  $\mathcal{T} \in S^d(\mathbb{R}^n)$  fulfills  $\mathcal{T} \cdot \vec{v}^{\otimes(d-1)} = \lambda \vec{v}$ , where the dot operation denotes the partial contraction of  $\mathcal{T}$  by  $\vec{v}$  along all but one of the  $d$  modes. Eigenpairs of a tensor are important in many applications [288], as they are highly linked to the best rank-one approximation problem [335].

Here it is natural to look for relations between the normalized eigenvectors of a symmetric tensor  $\mathcal{T}$  and the generating vectors  $\{\vec{v}_j\}$  of its symmetric decomposition in Eq. (7.1). In the matrix case, i.e.,  $\mathcal{T} \in S^2(\mathbb{R}^n)$ , these sets coincide. In contrast, in the general setting  $d \geq 2$ , both sets of vectors usually only coincide if one imposes additional constraints on the vectors  $\{\vec{v}_j\}$ , like orthogonality. If  $\mathcal{T} \in S^d(\mathbb{R}^n)$  has a decomposition in the form of Eq. (7.1) such that  $\{\vec{v}_1, \dots, \vec{v}_r\} \subset \mathbb{R}^n$  is an orthonormal set, the tensor  $\mathcal{T}$  is called *odeco* tensor. The set of odeco tensors turns out to be structurally rich [336]. In particular, all eigenvectors  $\vec{v}_j$  of an odeco tensor are attractive fixed points of the tensor power iteration map, which is used to numerically solve the tensor eigenvalue problem. However, the set of odeco tensors forms a variety [336] which is of very small dimension, straiten its usefulness in the analysis of generic datasets.

In a recent work [337], significant progress has been made in the analysis of those symmetric tensors  $\mathcal{T}$  whose generating vectors  $\vec{v}_j \in \mathbb{R}^n$  constitute an overcomplete set which is still close to an orthonormal basis, e.g., a tight frame. The set of such frame decomposable tensors, in short, *fradeco* tensors, is significantly larger than the odeco class. However, the eigenvectors of a fradeco tensor usually deviate from the underlying frame elements and it is an open problem under which conditions the eigenvectors of a fradeco tensor can be recovered by the tensor power method [338].

In this Chapter we focus on the special case of regular simplex tensors. These are defined as tensors whose symmetric decomposition in Eq. (7.1) uses equal weights  $\lambda_j = 1$  and is induced by an overcomplete equiangular set of  $n + 1$  vectors  $\vec{v}_j$  from  $\mathbb{R}^n$ . By reformulating the tensor eigenvalue problem as an algebraic set of equations in the barycentric coordinates of the eigenvector with respect to the frame elements  $\{\vec{v}_j\}$ , we develop a full analysis of the real eigenstructure of a regular simplex tensor with local dimension  $n \geq 2$  and an arbitrary number of modes  $d \geq 2$ . We begin by formalizing the notion of simplex frames, tensor eigenvectors and the tensor power iteration in Section 7.2. We proceed by translating the eigenvector property into an equivalent system of algebraic conditions in Section 7.3, which allows us to enumerate all possible normalized eigenpairs. Afterwards, we discuss in Section 7.4 the special case  $n = 2$  and in Section 7.5 the case  $n = 3$  in more detail. In Section 7.6, we study the robustness of all normalized eigenvectors with respect to the tensor power iteration.

## 7.2 Mathematical concepts and notation

A real-valued tensor or hypermatrix of order  $d \in \mathbb{N}$  and with local dimensions  $n_1, \dots, n_d \in \mathbb{N}$  is a  $d$ -variate data field  $\mathcal{T} \in \otimes_{j=1}^d \mathbb{R}^{n_j}$ . Clearly, in the special case where the number of modes  $d$  is equal to 1 or 2,  $\mathcal{T}$  is a column vector or a matrix respectively. Similar to our treatment of a quantum state, we denote the  $(i_1, \dots, i_d)$ -th entry of a tensor  $\mathcal{T}$  by  $\mathcal{T}_{i_1, \dots, i_d}$ . If all mode dimensions  $n_j$  are equal to  $n \in \mathbb{N}$ , the tensor  $\mathcal{T} \in (\mathbb{R}^n)^{\otimes d}$  is called cubic. Further, a cubic tensor  $\mathcal{T}$  is called (super)symmetric if

$$\mathcal{T}_{i_1, \dots, i_d} = \mathcal{T}_{i_{\sigma(1)}, \dots, i_{\sigma(d)}} \quad \text{for all permutations } \sigma : \{1, \dots, d\} \rightarrow \{1, \dots, d\}. \quad (7.2)$$

The set of all symmetric tensors of order  $d$  and local dimension  $n$  will be denoted by  $S^d(\mathbb{R}^n)$ . For a given vector  $\vec{v} \in \mathbb{R}^n$ , the  $d$ -fold tensor product  $\vec{v}^{\otimes d} \in S^d(\mathbb{R}^n)$  is called a symmetric rank-1 tensor. It can be easily seen that each symmetric tensor  $\mathcal{T} \in S^d(\mathbb{R}^n)$  admits a finite symmetric decomposition as in Eq. (7.1) where  $\vec{v}_1, \dots, \vec{v}_r \in \mathbb{R}^n$  with  $\|\vec{v}_j\| = 1$  and  $\lambda_j \in \mathbb{R}$  for  $1 \leq j \leq r$ . The smallest possible number  $r \in \mathbb{N}$  for which a decomposition of the form in Eq. (7.1) exists is called the symmetric rank of  $\mathcal{T}$ . If the number of modes  $d$  is odd, one can assume that all weights  $\lambda_j$  which appear in Eq. (7.1) are positive.

If  $\mathcal{T} \in S^d(\mathbb{R}^n)$  has a decomposition of the form in Eq. (7.1) such that the vectors  $\{\vec{v}_1, \dots, \vec{v}_r\}$  form an orthonormal set,  $\mathcal{T}$  is called orthogonally decomposable or, in short, an odeco tensor. However, the set of odeco tensors is relatively small since the symmetric rank of an odeco tensor cannot exceed the local dimension  $n$ .

### 7.2.1 Tensor eigenvalues and the tensor power method

A vector  $\vec{v} \in \mathbb{R}^n \setminus \{\vec{0}\}$  is called a real eigenvector<sup>1</sup> of  $\mathcal{T} \in S^d(\mathbb{R}^n)$  with eigenvalue  $\mu \in \mathbb{R}$  [280, 334] if

$$\mathcal{T} \cdot \vec{v}^{\otimes(d-1)} = \mu \vec{v}. \quad (7.3)$$

<sup>1</sup>Here it is important to notice that there exists a huge variety of definitions of tensor eigenvalues [288]. First, in its most general form, a number  $\mu \in \mathbb{C}$  is an eigenvalue of a not necessarily symmetric but cubic tensor  $\mathcal{T} \in (\mathbb{R}^n)^{\otimes d}$ , if  $\mu$  together with a nonzero vector  $\vec{v} \in \mathbb{C}^n$  are solutions of the homogeneous polynomial equations  $\mathcal{T} \cdot \vec{v}^{\otimes(d-1)} = \mu \vec{v}$ , where the contraction is with respect to the first  $d-1$  modes. In this case, one calls  $\vec{v}$  an eigenvector of  $\mathcal{T}$  associated with the eigenvalue  $\mu$ . The eigenvalue  $\mu \in \mathbb{C}$  is called an  $H$ -eigenvalue if it has a real eigenvector  $\vec{v}$ , which is then called  $H$ -eigenvector. As the tensor  $\mathcal{T}$  is real-valued, it immediately follows that if  $\vec{v}$  is an  $H$ -eigenvector, the corresponding  $H$ -eigenvalue is real. An eigenvalue that is not an  $H$ -eigenvalue is called  $N$ -eigenvalue. Notice that even though an  $H$ -eigenvalue is a real number, a real eigenvalue is not necessarily an  $H$ -eigenvalue. Further, there are  $E$ - and  $Z$ -eigenvalues, where an additional normalization constraint is incorporated into the definition. With the same notation as above,  $\mu \in \mathbb{C}$  is called an  $E$ -eigenvalue of  $\mathcal{T}$  if  $\mu$  together with  $\vec{v} \in \mathbb{C}^n$  solves the system  $\mathcal{T} \cdot \vec{v}^{\otimes(d-1)} = \mu \vec{v}$  with  $\vec{v}^\top \vec{v} = 1$ . An  $E$ -eigenvalue is called a  $Z$ -eigenvalue if it has a real  $E$ -eigenvector. Again, a  $Z$ -eigenvalue is a real  $E$ -eigenvalue but a real  $E$ -eigenvalue is not necessarily a  $Z$ -eigenvalue. See also Ref. [288].



Here  $\mathcal{T} \cdot \vec{v}^{\otimes(d-1)}$  denotes the partial contraction of  $\mathcal{T}$  by  $\vec{v}$  along all but one of the  $d$  modes,

$$(\mathcal{T} \cdot \vec{v}^{\otimes(d-1)})_j := \sum_{i_1, \dots, i_{d-1}=1}^n \mathcal{T}_{i_1, \dots, i_{d-1}, j} \vec{v}_{i_1} \cdots \vec{v}_{i_{d-1}}. \quad (7.4)$$

It is important to note that due to the symmetry of  $\mathcal{T}$ , it is irrelevant which particular modes are used in the  $(d-1)$ -fold sum in Eq. (7.4). A tuple  $(\vec{v}, \mu)$  consisting of an eigenvector  $\vec{v} \in \mathbb{R}^n \setminus \{\vec{0}\}$  and an associated eigenvalue  $\mu \in \mathbb{R}$  is called an eigenpair of  $\mathcal{T}$ . If in addition  $\|\vec{v}\| = 1$ , the eigenpair  $(\vec{v}, \mu)$  is called normalized eigenpair<sup>2</sup>. Normalized eigenpairs can be understood as critical points of the map  $\vec{v} \mapsto \langle \mathcal{T}, \vec{v}^{\otimes d} \rangle$  under the norm constraint  $\|\vec{v}\| = 1$ .

By construction, the left-hand side of Eq. (7.3) is  $(d-1)$ -homogeneous in the coordinates of the vector  $\vec{v}$  and therefore each eigenpair  $(\vec{v}, \mu)$  of  $\mathcal{T}$  induces the eigenpair  $(t\vec{v}, t^{d-2}\mu)$  for  $t \neq 0$ . In particular, if  $(\vec{v}, \mu)$  is an eigenpair of  $\mathcal{T}$ , we obtain with  $t = \|\vec{v}\|^{-1}$  that

$$\left( \frac{\vec{v}}{\|\vec{v}\|}, \frac{\mu}{\|\vec{v}\|^{d-2}} \right) \quad (7.5)$$

is a normalized eigenpair of  $\mathcal{T}$ .

Similarly, by normalizing both sides in Eq. (7.3), one can see that a necessary condition for a normalized vector  $\vec{v} \in \mathbb{R}^n$  to be an eigenvector of  $\mathcal{T} \in S^d(\mathbb{R}^n)$  with eigenvalue  $\mu > 0$  is that  $\vec{v}$  is a fixed point of the map

$$\varphi : \mathbb{R}^n \setminus \{\vec{0}\} \rightarrow \mathbb{R}^n \setminus \{\vec{0}\}, \quad \varphi(\vec{v}) := \frac{\mathcal{T} \cdot \vec{v}^{\otimes(d-1)}}{\|\mathcal{T} \cdot \vec{v}^{\otimes(d-1)}\|}. \quad (7.6)$$

More generally, if  $(\vec{v}, \mu)$  is a normalized eigenpair of  $\mathcal{T}$  with eigenvalue  $\mu \neq 0$ , we have  $s\varphi(\vec{v}) = \vec{v}$  with  $s = \text{sign}(\mu) = \mu/|\mu|^{-1}$ . The associated canonical fixed point iteration

$$\vec{v}^{(j+1)} := \varphi(\vec{v}^{(j)}), \quad \text{for } j = 0, 1, \dots \quad (7.7)$$

is called tensor power iteration [339]. For the matrix case, i.e.,  $\mathcal{T} \in S^2(\mathbb{R}^n)$  this iteration converges for any starting point  $\vec{v}^{(0)} \in \mathbb{R}^n$  to an eigenvector corresponding to the largest eigenvalue of  $\mathcal{T}$  in modulus. This changes for the more general case  $d > 2$ . More precisely, there exists a distinguished class of eigenpairs with respect to their behavior under this iteration. A unit vector  $\vec{v} \in \mathbb{R}^n$  is called a robust eigenvector of  $\mathcal{T} \in S^d(\mathbb{R}^n)$  if there exists an open neighborhood of  $\vec{v}$  such that the iterates that appear in Eq. (7.7) starting with any  $\vec{w}$  from this neighborhood converge to  $\vec{v}$ . Obviously, non-robust eigenvectors  $\vec{v}$  cannot be computed by using the tensor power iteration unless the starting point equals  $\vec{v}$ . From an analytical viewpoint, the robustness of a fixed point of the map in Eq. (7.6) can be quantified by the spectral radius of the Jacobian of  $\varphi$  evaluated at that particular fixed point.

<sup>2</sup>The notion of an eigenpair induced by Eq. (7.3) corresponds to the class of  $H$ -eigenvalues and  $H$ -eigenvectors, while a normalized eigenpair induced by Eq. (7.5) corresponds to  $Z$ -eigenvalues and  $Z$ -eigenvectors.

### 7.2.2 Frames and simplex tensors

The set of odeco tensors can be substantially enlarged by replacing the set of orthonormal vectors  $\{\vec{v}_1, \dots, \vec{v}_r\}$  in the symmetric decomposition by a so called frame. A family of  $r \geq n$  vectors  $\{\vec{v}_1, \dots, \vec{v}_r\} \subset \mathbb{R}^n$  is called a frame for  $\mathbb{R}^n$  if there exist constants  $B \geq A > 0$  such that

$$A\|\vec{v}\|^2 \leq \sum_{j=1}^r |\langle \vec{v}, \vec{v}_j \rangle|^2 \leq B\|\vec{v}\|^2 \quad (7.8)$$

for all vectors  $\vec{v} \in \mathbb{R}^n$ . A frame  $\{\vec{v}_1, \dots, \vec{v}_r\}$  with equal frame constants  $A = B$  is called a tight frame and a tight frame  $\{\vec{v}_1, \dots, \vec{v}_r\}$  is called a unit-norm tight frame if additionally  $\|\vec{v}_j\| = 1$  for all  $1 \leq j \leq r$ . It is well known [340] that a set  $\{\vec{v}_1, \dots, \vec{v}_r\}$  is a unit-norm tight frame for  $\mathbb{R}^n$  if and only if

$$VV^\top = A\mathbb{1}, \quad \text{where } V := (\vec{v}_1 \cdots \vec{v}_r) \in \mathbb{R}^{n \times r}. \quad (7.9)$$

Typical examples of unit-norm tight frames in  $\mathbb{R}^n$  are given by orthonormal bases  $\{\vec{v}_1, \dots, \vec{v}_n\}$ , with  $r = n$  and  $A = 1$ , and so called simplex frames  $\{\vec{v}_1, \dots, \vec{v}_{n+1}\} \subset \mathbb{R}^n$ , where  $r = n + 1$ ,  $A = \frac{n+1}{n}$  and  $\vec{v}_j \in \mathbb{R}^n$  is the orthogonal projection of the  $j$ -th unit vector  $\vec{e}_j \in \mathbb{R}^{n+1}$  onto the orthogonal complement of  $\vec{1}_{n+1} := (1, \dots, 1) \in \mathbb{R}^{n+1}$ , and subsequent normalization. More precisely, we have

$$\vec{v}_k = \begin{cases} \sqrt{1 + \frac{1}{n}} \vec{e}_j - \frac{1}{n^{3/2}} (\sqrt{n+1} - 1) \vec{1}_n & , 1 \leq j \leq n, \\ -\frac{1}{\sqrt{n}} \vec{1}_n, & j = n + 1. \end{cases} \quad (7.10)$$

Further we have

$$VV^\top = \frac{n+1}{n} \mathbb{1}, \quad V^\top V = \frac{n+1}{n} \mathbb{1} - \frac{1}{n} \vec{1}_{n+1} \vec{1}_{n+1}^\top, \quad (7.11)$$

and the nullspace of  $V^\top V$  and of  $V$  is spanned by  $\vec{v}_{n+1}$ .

## 7.3 Characterizing eigenpairs of regular simplex tensors

In the following we will perform an exhaustive search for eigenpairs of the simplex tensor

$$\mathcal{T} = \sum_{j=1}^{n+1} \vec{v}_j^{\otimes d}, \quad (7.12)$$

where  $n, d \geq 2$  and  $\{\vec{v}_1, \dots, \vec{v}_{n+1}\} \subset \mathbb{R}^n$  is a simplex frame. We will start our analysis by looking at the generic case  $d, n \geq 2$ . By using the linear independence of each  $n$ -element subset of the simplex frame  $\{\vec{v}_1, \dots, \vec{v}_{n+1}\}$ , we can easily deduce the following system of equations for the coordinates of an eigenvector  $\vec{v} \in \mathbb{R}^n \setminus \{\vec{0}\}$  of  $\mathcal{T}$  with respect to the basis  $\{\vec{v}_1, \dots, \vec{v}_n\}$  of  $\mathbb{R}^n$ .

**Lemma 46.**  $(\vec{v}, \mu)$  is an eigenpair of a simplex tensor  $\mathcal{T}$  given by Eq. (7.12) if and only if  $\vec{v} = \sum_{k=1}^n \alpha_k \vec{v}_k$  for some  $\alpha_1, \dots, \alpha_n \in \mathbb{R}$  with  $\sum_{k=1}^n |\alpha_k| > 0$  and

$$\mu \alpha_k = \left( \alpha_k - \frac{1}{n} \sum_{\substack{j=1 \\ j \neq k}}^n \alpha_j \right)^{d-1} - \left( -\frac{1}{n} \sum_{j=1}^n \alpha_j \right)^{d-1}, \quad 1 \leq k \leq n. \quad (7.13)$$

*Proof.*  $\{\vec{v}_1, \dots, \vec{v}_n\}$  is a basis for  $\mathbb{R}^n$ , so that each eigenvector  $\vec{v} \in \mathbb{R}^n \setminus \{\vec{0}\}$  has a unique representation  $\vec{v} = \sum_{j=1}^n \alpha_j \vec{v}_j$  with  $\sum_{j=1}^n |\alpha_j| > 0$ . By inserting this representation into the eigenvector equation Eq. (7.3), and by using the normalization  $\|\vec{v}_j\| = 1$ , Eq. (7.11) and  $\sum_{j=1}^{n+1} \vec{v}_j = \vec{0}$ , we obtain that

$$\begin{aligned} \mu \sum_{k=1}^n \alpha_k \vec{v}_k &= \sum_{k=1}^{n+1} \left\langle \vec{v}_k, \sum_{j=1}^n \alpha_j \vec{v}_j \right\rangle^{d-1} \vec{v}_k \\ &= \sum_{k=1}^n \left( \alpha_k - \frac{1}{n} \sum_{1 \leq j \neq k \leq n} \alpha_j \right)^{d-1} \vec{v}_k + \left( -\frac{1}{n} \sum_{j=1}^n \alpha_j \right)^{d-1} \vec{v}_{n+1} \\ &= \sum_{k=1}^n \left( \left( \alpha_k - \frac{1}{n} \sum_{\substack{j=1 \\ j \neq k}}^n \alpha_j \right)^{d-1} - \left( -\frac{1}{n} \sum_{j=1}^n \alpha_j \right)^{d-1} \right) \vec{v}_k, \end{aligned}$$

which yields Eq. (7.13) after using the linear independence of  $\{\vec{v}_1, \dots, \vec{v}_n\}$ .  $\square$

By using that each elementary tensor  $\vec{v}_k^{\otimes d}$  is weighted equally within  $\mathcal{T}$ , we can deduce the following permutation symmetry of all eigenpairs.

**Lemma 47.** Let  $(\vec{v}, \mu)$  be an eigenpair of simplex tensor  $\mathcal{T}$  given by Eq. (7.12) with  $\vec{v} = \sum_{k=1}^n \alpha_k \vec{v}_k$  for certain  $\alpha_k \in \mathbb{R}$ . Then for each permutation  $\sigma$  of  $\{1, \dots, n+1\}$ , also the vectors  $\sum_{k=1}^n \alpha_k \vec{v}_{\sigma(k)}$  are eigenvectors of  $\mathcal{T}$  with the same eigenvalue  $\mu$ .

*Proof.*  $(\vec{v}, \mu)$  is an eigenpair with  $\vec{v} = \sum_{k=1}^n \alpha_k \vec{v}_k$ , so that Eq. (7.13) holds true by Lemma 46. Let  $\sigma$  be a permutation of  $\{1, \dots, n+1\}$ . If  $p := \sigma(n+1) = n+1$ , we have  $\{1, \dots, n\} = \{\sigma(1), \dots, \sigma(n)\}$ , so that Eq. (7.13) holds true for all  $\alpha_{\sigma^{-1}(k)}$ ,  $1 \leq k \leq n$ , i.e., Lemma 46 yields that

$$\sum_{k=1}^n \alpha_k \vec{v}_{\sigma(k)} = \sum_{k=1}^n \alpha_{\sigma^{-1}(k)} \vec{v}_k$$

is an eigenvector of  $\mathcal{T}$  with eigenvalue  $\mu$ . If  $p = \sigma(n+1) \in \{1, \dots, n\}$ , we have  $q := \sigma^{-1}(n+1) \in \{1, \dots, n\}$ . By using that  $\sum_{k=1}^{n+1} \vec{v}_k = \vec{0}$  and because of the equivalence

$$1 \leq \sigma^{-1}(k) \leq n \wedge \sigma^{-1}(k) \neq q \quad \Leftrightarrow \quad 1 \leq k \leq n \wedge k \neq p, \quad (7.14)$$

we can write

$$\sum_{k=1}^n \alpha_k \vec{v}_{\sigma(k)} = \alpha_q \vec{v}_{n+1} + \sum_{\substack{k=1 \\ k \neq q}}^n \alpha_k \vec{v}_{\sigma(k)} = -\alpha_q \vec{v}_p + \sum_{\substack{k=1 \\ k \neq p}}^n (\alpha_{\sigma^{-1}(k)} - \alpha_q) \vec{v}_k.$$

Therefore, it remains to prove that Eq. (7.13) holds true for

$$\beta_k := \begin{cases} \alpha_{\sigma^{-1}(k)} - \alpha_q, & 1 \leq k \neq p \leq n, \\ -\alpha_q, & k = p, \end{cases}$$

because then the claim follows by an application of Lemma 46. If  $k = p$ , we compute that by means of Eq. (7.14) and Eq. (7.13),

$$\begin{aligned} & \left( \beta_p - \frac{1}{n} \sum_{\substack{j=1 \\ j \neq p}}^n \beta_j \right)^{d-1} - \left( -\frac{1}{n} \sum_{j=1}^n \beta_j \right)^{d-1} \\ &= \left( -\alpha_q - \frac{1}{n} \sum_{\substack{j=1 \\ j \neq p}}^n (\alpha_{\sigma^{-1}(j)} - \alpha_q) \right)^{d-1} - \left( \frac{1}{n} \alpha_q - \frac{1}{n} \sum_{\substack{j=1 \\ j \neq p}}^n (\alpha_{\sigma^{-1}(j)} - \alpha_q) \right)^{d-1} \\ &= \left( -\frac{1}{n} \alpha_q - \frac{1}{n} \sum_{\substack{j=1 \\ j \neq p}}^n \alpha_{\sigma^{-1}(j)} \right)^{d-1} - \left( \alpha_q - \frac{1}{n} \sum_{\substack{j=1 \\ j \neq p}}^n \alpha_{\sigma^{-1}(j)} \right)^{d-1} \\ &= \left( -\frac{1}{n} \sum_{j=1}^n \alpha_j \right)^{d-1} - \left( \alpha_q - \frac{1}{n} \sum_{\substack{j=1 \\ j \neq q}}^n \alpha_j \right)^{d-1} = -\alpha_q = \beta_p. \end{aligned}$$

If  $1 \leq k \leq n$  and  $k \neq p$ , we can argue in a similar way, again using Eq. (7.14):

$$\begin{aligned} & \left( \beta_k - \frac{1}{n} \sum_{\substack{j=1 \\ j \neq k}}^n \beta_j \right)^{d-1} - \left( -\frac{1}{n} \sum_{j=1}^n \beta_j \right)^{d-1} \\ &= \left( \alpha_{\sigma^{-1}(k)} - \alpha_q + \frac{1}{n} \alpha_q - \frac{1}{n} \sum_{\substack{j=1 \\ j \notin \{k,p\}}}^n (\alpha_{\sigma^{-1}(j)} - \alpha_q) \right)^{d-1} - \left( \frac{1}{n} \alpha_q - \frac{1}{n} \sum_{\substack{j=1 \\ j \neq p}}^n (\alpha_{\sigma^{-1}(j)} - \alpha_q) \right)^{d-1} \\ &= \left( \alpha_{\sigma^{-1}(k)} - \frac{1}{n} \alpha_q - \frac{1}{n} \sum_{\substack{j=1 \\ j \notin \{k,p\}}}^n \alpha_{\sigma^{-1}(j)} \right)^{d-1} - \left( \alpha_q - \frac{1}{n} \sum_{\substack{j=1 \\ j \neq p}}^n \alpha_{\sigma^{-1}(j)} \right)^{d-1} \\ &= \left( \alpha_{\sigma^{-1}(k)} - \frac{1}{n} \sum_{\substack{j=1 \\ j \neq \sigma^{-1}(k)}}^n \alpha_j \right)^{d-1} - \left( \alpha_q - \frac{1}{n} \sum_{\substack{j=1 \\ j \neq q}}^n \alpha_j \right)^{d-1} = \alpha_{\sigma^{-1}(k)} - \alpha_q = \beta_k. \end{aligned}$$

□

In view of the fact that  $\mathbb{R}^n$  can be decomposed into the conical hulls

$$\left\{ \sum_{k=1}^n \alpha_k \vec{v}_{\sigma(k)} : \alpha_k \geq 0 \right\}, \quad \sigma : \{1, \dots, n+1\} \rightarrow \{1, \dots, n+1\} \text{ permutation,}$$

Lemma 47 tells us that it suffices to compute all eigenpairs  $(\vec{v}, \mu)$  with eigenvectors from the set

$$\left\{ \sum_{k=1}^n \alpha_k \vec{v}_k : \alpha_k \geq 0, \sum_{j=1}^n |\alpha_j| > 0 \right\}$$

of all nontrivial conical combinations from the linearly independent set  $\{\vec{v}_1, \dots, \vec{v}_n\}$ . What is more, by the  $(d-1)$ -homogeneity of the right-hand side of Eq. (7.13), it is sufficient to consider all convex combinations

$$\vec{v} = \sum_{k=1}^n s_k \vec{v}_k, \quad 0 \leq s_k \leq 1, \quad \sum_{k=1}^n s_k = 1 \quad (7.15)$$

from  $\{\vec{v}_1, \dots, \vec{v}_n\}$  as eigenvector candidates. By inserting this very family of vectors into Lemma 46, we obtain the following eigenpair conditions.

**Lemma 48.** *Let  $0 \leq s_k \leq 1$  for  $1 \leq k \leq n-1$ , and  $s_n := 1 - \sum_{k=1}^{n-1} s_k$ . Then  $(\sum_{k=1}^n s_k \vec{v}_k, \mu)$  with  $\mu \in \mathbb{R}$  is an eigenpair of simplex tensor  $\mathcal{T}$  given by Eq. (7.12) if and only if*

$$\mu s_k = \frac{1}{n^{d-1}} \left( ((n+1)s_k - 1)^{d-1} - (-1)^{d-1} \right), \quad 1 \leq k \leq n. \quad (7.16)$$

*Proof.* The claim follows by inserting  $\alpha_k = s_k$  into Eq. (7.13), and by exploiting that  $\sum_{k=1}^n s_k = 1$ .  $\square$

By means of the function

$$g(s) := ((n+1)s - 1)^{d-1} - (-1)^{d-1}, \quad s \in \mathbb{R}, \quad d \geq 2, \quad (7.17)$$

we can rewrite the eigencondition in Eq. (7.16) in a compact way as

$$\mu n^{d-1} s_k = g(s_k), \quad 1 \leq k \leq n. \quad (7.18)$$

The eigenvalue  $\mu$  of an eigenpair  $(\sum_{k=1}^n s_k \vec{v}_k, \mu)$ ,  $s_k \geq 0$ ,  $\sum_{k=1}^n s_k = 1$ , can hence be computed by

$$\mu = \frac{1}{n^{d-1}} \sum_{k=1}^n g(s_k). \quad (7.19)$$

However, we still have to solve the remaining system of eigenvalue conditions

$$\begin{cases} s_k \sum_{j=1}^n g(s_j) = g(s_k), & 1 \leq k \leq n-1, \\ s_n = 1 - \sum_{j=1}^{n-1} s_j, \\ s_k \geq 0, & 1 \leq k \leq n. \end{cases} \quad (7.20)$$

**Example 49.** *If  $n = 2$ , Eq. (7.20) reads as the single equation in  $s = s_1$*

$$s(g(s) + g(1-s)) = g(s), \quad 0 \leq s \leq 1. \quad (7.21)$$

*If  $n = 3$ , Eq. (7.20) reads as the coupled system of equations in  $s = s_1$  and  $t = s_2$*

$$\begin{cases} s(g(s) + g(t) + g(1-s-t)) = g(s), \\ t(g(s) + g(t) + g(1-s-t)) = g(t), \end{cases} \quad s, t \geq 0, \quad s+t \leq 1. \quad (7.22)$$

In order to deduce the solution set of Eq. (7.20), let us first analyze the auxiliary function  $g$ .

**Lemma 50.** *Let  $n, d \geq 2$ , and let  $g$  be defined as in Eq. (7.17).*

1. We have  $g(0) = 0$ ,  $g(\frac{1}{n+1}) = (-1)^d$ ,  $g(\frac{2}{n+1}) = 1 + (-1)^d$ , and  $g(1) = n^{d-1} + (-1)^d > 0$ .
2. If  $d \geq 3$ , we have  $g'(\frac{1}{n+1}) = 0$ .
3. If  $d$  is even,  $g$  is strictly increasing and we have  $g(s) > 0$  for all  $s > 0$ .
4. If  $d$  is odd,  $g$  is strictly convex with a unique local minimum at  $s = \frac{1}{n+1}$ , and we have  $g(\frac{2}{n+1}) = 0$ , so that  $g(s) < 0$  for  $0 < s < \frac{2}{n+1}$  and  $g(s) > 0$  for  $\frac{2}{n+1} < s \leq 1$ .
5. The polynomial  $p : s \mapsto \frac{g(s)}{s}$  is well-defined. If  $d = 2$ ,  $p(s) = n + 1$  is constant. If  $d$  is odd,  $p$  is strictly increasing on  $[0, \infty)$ . If  $d \geq 4$  is even, there exists a point  $s^* \in [\frac{1}{n}, \frac{2}{n+1})$  such that  $p$  is strictly decreasing on  $[0, s^*]$  and strictly increasing on  $[s^*, \infty)$ . We have  $s^* = \frac{1}{n}$  if and only if  $(n, d) = (2, 4)$ .
6. If  $d$  is odd, the  $m$ -variate polynomial

$$\mathbb{R}^m \ni (s_1, \dots, s_m) \mapsto g\left(1 - \sum_{k=1}^m s_k\right) + \sum_{k=1}^m g(s_k)$$

is strictly convex, with unique minimum at  $\vec{s}^* := (\frac{1}{m+1}, \dots, \frac{1}{m+1})$  and value  $(m+1)g(\frac{1}{m+1})$ ,  $\vec{s}^*$  lying in the interior of the  $m$ -dimensional unit simplex

$$\Delta_m := \{(s_1, \dots, s_m) : s_k \geq 0, \sum_{j=1}^m s_j \leq 1\}. \quad (7.23)$$

*Proof.* Claim (1) follows directly from Eq. (7.17). To prove (2) let  $d \geq 3$ . We have  $g'(s) = (n+1)(d-1)((n+1)s-1)^{d-2}$ , so  $g'(\frac{1}{n+1}) = 0$ . To (3). If  $d$  is even, we see that

$$g(s) = ((n+1)s-1)^{d-1} + 1, \quad s \in \mathbb{R},$$

is a composition of the strictly increasing functions  $s \mapsto (n+1)s-1$  and  $t \mapsto t^{d-1} + 1$ . Therefore,  $s > 0$  implies that  $g(s) > g(0) = 0$ . To (4). If  $d$  is odd,  $g$  is strictly convex as a sum of the strictly convex function  $s \mapsto ((n+1)s-1)^{d-1}$  and a constant. In view of  $g'(0) = (n+1)(d-1)(-1)^d < 0$  and of  $g(1) > 0$ , see (1), the convexity of  $g$  implies that there exists exactly one further zero of  $g$  in the open interval  $(0, 1)$ , namely  $s = \frac{2}{n+1}$ , because the oddity of  $d$  and the identity

$$a^k - b^k = (a-b) \sum_{j=0}^{k-1} a^j b^{k-1-j}$$

entail that

$$g(s) = ((n+1)s - 1)^{d-1} - 1 = ((n+1)s - 2) \sum_{j=0}^{d-2} ((n+1)s - 1)^j.$$

By the continuity and strict convexity of  $g$ , it follows that  $g(s) < 0$  for  $0 < s < \frac{2}{n+1}$  and  $g(s) > 0$  for  $\frac{2}{n+1} < s \leq 1$ . To (5). In view of  $g(0) = 0$ , see (1),  $p(s) := \frac{g(s)}{s}$  defines a polynomial of degree  $d-2$ , and  $p(s) = n+1$  if  $d=2$ . If  $d \geq 3$ , we compute that for  $s > 0$ ,

$$\begin{aligned} p'(s) &= \frac{g'(s)s - g(s)}{s^2} \\ &= \frac{(d-1)(n+1)s((n+1)s-1)^{d-2} + (-1)^{d-1} - ((n+1)s-1)^{d-1}}{s^2} \\ &= \frac{(d-2)((n+1)s-1)^{d-1} + (d-1)((n+1)s-1)^{d-2} + (-1)^{d-1}}{s^2}. \end{aligned}$$

The derivative of the numerator  $g'(s)s - g(s)$  reads as

$$\frac{d}{ds}(g'(s)s - g(s)) = g''(s)s = (d-1)(d-2)(n+1)^2((n+1)s-1)^{d-3}s,$$

having a single zero at  $s=0$  and a  $(d-3)$ -fold zero at  $s = \frac{1}{n+1}$ . Hence, if  $d = 2k+1$  is odd,  $k \geq 1$ ,  $g'(s)s - g(s)$  is positive if  $s > 0$ , so  $p$  is strictly increasing on  $[0, \infty)$ . If  $d = 2k$  is even,  $k \geq 2$ ,  $g''(s)$  is negative on  $(0, \frac{1}{n+1}]$  and positive on  $(\frac{1}{n+1}, \infty)$ . Therefore, in view of

$$g'\left(\frac{1}{n+1}\right)\frac{1}{n+1} - g\left(\frac{1}{n+1}\right) = -1$$

and

$$g'\left(\frac{2}{n+1}\right)\frac{2}{n+1} - g\left(\frac{2}{n+1}\right) = 2d - 4 > 0,$$

there exists a  $s^* \in (\frac{1}{n+1}, \frac{2}{n+1})$  such that  $p$  is strictly decreasing on  $[0, s^*]$  and strictly increasing on  $[s^*, \infty)$ . Moreover,

$$g'\left(\frac{1}{n}\right)\frac{1}{n} - g\left(\frac{1}{n}\right) = \frac{(d-1)(n+1) - 1}{n^{d-1}} - 1$$

is nonpositive, and negative if and only if  $(n, d) = (2, 4)$ . This can be seen as follows: Setting

$$c_{n,d} := \frac{(d-1)(n+1) - 1}{n^{d-1}}, \quad n \geq 2, \quad d \geq 4,$$

we observe that

$$c_{n,4} = \frac{3n+2}{n^3} = \frac{3 + \frac{2}{n}}{n^2} \leq \frac{4}{n^2} \leq 1,$$

with equality if and only if  $n = 2$ , and

$$\frac{c_{n,d+1}}{c_{n,d}} = \frac{d(n+1)-1}{n((d-1)(n+1)-1)} \leq \frac{dn+d-1}{dn+d-1+n} < 1, \quad n \geq 2, \quad d \geq 4.$$

Therefore, if  $n \geq 2$  and  $d = 2k \geq 4$  is even, we have  $\frac{1}{n} \leq s^*$ , with equality if and only if  $(n, d) = (2, 4)$ . To (6). The  $m$ -variate polynomial

$$f(\vec{s}) := g\left(1 - \sum_{k=1}^m s_k\right) + \sum_{k=1}^m g(s_k), \quad \vec{s} = (s_1, \dots, s_m),$$

is convex as a sum of  $m+1$  convex functions.  $f$  is strictly convex because if  $\vec{s}, \vec{w} \in \mathbb{R}^m$  with  $\vec{s} \neq \vec{w}$ , we have  $s_k \neq w_k$  for at least one  $1 \leq k \leq m$ , so that for each  $0 < \lambda < 1$ , the strict convexity of  $g$  implies that

$$\begin{aligned} & f(\lambda\vec{s} + (1-\lambda)\vec{w}) \\ &= g\left(1 - \sum_{k=1}^m (\lambda s_k + (1-\lambda)w_k)\right) + \sum_{k=1}^m g(\lambda s_k + (1-\lambda)w_k) \\ &= g\left(\lambda\left(1 - \sum_{k=1}^m s_k\right) + (1-\lambda)\left(1 - \sum_{k=1}^m w_k\right)\right) + \sum_{k=1}^m g(\lambda s_k + (1-\lambda)w_k) \\ &< \lambda g\left(1 - \sum_{k=1}^m s_k\right) + (1-\lambda)g\left(1 - \sum_{k=1}^m w_k\right) + \sum_{k=1}^m (\lambda g(s_k) + (1-\lambda)g(w_k)) \\ &= \lambda f(\vec{s}) + (1-\lambda)f(\vec{w}). \end{aligned}$$

Further,  $f$  is bounded from below because  $g$  is, and the minimality condition

$$\vec{0} = \nabla f(\vec{s}^*) = \left(g'(s_k^*) - g'\left(1 - \sum_{j=1}^m s_j^*\right)\right)_{1 \leq k \leq m}$$

together with the injectivity of  $g'$  imply that  $s_k^* = \frac{1}{m+1}$  for all  $1 \leq k \leq m$ , and hence  $f(\vec{s}^*) = (m+1)g(\frac{1}{m+1})$ .  $\square$

By means of Lemma 50, we are now able to enumerate all solutions of the system in Eq. (7.20), i.e., all zeros  $\vec{s} := (s_1, \dots, s_{n-1})$  of the vector function

$$\vec{h}(\vec{s}) := \left(s_k \left(g\left(1 - \sum_{j=1}^{n-1} s_j\right) + \sum_{j=1}^{n-1} g(s_j)\right) - g(s_k)\right)_{1 \leq k \leq n-1} \quad (7.24)$$

in the  $(n-1)$ -dimensional unit simplex  $\Delta_{n-1} = \text{conv}\{\vec{0}, \vec{e}_1, \dots, \vec{e}_{n-1}\}$  from Eq. (7.23).

**Proposition 51.** *Let  $d \geq 2$ , and let  $\vec{h}$  be defined as in Eq. (7.24).*

1. *The function  $\vec{h}$  vanishes at least at those  $\vec{s} \in \Delta_{n-1}$  such that with  $s_n := 1 - \sum_{k=1}^{n-1} s_k$ , there exists a nonempty subset  $K \subseteq \{1, \dots, n\}$  and*

$$s_k = \begin{cases} \frac{1}{|K|}, & k \in K, \\ 0, & k \in \{1, \dots, n\} \setminus K. \end{cases} \quad (7.25)$$



2. If  $d = 2$ ,  $\vec{h}$  is identically zero.
3. If  $d$  is odd, there are no further zeros of  $\vec{h}$  in  $\Delta_{n-1}$  than those from (1).
4. If  $d \geq 4$  is even,  $\vec{h}$  vanishes at  $\vec{s} \in \Delta_{n-1}$  if and only if with  $s^* \in [\frac{1}{n}, \frac{2}{n+1})$  from Lemma 50(5) and  $s_n := 1 - \sum_{k=1}^{n-1} s_k$ , there exist disjoint subsets  $K_1 \subset \{1, \dots, n\}$  and  $K_2 \subset \{1, \dots, n\}$ , at least one of these being nonempty, such that either

$$K_1 = \emptyset, \quad s_k = \begin{cases} \frac{1}{|K_2|} > s^*, & k \in K_2, \\ 0, & k \in \{1, \dots, n\} \setminus K_2, \end{cases} \quad (7.26)$$

or

$$K_2 = \emptyset, \quad s_k = \begin{cases} \frac{1}{|K_1|} \leq s^*, & k \in K_1, \\ 0, & k \in \{1, \dots, n\} \setminus K_1, \end{cases} \quad (7.27)$$

or

$$K_1, K_2 \neq \emptyset, \quad s_k = \begin{cases} s_{k_1}, & k \in K_1, \\ s_{k_2}, & k \in K_2, \\ 0, & k \in \{1, \dots, n\} \setminus (K_1 \cup K_2), \end{cases} \quad (7.28)$$

where  $s_{k_1} \in (0, s^*)$  is a zero of the polynomial

$$r(s) := \frac{g(s)}{s} - \frac{|K_2|g\left(\frac{1-|K_1|s}{|K_2|}\right)}{1-|K_1|s} \quad (7.29)$$

and

$$s_{k_2} = \frac{1-|K_1|s_{k_1}}{|K_2|} \quad (7.30)$$

is contained in  $(s^*, 1]$ .

*Proof.* To prove (1), let  $\emptyset \neq K \subseteq \{1, \dots, n\}$ , and let  $s_k \in [0, 1]$  be given as in Eq. (7.25). Then we have

$$s_k \geq 0, \quad \sum_{k=1}^n s_k = 1, \quad \sum_{k=1}^{n-1} s_k = 1 - s_n \leq 1,$$

i.e.,  $\vec{s} := (s_1, \dots, s_{n-1}) \in \Delta_{n-1}$ . By using that  $g(0) = 0$ , see Lemma 50 (1), we compute that for all  $1 \leq k \leq n$ , regardless of whether  $k \in K$  or  $k \notin K$ ,

$$s_k \sum_{j=1}^n g(s_j) = s_k \sum_{j \in K} g(s_j) = s_k |K| g\left(\frac{1}{|K|}\right) = g(s_k),$$

so that  $\vec{h}(\vec{s}) = \vec{0}$ . For the proof of (2), assume that  $d = 2$ . In this case, we have  $g(s) = (n+1)s$ . For each  $\vec{s} = (s_1, \dots, s_{n-1}) \in \Delta_{n-1}$ , we obtain that with  $s_n := 1 - \sum_{j=1}^{n-1} s_j$ ,

$$s_k \sum_{j=1}^n g(s_j) - g(s_k) = s_k (n+1) \sum_{j=1}^n s_j - (n+1)s_k = 0,$$

so that  $\vec{h}$  vanishes identically. To (3). Suppose that  $d$  is odd and that  $\vec{s} \in \Delta_{n-1}$  is a zero of  $\vec{h}$ . Then with  $s_n := 1 - \sum_{k=1}^{n-1} s_k$ , the set  $K := \{1 \leq k \leq n : s_k \neq 0\}$  is nonempty, and  $g(0) = 0$  yields

$$\frac{g(s_k)}{s_k} = \sum_{j=1}^n g(s_j) = \sum_{j \in K} g(s_j), \quad k \in K. \quad (7.31)$$

The left-hand side of Eq. (7.31) is strictly increasing in  $s_k$ , and the right-hand side is strictly convex in  $(s_j)_{j \in K}$  with lower bound  $|K|g(\frac{1}{|K|})$ , see Lemma 50 (5)/(4). Therefore, we obtain that

$$s_k \geq \frac{1}{|K|}, \quad k \in K,$$

which, in view of  $\sum_{k \in K} s_k = 1$ , is only achievable if  $s_k$  is of the form given in Eq. (7.25). As to (4), if  $d \geq 4$  is even and  $\vec{s} \in \Delta_{n-1}$  is a zero of  $\vec{h}$ , like in (3), with  $s_n := 1 - \sum_{k=1}^{n-1} s_k$  and  $K := \{1 \leq k \leq n : s_k \neq 0\} \neq \emptyset$ , we have

$$\frac{g(s_k)}{s_k} = \sum_{j=1}^n g(s_j) = \sum_{j \in K} g(s_j), \quad k \in K.$$

Lemma 50 (5) tells us that for some  $s^* \in [\frac{1}{n}, \frac{2}{n+1})$ , the polynomial  $p(s) := \frac{g(s)}{s}$  of degree  $d-2$  is strictly decreasing on  $[0, s^*]$  and strictly increasing on  $[s^*, \infty)$ . Let us split  $K$  into

$$K = K_1 \cup K_2, \quad K_1 := \{k \in K : s_k \leq s^*\}, \quad K_2 := \{k \in K : s_k > s^*\}.$$

At least one of the subsets  $K_1$  and  $K_2$  is nonempty, because  $K$  is. We consider the three possible special cases separately.

- If  $K_1$  is empty, we obtain that by the injectivity of  $p$  on  $(s^*, \infty)$ , there exists  $k_2 \in K_2$  such that  $s_k = s_{k_2}$  for all  $k \in K_2$ . We obtain that

$$\frac{g(s_{k_2})}{s_{k_2}} = \sum_{j \in K} g(s_j) = |K_2|g(s_{k_2})$$

and hence

$$s_k = \frac{1}{|K_2|}, \quad k \in K_2,$$

after dividing both sides by  $g(s_{k_2}) > 0$ , which is the situation in Eq. (7.26).

- If  $K_2$  is empty, we can argue in an analogous way: the injectivity of  $p$  on  $[0, s^*]$  implies the existence of some  $k_1 \in K_1 \neq \emptyset$  with  $s_k = s_{k_1}$  for all  $k \in K_1$  and thus

$$s_k = \frac{1}{|K_1|}, \quad k \in K_1,$$

which is the situation in Eq. (7.27).

- Finally, assume that both  $K_1$  and  $K_2$  are nonempty. As in the previous special cases, the injectivity of  $p$  on  $[0, s^*]$  and on  $(s^*, \infty)$  implies the existence of certain  $k_1 \in K_1$  and  $k_2 \in K_2$  with  $s_k = s_{k_1}$  for all  $k \in K_1$  and  $s_k = s_{k_2}$  for all  $k \in K_2$ , such that

$$\frac{g(s_{k_1})}{s_{k_1}} = \frac{g(s_{k_2})}{s_{k_2}} = |K_1|g(s_{k_1}) + |K_2|g(s_{k_2}).$$

We have  $s_{k_1} \neq s^*$  because of

$$\frac{g(s_{k_1})}{s_{k_1}} = \frac{g(s_{k_2})}{s_{k_2}} > \frac{g(s^*)}{s^*}.$$

By using that  $|K_1|s_{k_1} + |K_2|s_{k_2} = \sum_{j=1}^n s_j = 1$ , we observe that  $s_{k_1} \in (0, s^*)$  is a zero of the even-degree polynomial

$$r(s) := p(s) - p\left(\frac{1 - |K_1|s}{|K_2|}\right) = \frac{g(s)}{s} - \frac{|K_2|g\left(\frac{1 - |K_1|s}{|K_2|}\right)}{1 - |K_1|s}$$

from Eq. (7.29), and we have

$$s_{k_2} = \frac{1 - |K_1|s_{k_1}}{|K_2|},$$

showing Eq. (7.28) and Eq. (7.30).

Conversely, assume that arbitrary disjoint subsets  $K_1, K_2 \subset \{1, \dots, n\}$  are given, with  $K_1 \neq \emptyset$  or  $K_2 \neq \emptyset$ . If  $K_1 = \emptyset$  or if  $K_2 = \emptyset$ , setting  $s_k \in [0, 1]$  as in Eq. (7.26) or in Eq. (7.27) and following the proof of part (1) with  $K$  replaced by  $K_1$  or by  $K_2$ , respectively, we see that  $\vec{h}(\vec{s}) = \vec{0}$ . If  $K_1$  and  $K_2$  are nonempty, and if  $s_{k_1} \in (0, s^*)$  is an arbitrary zero of the polynomial  $r$  from Eq. (7.29), such that  $s_{k_2}$  from Eq. (7.30) is contained in  $(s^*, 1]$ , we see that  $s_k$  from Eq. (7.28) fulfills

$$s_k \geq 0, \quad \sum_{k=1}^n s_k = |K_1|s_{k_1} + |K_2|s_{k_2} = 1, \quad \sum_{k=1}^{n-1} s_k = 1 - s_n \leq 1,$$

i.e.,  $\vec{s} := (s_1, \dots, s_{n-1}) \in \Delta_{n-1}$ . Moreover, regardless of whether  $k \in K_1$ ,  $k \in K_2$  or  $k \notin K_1 \cup K_2$ , we have

$$s_k \sum_{j=1}^n g(s_j) = s_k (|K_1|g(s_{k_1}) + |K_2|g(s_{k_2})) = g(s_k),$$

so that  $\vec{h}(\vec{s}) = \vec{0}$ . □

**Remark 52.** In case that  $d \geq 3$  is odd, the zeros of  $\vec{h}$  given by Eq. (7.25) are precisely the midpoints of the unit simplex  $\Delta_{n-1}$  and of all its lower-dimensional facets, see also Fig. 7.1. In particular, if  $n = 2$ , we obtain that  $h$  vanishes exactly at

$$s_1 \in \{0, \frac{1}{2}, 1\}.$$

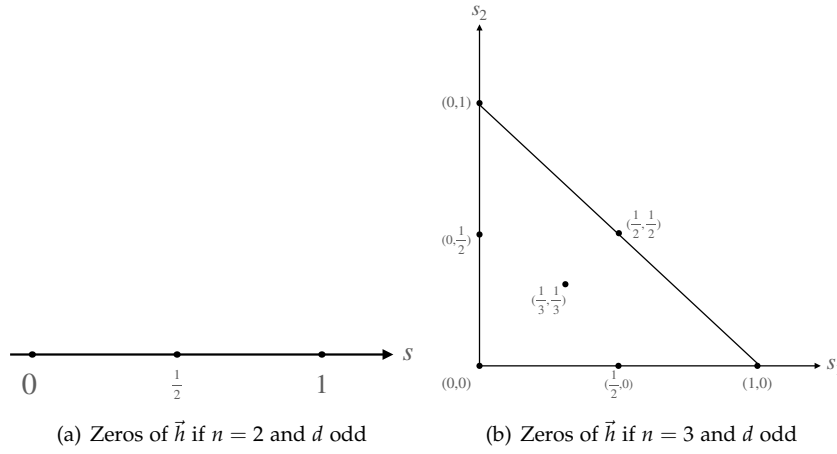


Figure 7.1: Geometric interpretation of the zero set of  $\vec{h}$  for  $d$  odd. The figure is taken from Ref. [H].

If  $n = 3$ ,  $\vec{h}$  vanishes exactly at

$$(s_1, s_2) \in \{(0,0), (1,0), (0,1), (\frac{1}{2}, 0), (0, \frac{1}{2}), (\frac{1}{2}, \frac{1}{2}), (\frac{1}{3}, \frac{1}{3})\}.$$

If  $d \geq 4$  is even, the situations in Eq. (7.26), Eq. (7.27) or Eq. (7.28) can only occur if the respective conditions  $\frac{1}{|K_2|} > s^*$ ,  $\frac{1}{|K_1|} \leq s^*$  or  $s_{k_1} < s^* < s_{k_2}$  are fulfilled.

**Example 53.** If  $n = 3$  and  $d = 4$ , we compute that

$$p(s) = \frac{g(s)}{s} = \frac{(4s-1)^3 + 1}{s} = 64s^2 - 48s + 12$$

is strictly convex with a unique global minimum at  $s^* = \frac{3}{8} < \frac{1}{2}$ . Therefore, the situation in Eq. (7.26) occurs if and only if  $K_1 = \emptyset$  and  $|K_2| \in \{1, 2\}$  and therefore  $K_2 \in \{\{1\}, \{2\}, \{3\}, \{1, 2\}, \{1, 3\}, \{2, 3\}\}$ , which corresponds to the subset of zeros

$$\{(1,0), (0,1), (0,0), (\frac{1}{2}, \frac{1}{2}), (\frac{1}{2}, 0), (0, \frac{1}{2})\}$$

of

$$\begin{aligned} \vec{h}(\vec{s}) &= \begin{pmatrix} s_1(g(s_1) + g(s_2) + g(1-s_1-s_2)) - g(s_1) \\ s_2(g(s_1) + g(s_2) + g(1-s_1-s_2)) - g(s_2) \end{pmatrix} \\ &= \begin{pmatrix} -192s_1^3s_2 - 192s_1^2s_2^2 + 32s_1^3 + 288s_1^2s_2 + 96s_1s_2^2 - 48s_1^2 - 96s_1s_2 + 16s_1 \\ -192s_1s_2^3 - 192s_1^2s_2^2 + 32s_2^3 + 288s_1s_2^2 + 96s_1^2s_2 - 48s_2^2 - 96s_1s_2 + 16s_2 \end{pmatrix}. \end{aligned}$$

The constraint  $\frac{1}{|K_1|} \leq s^* = \frac{3}{8}$  in situation (7.27) can only be fulfilled if  $K_1 = \{1, 2, 3\}$  and  $K_2 = \emptyset$ , which corresponds to the zero  $(\frac{1}{3}, \frac{1}{3})$  of  $\vec{h}$ . Finally, as it concerns the situation in

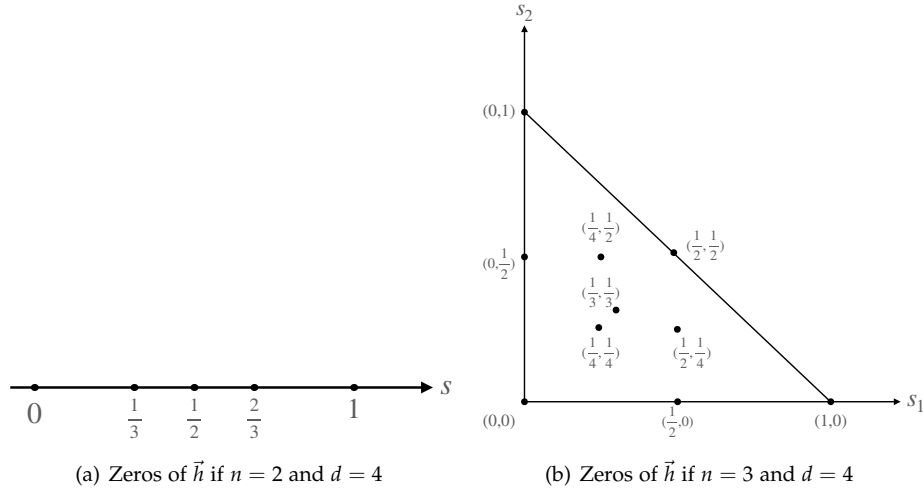


Figure 7.2: Geometric interpretation of the zero set of  $\vec{h}$  for  $d = 4$ . The figure is taken from Ref. [H].

Eq. (7.28), the only three possible configurations are  $(|K_1|, |K_2|) \in \{(1, 1), (1, 2), (2, 1)\}$ . If  $|K_1| = |K_2| = 1$ , the polynomial  $r$  from Eq. (7.29) reads as

$$r(s) = \frac{g(s)}{s} - \frac{g(1-s)}{1-s} = 32s - 16,$$

having the unique zero  $s = \frac{1}{2} > s^*$ , so that this case cannot occur. If  $|K_1| = 1$  and  $|K_2| = 2$ , we obtain

$$r(s) = \frac{g(s)}{s} - \frac{2g(\frac{1-s}{2})}{1-s} = 48s^2 - 40s + 8,$$

having the zeros  $\frac{1}{3} < s^*$  and  $\frac{1}{2} > s^*$ . But since  $s := \frac{1}{3}$  would correspond to  $\frac{1-s}{2} = \frac{1}{3}$  which is not strictly larger than  $s^*$ , this case cannot occur either. If  $|K_1| = 2$  and  $|K_2| = 1$ , we obtain

$$r(s) = \frac{g(s)}{s} - \frac{g(1-2s)}{1-2s} = -192s^2 + 112s - 16,$$

having the zeros  $\frac{1}{3}$  and  $\frac{1}{4}$ , both less than  $s^*$ . The first zero  $s := \frac{1}{3}$  of  $r$  would correspond to  $1 - 2s = \frac{1}{3}$ , which is not strictly larger than  $s^*$ , which again is not allowed. The second zero  $s_{k_1} := \frac{1}{4}$  of  $r$  yields the corresponding argument  $s_{k_2} := 1 - 2s_{k_1} = \frac{1}{2}$  and hence induces the remaining zeros

$$\left\{ \left( \frac{1}{4}, \frac{1}{4} \right), \left( \frac{1}{4}, \frac{1}{2} \right), \left( \frac{1}{2}, \frac{1}{4} \right) \right\}$$

of  $\vec{h}$ , see also Fig. 7.2.

We are now in the position to enumerate all normalized eigenpairs of simplex tensor  $\mathcal{T}$  given by Eq. (7.12) in the generic case  $n, d \geq 2$ .

**Theorem 54.** Let  $n, d \geq 2$ , and let  $\mathcal{T} = \sum_{k=1}^{n+1} \vec{v}_k^{\otimes d}$  according to Eq. (7.12).

1. If  $d = 2$ , each  $\vec{v} \in \mathbb{R}^n$  with  $\|\vec{v}\| = 1$  is an eigenvector of  $\mathcal{T}$ , with positive eigenvalue

$$\mu = 1 + \frac{1}{n}. \quad (7.32)$$

2. If  $d \geq 3$  is odd,  $\vec{v} \in \mathbb{R}^n$  with  $\|\vec{v}\| = 1$  is an eigenvector of  $\mathcal{T}$  if and only if there exists a nonempty subset  $K \subset \{1, \dots, n+1\}$  of cardinality at most  $n$ , such that

$$\vec{v} = \frac{\sum_{k \in K} \vec{v}_k}{\left\| \sum_{k \in K} \vec{v}_k \right\|} = \frac{\sum_{k \in K} \vec{v}_k}{\sqrt{\frac{|K|(n+1-|K|)}{n}}}, \quad (7.33)$$

and the corresponding eigenvalue is given by

$$\mu = \frac{(n+1-|K|)^{d-1} - |K|^{d-1}}{n^{d/2}(|K|(n+1-|K|))^{d/2-1}}. \quad (7.34)$$

The eigenvalue 0 corresponds to the case  $2|K| = n+1$ . Therefore, if  $n$  is even, all eigenvalues are different from 0.

3. If  $d \geq 4$  is even,  $\vec{v} \in \mathbb{R}^n$  with  $\|\vec{v}\| = 1$  is an eigenvector of  $\mathcal{T}$  if and only if one of the following two conditions is met: Either there exists a nonempty subset  $K \subset \{1, \dots, n+1\}$  of cardinality at most  $n$ , such that  $\vec{v}$  has the form (7.33), with positive eigenvalue

$$\mu = \frac{(n+1-|K|)^{d-1} + |K|^{d-1}}{n^{d/2}(|K|(n+1-|K|))^{d/2-1}} \quad (7.35)$$

or there exist nonempty, disjoint subsets  $K_1 \subset \{1, \dots, n+1\}$  and  $K_2 \subset \{1, \dots, n+1\}$ , each of cardinality at most  $n$ , such that with  $s^* \in [\frac{1}{n}, \frac{2}{n+1})$  from Lemma 50 (5) and  $0 < s_{k_1} \leq s^* < s_{k_2} \leq 1$  with  $|K_1|s_{k_1} + |K_2|s_{k_2} = 1$  and  $\frac{g(s_{k_1})}{s_{k_1}} = \frac{g(s_{k_2})}{s_{k_2}}$ , we have

$$\vec{v} = \frac{s_{k_1} \sum_{k \in K_1} \vec{v}_k + s_{k_2} \sum_{k \in K_2} \vec{v}_k}{\left\| s_{k_1} \sum_{k \in K_1} \vec{v}_k + s_{k_2} \sum_{k \in K_2} \vec{v}_k \right\|} = \frac{s_{k_1} \sum_{k \in K_1} \vec{v}_k + s_{k_2} \sum_{k \in K_2} \vec{v}_k}{\sqrt{\frac{(n+1)s_{k_1}^2 |K_1| + (n+1)s_{k_2}^2 |K_2| - 1}{n}}} \quad (7.36)$$

with positive eigenvalue

$$\mu = \frac{|K_1|((n+1)s_{k_1} - 1)^d + |K_2|((n+1)s_{k_2} - 1)^d + n + 1 - |K_1| - |K_2|}{n^{d/2}((n+1)s_{k_1}^2 |K_1| + (n+1)s_{k_2}^2 |K_2| - 1)^{d/2}}. \quad (7.37)$$

*Proof.* To prove (1), observe that by Eq. (7.11) we have  $\mathcal{T} = (1 + \frac{1}{n})\mathbf{1}$  for  $d = 2$ . For the proof of (2), assume that  $d \geq 3$  is odd. The first identity in Eq. (7.33) follows by an application of Lemma 47 and Proposition 51 (3). The second identity in Eq. (7.33) can be verified by using Eq. (7.11), which yields

$$\left\| \sum_{k \in K} \vec{v}_k \right\|^2 = \sum_{k \in K} \sum_{j \in K} \langle \vec{v}_k, \vec{v}_j \rangle = \sum_{k \in K} \left( 1 - \frac{|K| - 1}{n} \right) = \frac{|K|(n+1-|K|)}{n}.$$

As to the corresponding eigenvalue  $\mu$  of  $\vec{v} = \frac{\vec{z}}{\|\vec{z}\|}$ , where  $\vec{z} := \sum_{k \in K} \vec{v}_k$ , we can use that by the oddity of  $d$ ,

$$\begin{aligned} \langle \mathcal{T} \cdot \vec{z}^{\otimes(d-1)}, \vec{z} \rangle &= \sum_{k \in K} \left\langle \sum_{j \in K} \vec{v}_j, \vec{v}_k \right\rangle^d + \sum_{\substack{1 \leq k \leq n+1 \\ k \notin K}} \left\langle \sum_{j \in K} \vec{v}_j, \vec{v}_k \right\rangle^d \\ &= \sum_{k \in K} \left(1 - \frac{|K| - 1}{n}\right)^d + \sum_{\substack{1 \leq k \leq n+1 \\ k \notin K}} \left(-\frac{|K|}{n}\right)^d \\ &= \frac{|K|(n+1-|K|)}{n^d} \left( (n+1-|K|)^{d-1} - |K|^{d-1} \right), \end{aligned}$$

which yields that the eigenvalue of  $\vec{z}$  is

$$\frac{\langle \mathcal{T} \cdot \vec{z}^{\otimes(d-1)}, \vec{z} \rangle}{\|\vec{z}\|^2} = \frac{(n+1-|K|)^{d-1} - |K|^{d-1}}{n^{d-1}},$$

from which we can deduce Eq. (7.34) by an application of Eq. (7.5). To (3). If  $d \geq 4$  is even, by invoking Lemma 47, the first family of normalized eigenvectors given by Eq. (7.33) corresponds to the situations in Eq. (7.26) and Eq. (7.27) from Proposition 51. By using that  $d$  is even, similar to the reasoning in (2), we can compute that with  $\vec{z} := \sum_{k \in K} \vec{v}_k$ , we have

$$\langle \mathcal{T} \cdot \vec{z}^{\otimes(d-1)}, \vec{z} \rangle = \frac{|K|(n+1-|K|)}{n^d} \left( (n+1-|K|)^{d-1} + |K|^{d-1} \right),$$

which yields Eq. (7.35) after normalization. The second family of normalized eigenvectors given by Eq. (7.36) corresponds to situation in Eq. (7.28), and the second identity in Eq. (7.36) follows via a similar argument as in (2), by setting

$$\vec{z} := s_{k_1} \sum_{k \in K_1} \vec{v}_k + s_{k_2} \sum_{k \in K_2} \vec{v}_k$$

and by using Eq. (7.11),  $K_1 \cap K_2 = \emptyset$  and  $s_{k_1}|K_1| + s_{k_2}|K_2| = 1$ , that

$$\begin{aligned} \|\vec{z}\|^2 &= s_{k_1}^2 \left\| \sum_{k \in K_1} \vec{v}_k \right\|^2 + 2s_{k_1}s_{k_2} \left\langle \sum_{k \in K_1} \vec{v}_k, \sum_{j \in K_2} \vec{v}_j \right\rangle + s_{k_2}^2 \left\| \sum_{k \in K_2} \vec{v}_k \right\|^2 \\ &= s_{k_1}^2 \frac{|K_1|(n+1-|K_1|)}{n} - 2s_{k_1}s_{k_2} \frac{|K_1||K_2|}{n} + s_{k_2}^2 \frac{|K_2|(n+1-|K_2|)}{n} \\ &= \frac{1}{n} \left( (n+1)s_{k_1}^2|K_1| + (n+1)s_{k_2}^2|K_2| - 1 \right). \end{aligned}$$

As it concerns the corresponding eigenvalue  $\mu$  of the vector  $\vec{v} = \frac{\vec{z}}{\|\vec{z}\|}$ , we combine that

$s_{k_1}|K_1| + s_{k_2}|K_2| = 1$ ,  $K_1 \cap K_2 = \emptyset$  and that  $d$  is even, which yields the positive number

$$\begin{aligned}
\langle \mathcal{T} \cdot \vec{z}^{\otimes(d-1)}, \vec{z} \rangle &= \sum_{k \in K_1} \langle \vec{z}, \vec{v}_k \rangle^d + \sum_{k \in K_2} \langle \vec{z}, \vec{v}_k \rangle^d + \sum_{\substack{1 \leq k \leq n+1 \\ k \notin K_1 \cup K_2}} \langle \vec{z}, \vec{v}_k \rangle^d \\
&= \sum_{k \in K_1} \left( s_{k_1} \left( 1 - \frac{|K_1| - 1}{n} \right) - s_{k_2} \frac{|K_2|}{n} \right)^d \\
&\quad + \sum_{k \in K_2} \left( s_{k_2} \left( 1 - \frac{|K_2| - 1}{n} \right) - s_{k_1} \frac{|K_1|}{n} \right)^d \\
&\quad + \sum_{\substack{1 \leq k \leq n+1 \\ k \notin K_1 \cup K_2}} \left( -\frac{s_{k_1}|K_1| + s_{k_2}|K_2|}{n} \right)^d \\
&= \frac{|K_1|((n+1)s_{k_1} - 1)^d + |K_2|((n+1)s_{k_2} - 1)^d + n + 1 - |K_1| - |K_2|}{n^d},
\end{aligned}$$

and thus Eq. (7.37) after normalization.  $\square$

## 7.4 Eigenstructure for local dimension $n = 2$

In case that the local dimension  $n$  is equal to 2, the results from the previous generic analysis concretise as follows. The single barycentric coordinate  $s \in [0, 1]$  of an eigenvector  $\vec{v} = s\vec{v}_1 + (1-s)\vec{v}_2$  has to solve the nonlinear equation Eq. (7.21). We will therefore enumerate all real zeros of the expression

$$h(s) := (1-s)g(s) - sg(1-s) = g(s) - s(g(s) + g(1-s)), \quad s \in \mathbb{R}, \quad (7.38)$$

in the following proposition.

**Proposition 55.** *Let  $d \geq 2$ . Then  $h$  from Eq. (7.38) is antisymmetric with respect to  $s = \frac{1}{2}$ , i.e.,*

$$h(s) = -h(1-s), \quad s \in \mathbb{R}. \quad (7.39)$$

Moreover, depending on the parity of  $d$ ,  $h$  has the following properties.

- (1) *If  $d$  is even,  $h$  vanishes at least at  $s \in \{0, \frac{1}{3}, \frac{1}{2}, \frac{2}{3}, 1\}$ . In particular,  $h$  is the zero polynomial if  $d \in \{2, 4\}$ , and for each even  $d \geq 2$ , we can factorize  $h(s)$  into*

$$h(s) = -9s(s-1)(2s-1)(3s-1)(3s-2) \sum_{\substack{p, q \geq 0 \\ p+q \leq d/2-3}} (3s-1)^{2p}(3s-2)^{2q}, \quad (7.40)$$

*so that if  $d \geq 6$  is even,  $h$  does not have further real zeros than  $\{0, \frac{1}{3}, \frac{1}{2}, \frac{2}{3}, 1\}$ .*

- (2) *If  $d$  is odd,  $h$  vanishes at  $s \in \{0, \frac{1}{2}, 1\}$ , and we have  $h(s) < 0$  for  $0 < s < \frac{1}{2}$  and  $h(s) > 0$  for  $\frac{1}{2} < s < 1$ .*



*Proof.*  $h$  obviously fulfills Eq. (7.39), which yields  $h(\frac{1}{2}) = 0$ . Moreover, for each  $d \geq 2$  we have  $\varphi(0) = 0$  by Lemma 50 (1), which yields  $h(0) = -h(1) = 0$ . To prove (1), let  $d \geq 2$  be even. Then Lemma 50 (1) implies that  $g(\frac{2}{3}) = 2g(\frac{1}{3})$ , so that  $h(\frac{1}{3}) = -h(\frac{2}{3}) = 0$ . Therefore,  $h$  vanishes at  $s \in \{0, \frac{1}{3}, \frac{1}{2}, \frac{2}{3}, 1\}$ . If  $d \in \{2, 4\}$ , by the fundamental theorem of algebra,  $h$  vanishes identically. We will then prove Eq. (7.40) by induction over even dimensions  $d \geq 2$ . To this end, let us define

$$h_k(s) := (1-s)((3s-1)^{2k-1} + 1) - s((2-3s)^{2k-1} + 1), \quad k = 1, 2, \dots,$$

which is equal to  $h(s)$  from Eq. (7.38) for  $d = 2k$ . We already know from the previous reasoning that  $h_1$  and  $h_2$  vanish identically. Assume now that for some  $k \geq 2$ , the factorization

$$h_k(s) = -9s(s-1)(2s-1)(3s-1)(3s-2) \sum_{\substack{p, q \geq 0 \\ p+q \leq k-3}} (3s-1)^{2p}(3s-2)^{2q}$$

holds true. Then we can compute that

$$\begin{aligned} h_{k+1} - h_k(s) &= (1-s)((3s-1)^{2k+1} - (3s-1)^{2k-1}) - s((2-3s)^{2k+1} - (2-3s)^{2k-1}) \\ &= (1-s)(3s-1)^{2k-1}((3s-1)^2 - 1) - s(2-3s)^{2k-1}((2-3s)^2 - 1) \\ &= -3s(s-1)(3s-1)(3s-2)((3s-1)^{2(k-1)} - (3s-2)^{2(k-1)}) \\ &= -9s(s-1)(3s-1)(3s-2)(2s-1) \sum_{p=0}^{k-2} (3s-1)^{2p}(3s-2)^{2(k-2-p)}, \end{aligned}$$

which yields the desired factorization of  $h_{k+1}$ , thereby proving Eq. (7.40) for all even  $d \geq 2$ . If  $d \geq 6$  is even, the trailing sum in Eq. (7.40) is nonempty and positive, so that  $h$  does not have other real zeros than  $\{0, \frac{1}{3}, \frac{1}{2}, \frac{2}{3}, 1\}$ . For the proof of (2) assume  $d = 2k + 1$  for  $k \geq 1$ . Then Lemma 50(4) tells us that  $g$  is strictly convex. Therefore,

$$g(s) + g(1-s) \geq 2g(\frac{1}{2}s + \frac{1}{2}(1-s)) = 2g(\frac{1}{2}) = 2^{1-2k} - 2, \quad s \in \mathbb{R},$$

with equality if and only if  $s = \frac{1}{2}$ . On the other hand, by Lemma 50 (5), we know that  $s \mapsto \frac{g(s)}{s}$  is strictly increasing on  $(0, \infty)$ , so that

$$\frac{g(s)}{s} \leq \frac{g(\frac{1}{2})}{\frac{1}{2}} = 2g(\frac{1}{2}) = 2^{1-2k} - 2, \quad 0 < s \leq \frac{1}{2}.$$

By combining both estimates, we obtain that

$$g(s) + g(1-s) \geq \frac{g(s)}{s}, \quad 0 < s \leq \frac{1}{2},$$

with equality only if  $s = \frac{1}{2}$ , thereby proving the claim because of the antisymmetry of  $h$ .  $\square$

We are now in the position to enumerate all normalized eigenpairs of  $\mathcal{T}$  in the case  $n = 2$ .

**Theorem 56.** Let  $d \geq 2$  and  $n = 2$ , and let  $\mathcal{T} = \sum_{k=1}^3 \bar{v}_k^{\otimes d}$  according to Eq. (7.12).

(1) If  $d \in \{2, 4\}$ , each  $\bar{v} \in \mathbb{R}^2 \setminus \{0\}$  with  $\|\bar{v}\|_2 = 1$  is an eigenvector of  $\mathcal{T}$ , with positive eigenvalue

$$\mu = \begin{cases} \frac{3}{2}, & d = 2, \\ \frac{9}{8}, & d = 4. \end{cases} \quad (7.41)$$

(2) If  $d \geq 6$  is even, there are exactly 12 normalized eigenpairs  $(\bar{v}, \mu)$  of  $\mathcal{T}$ , given by

$$\{(\pm \bar{v}_k, 1 + 2^{1-d}) : 1 \leq k \leq 3\} \cup \{(\bar{v}_k + 2\bar{v}_j)/\sqrt{3}, 3^{d/2}2^{1-d}\} : 1 \leq k \neq j \leq 3\}, \quad (7.42)$$

and the corresponding eigenvalues are positive.

(3) If  $d \geq 3$  is odd, there are exactly 6 normalized eigenpairs  $(\bar{v}, \mu)$  of  $\mathcal{T}$ , given by

$$\{(\pm \bar{v}_k, \pm(1 - 2^{1-d})) : 1 \leq k \leq 3\}. \quad (7.43)$$

*Proof.* First observe that Proposition 55 yields all eigenvectors from the convex hull  $\{s\bar{v}_1 + (1-s)\bar{v}_2 : 0 \leq s \leq 1\}$  of  $\bar{v}_1$  and  $\bar{v}_2$ , which can then be mapped to normalized eigenvectors within the conical hull  $\{s\bar{v}_1 + t\bar{v}_2 : s, t \geq 0\}$  by means of Eq. (7.5). All eigenvectors in the other three sectors of  $\mathbb{R}^2$  are then found by rotation, i.e., by cyclic shifts of  $\bar{v}_1, \bar{v}_2, \bar{v}_3$ . As to (1), if  $d \in \{2, 4\}$ , Proposition 55 (1) shows that each  $\bar{v} \in \mathbb{R}^2 \setminus \{0\}$  is an eigenvector. If  $d = 2$ , we can represent each  $\bar{v} \in \mathbb{R}^2 \setminus \{0\}$  with  $\|\bar{v}\|_2 = 1$  as  $\bar{v} = \alpha\bar{v}_1 + \beta\bar{v}_2$  with  $\alpha^2 - \alpha\beta + \beta^2 = 1$ . By using that  $\langle \bar{v}_j, \bar{v}_k \rangle$  is equal to 1 if  $j = k$  and equal to  $-\frac{1}{2}$  if  $j \neq k$ , and by using  $\bar{v}_1 + \bar{v}_2 + \bar{v}_3 = \vec{0}$ , we can compute

$$\begin{aligned} \mathcal{T} \cdot \bar{v} &= \sum_{k=1}^3 \langle \bar{v}, \bar{v}_k \rangle \bar{v}_k = \left(\alpha - \frac{\beta}{2}\right)\bar{v}_1 + \left(\beta - \frac{\alpha}{2}\right)\bar{v}_2 + \left(-\frac{\alpha}{2} - \frac{\beta}{2}\right)\bar{v}_3 \\ &= \frac{3}{2}\alpha\bar{v}_1 + \frac{3}{2}\beta\bar{v}_2 = \frac{3}{2}\bar{v}. \end{aligned}$$

If  $d = 4$ , a similar computation yields

$$\begin{aligned} \mathcal{T} \cdot \bar{v}^{\otimes 3} &= \sum_{k=1}^3 \langle \bar{v}, \bar{v}_k \rangle^3 \bar{v}_k = \left(\alpha - \frac{\beta}{2}\right)^3 \bar{v}_1 + \left(\beta - \frac{\alpha}{2}\right)^3 \bar{v}_2 + \left(-\frac{\alpha}{2} - \frac{\beta}{2}\right)^3 \bar{v}_3 \\ &= \frac{9}{8}(\alpha^3 - \alpha^2\beta + \alpha\beta^2)\bar{v}_1 + \frac{9}{8}(\beta^3 - \beta^2\alpha + \beta\alpha^2)\bar{v}_2 = \frac{9}{8}\bar{v}. \end{aligned}$$

For the proof of (2), let  $d \geq 6$  even. Then Proposition 55 (1) tells us that all normalized eigenvectors from the conical hull of  $\bar{v}_1$  and  $\bar{v}_2$  are given by

$$\bar{v}_1, \quad \frac{\bar{v}_1 + 2\bar{v}_2}{\sqrt{3}}, \quad \bar{v}_1 + \bar{v}_2 = -\bar{v}_3, \quad \frac{2\bar{v}_1 + \bar{v}_2}{\sqrt{3}}, \quad \bar{v}_2.$$

The corresponding eigenvalues can be computed by using Eq. (7.19) and Eq. (7.5), they read as

$$1 + 2^{1-d}, \quad 3^{d/2}2^{1-d}, \quad 1 + 2^{1-d}, \quad 3^{d/2}2^{1-d}, \quad 1 + 2^{1-d},$$

respectively. In order to prove (3), assume that  $d \geq 3$  is odd. We can proceed in a similar way as in (2). By Proposition 55 (2), the normalized eigenvectors from the conical hull of  $\vec{v}_1$  and  $\vec{v}_2$  are given by  $\vec{v}_1, \vec{v}_1 + \vec{v}_2 = -\vec{v}_3, \vec{v}_2$ , and the corresponding eigenvalues

$$1 - 2^{1-d}, \quad -1 + 2^{1-d}, \quad 1 - 2^{1-d},$$

respectively, can be inferred from Eq. (7.19) and Eq. (7.5).  $\square$

**Corollary 57.** *If  $n = 2$  and  $d \geq 3$  is odd,  $\mathcal{T} = \sum_{k=1}^3 \vec{v}_k^{\otimes d}$  is not odeco.*

*Proof.* If  $d \geq 3$  is odd, there is no orthogonal set  $\{\vec{w}_1, \vec{w}_2\} \subset \mathbb{R}^2$  of eigenvectors, so that  $\mathcal{T}$  cannot be orthogonally decomposable.  $\square$

## 7.5 Eigenstructure for local dimension $n = 3$

If  $n = 3$ , the eigenstructure of the simplex tensor  $\mathcal{T} = \sum_{k=1}^4 \vec{v}_k^{\otimes d}$  is slightly more complicated as in the previous case  $n = 2$ , but still amenable to a concrete analysis. From Proposition 51 and Theorem 54, we can deduce the following theorem.

**Theorem 58.** *Let  $d \geq 2$  and  $n = 3$ , and let  $\mathcal{T} = \sum_{k=1}^4 \vec{v}_k^{\otimes d}$ .*

(1) *If  $d = 2$ , each  $\vec{v} \in \mathbb{R}^3$  with  $\|\vec{v}\| = 1$  is an eigenvector of  $\mathcal{T}$ , with eigenvalue*

$$\mu = \frac{4}{3}. \quad (7.44)$$

(2) *If  $d \geq 3$  is odd,  $\vec{v} \in \mathbb{R}^3$  with  $\|\vec{v}\| = 1$  is an eigenvector of  $\mathcal{T}$  if and only if there exists a nonempty subset  $K \subset \{1, \dots, 4\}$  of cardinality at most 3, such that*

$$\vec{v} = \frac{\sum_{k \in K} \vec{v}_k}{\left\| \sum_{k \in K} \vec{v}_k \right\|} = \frac{\sum_{k \in K} \vec{v}_k}{\sqrt{\frac{|K|(4-|K|)}{3}}}, \quad (7.45)$$

and the corresponding eigenvalue is given by

$$\mu = \frac{(4 - |K|)^{d-1} - |K|^{d-1}}{3^{d/2}(|K|(4 - |K|))^{d/2-1}}. \quad (7.46)$$

We have  $\mu = 0$  if and only if  $|K| = 2$ , with three linearly independent eigenvectors

$$\sqrt{\frac{3}{4}}(\vec{v}_1 + \vec{v}_2), \quad \sqrt{\frac{3}{4}}(\vec{v}_1 + \vec{v}_3), \quad \sqrt{\frac{3}{4}}(\vec{v}_1 + \vec{v}_4). \quad (7.47)$$

(3) If  $d \geq 4$  is even,  $\vec{v} \in \mathbb{R}^3$  with  $\|\vec{v}\| = 1$  is an eigenvector of  $\mathcal{T}$  if and only if one of the following two conditions is met: Either there exists a nonempty subset  $K \subset \{1, \dots, 4\}$  of cardinality at most 3, such that  $\vec{v}$  has the form of Eq. (7.45), with positive eigenvalue

$$\mu = \frac{(4 - |K|)^{d-1} + |K|^{d-1}}{3^{d/2}(|K|(4 - |K|))^{d/2-1}}. \quad (7.48)$$

or there exist nonempty, disjoint subsets  $K_1, K_2 \subset \{1, \dots, 4\}$ , such that  $K_1 \cup K_2$  has cardinality at most 3, and with  $s^* \in [\frac{1}{3}, \frac{1}{2})$  from Lemma 50 (5) and  $0 < s_{k_1} \leq s^* < s_{k_2} \leq 1$  with  $|K_1|s_{k_1} + |K_2|s_{k_2} = 1$  and  $\frac{g(s_{k_1})}{s_{k_1}} = \frac{g(s_{k_2})}{s_{k_2}}$ , we have

$$\vec{v} = \frac{s_{k_1} \sum_{k \in K_1} \vec{v}_k + s_{k_2} \sum_{k \in K_2} \vec{v}_k}{\left\| s_{k_1} \sum_{k \in K_1} \vec{v}_k + s_{k_2} \sum_{k \in K_2} \vec{v}_k \right\|} = \frac{s_{k_1} \sum_{k \in K_1} \vec{v}_k + s_{k_2} \sum_{k \in K_2} \vec{v}_k}{\sqrt{\frac{4s_{k_1}^2 |K_1| + 4s_{k_2}^2 |K_2| - 1}{3}}} \quad (7.49)$$

with positive eigenvalue

$$\mu = \frac{|K_1|(4s_{k_1} - 1)^d + |K_2|(4s_{k_2} - 1)^d + 4 - |K_1| - |K_2|}{3^{d/2}(4s_{k_1}^2 |K_1| + 4s_{k_2}^2 |K_2| - 1)^{d/2}}. \quad (7.50)$$

*Proof.* Claim (1) directly follows from Theorem 54 (1), and we have  $\mathcal{T} = \frac{4}{3}\mathbf{1}$ . For the proof of (2), observe that Eq. (7.45) and Eq. (7.46) directly follow from part (2) of Theorem 54. We have  $\mu = 0$  if and only if  $4 - |K| = |K|$ , i.e.,  $|K| = 2$ . This yields  $\binom{4}{2} = 6$  possibilities for  $K$ ,  $K \in \{\{1, 2\}, \{1, 3\}, \{1, 4\}, \{2, 3\}, \{2, 4\}, \{3, 4\}\}$ , but due to  $\sum_{k=1}^4 \vec{v}_k = \vec{0}$ , there are three collinear pairs

$$(\vec{v}_1 + \vec{v}_2, \vec{v}_3 + \vec{v}_4), \quad (\vec{v}_1 + \vec{v}_3, \vec{v}_2 + \vec{v}_4), \quad (\vec{v}_1 + \vec{v}_4, \vec{v}_2 + \vec{v}_3)$$

of eigenvectors, from which we can deduce the linearly independent set in Eq. (7.47) of normalized eigenvectors with zero eigenvalue. Claim (3) is part of Theorem 54 (3).  $\square$

**Example 59.** If  $n = 3$  and  $d = 4$ , Theorem 58 parametrizes two families of normalized eigenpairs. On the one hand, the  $2^4 - 2 = 14$  nonempty subsets

$$K \in \{\{1\}, \{2\}, \{3\}, \{4\}, \{1, 2\}, \{1, 3\}, \{1, 4\}, \{2, 3\}, \\ \{2, 4\}, \{3, 4\}, \{1, 2, 3\}, \{1, 2, 4\}, \{1, 3, 4\}, \{2, 3, 4\}\}$$

of  $\{1, \dots, 4\}$  with at most 3 elements induce the normalized eigenvectors

$$\vec{v} \in \{\vec{v}_1, \vec{v}_2, \vec{v}_3, \vec{v}_4, \sqrt{\frac{3}{4}}(\vec{v}_1 + \vec{v}_2), \sqrt{\frac{3}{4}}(\vec{v}_1 + \vec{v}_3), \sqrt{\frac{3}{4}}(\vec{v}_1 + \vec{v}_4), \sqrt{\frac{3}{4}}(\vec{v}_2 + \vec{v}_3), \sqrt{\frac{3}{4}}(\vec{v}_2 + \vec{v}_4), \\ \sqrt{\frac{3}{4}}(\vec{v}_3 + \vec{v}_4), \vec{v}_1 + \vec{v}_2 + \vec{v}_3, \vec{v}_1 + \vec{v}_2 + \vec{v}_4, \vec{v}_1 + \vec{v}_3 + \vec{v}_4, \vec{v}_2 + \vec{v}_3 + \vec{v}_4\}$$

from Eq. (7.45), which by means of  $\sum_{k=1}^4 \vec{v}_k = \vec{0}$  can be grouped into 7 pairs of collinear eigenvectors. The respective eigenvalues read as  $\mu \in \{\frac{28}{27}, \frac{4}{9}, \frac{28}{27}\}$ , occurring with multiplicity 4, 6, 4,

respectively. On the other hand, a second family of normalized eigenpairs is given by pairs of nonempty, disjoint subsets  $K_1, K_2 \subset \{1, \dots, 4\}$  with  $|K_1| + |K_2| \leq 3$  and the constraints listed in Theorem 58 (3). There are 3 possibilities for the cardinalities  $|K_j|$  of  $K_j$ , namely

$$(|K_1|, |K_2|) \in \{(1, 1), (1, 2), (2, 1)\}.$$

By retracing Example 53, we see that only the case  $|K_1| = 2$  and  $|K_2| = 1$  and the corresponding zeros

$$\left(\frac{1}{4}, \frac{1}{4}\right), \quad \left(\frac{1}{4}, \frac{1}{2}\right), \quad \left(\frac{1}{2}, \frac{1}{4}\right)$$

of  $\vec{h}$  from Eq. (7.24) induce further normalized eigenvectors. Up to a normalization of  $\sqrt{\frac{3}{8}}$  they read

$$\begin{aligned} \vec{v} \in \{ & 2\vec{v}_1 + \vec{v}_2 + \vec{v}_3, 2\vec{v}_1 + \vec{v}_2 + \vec{v}_4, 2\vec{v}_1 + \vec{v}_3 + \vec{v}_4, 2\vec{v}_2 + \vec{v}_1 + \vec{v}_3, 2\vec{v}_2 + \vec{v}_1 + \vec{v}_4, \\ & 2\vec{v}_2 + \vec{v}_3 + \vec{v}_4, 2\vec{v}_3 + \vec{v}_1 + \vec{v}_2, 2\vec{v}_3 + \vec{v}_1 + \vec{v}_4, 2\vec{v}_3 + \vec{v}_2 + \vec{v}_4, 2\vec{v}_4 + \vec{v}_1 + \vec{v}_2, \\ & 2\vec{v}_4 + \vec{v}_1 + \vec{v}_3, 2\vec{v}_4 + \vec{v}_2 + \vec{v}_3\}, \end{aligned}$$

which can be grouped into 6 pairs of collinear eigenvectors, with eigenvalue  $\mu = \frac{8}{9}$  each.

## 7.6 Robustness analysis

In this section we study the robustness of all normalized eigenvectors of a regular simplex tensor  $\mathcal{T} = \sum_{k=1}^{n+1} \vec{v}_k^{\otimes d}$  with respect to the tensor power iteration mapping  $\varphi$  from Eq. (7.6), at least if  $n = 2$  and  $n = 3$ .

In order to assess the attractivity of a given fixed point  $\vec{x} \in \mathbb{R}^n$  of the differentiable mapping  $\varphi$ , we will determine the spectral radius of the Jacobian of  $\varphi$  at  $\vec{x}$ . We will check whether the Jacobian of  $\varphi$  at  $\vec{x}$  has spectral radius strictly less than one, which indicates local contractivity of  $\varphi$ , or strictly greater than one, which indicates local expansivity at least in one direction.

It is well-known, see also [338], that the Jacobian of  $\varphi$  at  $\vec{x} \in \mathbb{R}^n$  is given by

$$\varphi'(\vec{x}) = \frac{d-1}{\|\mathcal{T} \cdot \vec{x}^{\otimes(d-1)}\|} \left( \mathbf{1} - \frac{(\mathcal{T} \cdot \vec{x}^{\otimes(d-1)})(\mathcal{T} \cdot \vec{x}^{\otimes(d-1)})^\top}{\|\mathcal{T} \cdot \vec{x}^{\otimes(d-1)}\|^2} \right) \mathcal{T} \cdot \vec{x}^{\otimes(d-2)}, \quad (7.51)$$

where

$$\mathcal{T} \cdot \vec{x}^{\otimes(d-2)} := \left( \sum_{i_1, \dots, i_{d-2}=1}^n T_{i_1, \dots, i_{d-2}, j, k} x_{i_1} \cdots x_{i_{d-2}} x_j x_k \right)_{1 \leq j, k \leq n} \quad (7.52)$$

is the  $(d-2)$ -fold partial contraction of  $\mathcal{T}$  with  $\vec{x}^{\otimes(d-2)}$ . In the matrix case  $d = 2$  and  $\mathcal{T} = A$ , Eq. (7.51) is to be understood as

$$\varphi'(\vec{x}) = \frac{1}{\|A\vec{x}\|} \left( \mathbf{1} - \frac{A\vec{x}(A\vec{x})^\top}{\|A\vec{x}\|^2} \right) A. \quad (7.53)$$

If  $\vec{x} \in \mathbb{R}^n$  is a normalized eigenvector of  $\mathcal{T}$  with eigenvalue  $\mu$ , we obtain that

$$\varphi'(\vec{x}) = \frac{d-1}{|\mu|}(\mathbb{1} - \vec{x}\vec{x}^\top)\mathcal{T} \cdot \vec{x}^{\otimes(d-2)} = \frac{d-1}{|\mu|}(\mathcal{T} \cdot \vec{x}^{\otimes(d-2)} - \mu\vec{x}\vec{x}^\top), \quad (7.54)$$

which for  $d = 2$  and  $\mathcal{T} = A$  reduces to

$$\varphi'(\vec{x}) = \frac{1}{|\mu|}(\mathbb{1} - \vec{x}\vec{x}^\top)A. \quad (7.55)$$

Before we start with the eigenvalue analysis of  $\varphi'$  at a normalized tensor eigenvector  $\vec{x} \in \mathbb{R}^n$ , we note that by the very structure of Eq. (7.54), we have

$$\varphi'(\vec{x})\vec{x} = \vec{0}. \quad (7.56)$$

Therefore, the spectral radius  $\rho(\varphi'(\vec{x}))$  of the symmetric matrix  $\varphi'(\vec{x})$  is determined by the other eigenvectors of  $\varphi'(\vec{x})$  which are contained in the orthogonal complement  $\text{span}\{\vec{x}\}^\perp$ .

### 7.6.1 Local dimension $n = 2$

In case that the local dimension  $n$  is equal to two, we will now explicitly compute the spectral radius of the Jacobian of  $\varphi$  at each normalized eigenvector.

**Theorem 60.** *Let  $d \geq 2$ , and let  $\mathcal{T} = \sum_{k=1}^3 \vec{v}_k^{\otimes d}$  simplex tensor given by Eq. (7.12).*

- (1) *If  $d \in \{2, 4\}$ , the spectral radius of  $\varphi'$  at all normalized  $\vec{x} \in \mathbb{R}^2$  is equal to 1, i.e., there are no robust eigenvectors.*
- (2) *If  $d \geq 6$  is even, the spectral radius of  $\varphi'$  at the normalized eigenvectors  $\pm\vec{v}_k$  is equal to  $\frac{3(d-1)}{2^{d-1}+1}$ , i.e., these normalized eigenvectors are robust. The spectral radius of  $\varphi'$  at the normalized eigenvectors  $(\vec{v}_k + 2\vec{v}_j)/\sqrt{3}$ ,  $k \neq j$  is equal to  $\frac{d-1}{3}$ , i.e., those normalized eigenvectors are non-robust.*
- (3) *If  $d \geq 3$  is odd, the spectral radius of  $\varphi'$  at all normalized eigenvectors  $\pm\vec{v}_k$  is equal to  $\frac{3(d-1)}{2^{d-1}-1}$ , i.e., all normalized eigenvectors are non-robust for  $d = 3$ , and robust for  $d \geq 5$ .*

*Proof.* We begin with the proof of (1). If  $d = 2$ , we have  $\mathcal{T} = \sum_{k=1}^3 \vec{v}_k^{\otimes 2} = \frac{3}{2}\mathbb{1}$ . Therefore, each  $\vec{x} \in \mathbb{R}^2 \setminus \{\vec{0}\}$  with  $\|\vec{x}\| = 1$  is an eigenvector of  $\mathcal{T}$  with eigenvalue  $\mu = \frac{3}{2}$ . By Eq. (7.53), we obtain

$$\varphi'(\vec{x}) = \frac{1}{3/2}(\mathbb{1} - \vec{x}\vec{x}^\top)\frac{3}{2}\mathbb{1} = \mathbb{1} - \vec{x}\vec{x}^\top,$$

which has eigenvalues 0 and 1, and thus spectral radius 1. If  $d = 4$ , Theorem 56 (1) tells us that each  $\vec{x} = (x_1, x_2) \in \mathbb{R}^2 \setminus \{0\}$  with  $x_1^2 + x_2^2 = 1$  is an eigenvector of  $\mathcal{T}$ ,

and hence of  $\mathcal{T} \cdot \vec{x}^{\otimes(d-2)} \in \mathbb{R}^{2 \times 2}$ , with eigenvalue  $\mu = \frac{9}{8}$ . For  $\vec{w} := (x_2, -x_1) \perp \vec{x}$ , we compute that

$$\begin{aligned} (\mathcal{T} \cdot \vec{x}^{\otimes 2})\vec{w} &= \sum_{k=1}^3 \langle \vec{x}, \vec{v}_k \rangle^2 \langle \vec{w}, \vec{v}_k \rangle \vec{v}_k \\ &= x_1^2 x_2 \vec{e}_1 + \left( -\frac{x_1}{2} + \frac{x_2 \sqrt{3}}{2} \right)^2 \left( -\frac{x_2}{2} - \frac{x_1 \sqrt{3}}{2} \right) \left( -\frac{1}{2} \vec{e}_1 + \frac{\sqrt{3}}{2} \vec{e}_2 \right) \\ &\quad + \left( -\frac{x_1}{2} - \frac{x_2 \sqrt{3}}{2} \right)^2 \left( -\frac{x_2}{2} + \frac{x_1 \sqrt{3}}{2} \right) \left( -\frac{1}{2} \vec{e}_1 - \frac{\sqrt{3}}{2} \vec{e}_2 \right) \\ &= \begin{pmatrix} \frac{3}{8} x_1^2 x_2 + \frac{3}{8} x_2^3 \\ -\frac{3}{8} x_1^3 - \frac{3}{8} x_1 x_2^2 \end{pmatrix} = \frac{3}{8} \vec{w}, \end{aligned}$$

i.e.,  $\vec{w}$  is an eigenvector of  $\mathcal{T} \cdot \vec{x}^{\otimes 2}$  with eigenvalue  $\frac{3}{8}$ . By the orthogonality between  $\vec{x}$  and  $\vec{w}$ , we obtain  $\varphi'(\vec{x})\vec{w} = \frac{8}{3}(\mathbb{1} - \vec{x}\vec{x}^\top)(\mathcal{T} \cdot \vec{x}^{\otimes 2})\vec{w} = \vec{w}$ , so that by Eq. (7.56),  $\varphi'(\vec{x})$  has eigenvalues 0 and 1, and thus spectral radius 1. For the proof of (2), let  $d \geq 6$  be even. In this case Theorem 56 (1) tells us that there are two types of normalized eigenvectors, namely  $\pm \vec{v}_k$  and  $(\vec{v}_k + 2\vec{v}_j)/\sqrt{3}$  for  $1 \leq k \neq j \leq 3$ , with eigenvalues  $\mu_1 = 1 + 2^{1-d}$  and  $\mu_2 = 3^{d/2} 2^{1-d}$ , respectively. By using that  $\sum_{k=1}^3 \vec{v}_k \vec{v}_k^\top = \frac{3}{2} \mathbb{1}$ , we get

$$\begin{aligned} \mathcal{T} \cdot (\pm \vec{v}_k)^{\otimes(d-2)} &= \sum_{j=1}^3 \langle \vec{v}_k, \vec{v}_j \rangle^{d-2} \vec{v}_j \vec{v}_j^\top = \vec{v}_k \vec{v}_k^\top + \frac{1}{2^{d-2}} \sum_{j \neq k} \vec{v}_j \vec{v}_j^\top \\ &= \left( 1 - \frac{1}{2^{d-2}} \right) \vec{v}_k \vec{v}_k^\top + \frac{3}{2^{d-1}} \mathbb{1}. \end{aligned}$$

Therefore,  $\mathcal{T} \cdot (\pm \vec{v}_k)^{\otimes(d-2)}$  has the normalized eigenvectors  $\vec{v}_k$  and  $\vec{w} \in \text{span}\{\vec{v}_k\}^\perp$  with eigenvalues  $\mu_1$  and  $\frac{3}{2^{d-1}}$ , respectively. Hence,  $\mathcal{T} \cdot (\pm \vec{v}_k)^{\otimes(d-2)}$  and  $\mathbb{1} - \vec{v}_k \vec{v}_k^\top$  having the same eigenvectors,  $\varphi'(\pm \vec{v}_k)$  has the eigenvalues

$$0, \quad \frac{d-1}{|\mu_1|} \frac{3}{2^{d-1}} = \frac{3(d-1)}{2^{d-1} + 1}.$$

The spectral radius of  $\varphi'(\pm \vec{v}_k)$  is therefore given by

$$\rho(\varphi'(\pm \vec{v}_k)) = \frac{3(d-1)}{2^{d-1} + 1},$$

which is strictly less than 1 if and only if  $d \geq 6$ . Concerning the second eigenvector family, we proceed in an analogous way. If  $1 \leq k \neq j \leq 3$  are given, and if  $r \in \{1, 2, 3\}$  is the index with  $\{1, 2, 3\} \setminus \{k, j\} = \{r\}$ , we have

$$\langle \vec{v}_k + 2\vec{v}_j, \vec{v}_l \rangle = \begin{cases} 0, & l = k, \\ \frac{3}{2}, & l = j, \\ -\frac{3}{2}, & l = r, \end{cases}$$

which implies that

$$\begin{aligned} \mathcal{T} \cdot ((\vec{v}_k + 2\vec{v}_j)/\sqrt{3})^{\otimes(d-2)} &= 3^{1-d/2} \sum_{l=1}^3 \langle \vec{v}_k + 2\vec{v}_j, \vec{v}_l \rangle^{d-2} \vec{v}_l \vec{v}_l^\top \\ &= \frac{3^{d/2-1}}{2^{d-2}} (\vec{v}_j \vec{v}_j^\top + \vec{v}_r \vec{v}_r^\top) = \frac{3^{d/2}}{2^{d-1}} \mathbb{1} - \frac{3^{d/2-1}}{2^{d-2}} \vec{v}_k \vec{v}_k^\top. \end{aligned}$$

Therefore,  $\mathcal{T} \cdot ((\vec{v}_k + 2\vec{v}_j)/\sqrt{3})^{d-2}$  has the normalized eigenvectors  $(\vec{v}_k + 2\vec{v}_j)/\sqrt{3} \perp \vec{v}_k$  and  $\vec{v}_k$  with eigenvalues  $\mu_2$  and  $\frac{3^{d/2-1}}{2^{d-1}}$ , respectively. Hence,  $\mathcal{T} \cdot ((\vec{v}_k + 2\vec{v}_j)/\sqrt{3})^{\otimes(d-2)}$  and  $\mathbb{1} - (\vec{v}_k + 2\vec{v}_j)(\vec{v}_k + 2\vec{v}_j)^\top/3$  having the same eigenvectors,  $\varphi'((\vec{v}_k + 2\vec{v}_j)/\sqrt{3})$  has the eigenvalues

$$0, \quad \frac{d-1}{|\mu_2|} \frac{3^{d/2-1}}{2^{d-1}} = \frac{d-1}{3}.$$

The spectral radius of  $\varphi'((\vec{v}_k + 2\vec{v}_j)/\sqrt{3})$  is therefore equal to  $\frac{d-1}{3} > 1$ , so that the second family of normalized eigenvectors is not robust. It remains to prove (3). If  $d \geq 3$  is odd, Theorem 56 (2) tells us that the only normalized eigenvectors of  $\mathcal{T}$  are given by  $\pm \vec{v}_k$ , with eigenvalues  $\mu_\pm = \pm(1 - 2^{1-d})$ . By using that  $\sum_{k=1}^3 \vec{v}_k \vec{v}_k^\top = \frac{3}{2} \mathbb{1}$ , similar to the reasoning in (2), we get

$$\begin{aligned} \mathcal{T} \cdot (\pm \vec{v}_k)^{\otimes(d-2)} &= \pm \sum_{j=1}^3 \langle \vec{v}_k, \vec{v}_j \rangle^{d-2} \vec{v}_j \vec{v}_j^\top = \pm \left( \vec{v}_k \vec{v}_k^\top - \frac{1}{2^{d-2}} \sum_{j \neq k} \vec{v}_j \vec{v}_j^\top \right) \\ &= \pm \left( \left(1 + \frac{1}{2^{d-2}}\right) \vec{v}_k \vec{v}_k^\top - \frac{3}{2^{d-1}} \mathbb{1} \right). \end{aligned}$$

Therefore,  $\mathcal{T} \cdot (\pm \vec{v}_k)^{\otimes(d-2)}$  has the normalized eigenvectors  $\vec{v}_k$  and  $\vec{w} \in \text{span}\{\vec{v}_k\}^\perp$  with eigenvalues  $\mu_\pm$  and  $\mp \frac{3}{2^{d-1}}$ , respectively. Hence,  $\mathcal{T} \cdot (\pm \vec{v}_k)^{\otimes(d-2)}$  and  $\mathbb{1} - \vec{v}_k \vec{v}_k^\top$  having the same eigenvectors,  $\varphi'(\pm \vec{v}_k)$  has the eigenvalues

$$0, \quad \mp \frac{d-1}{|\mu_\pm|} \frac{3}{2^{d-1}} = \mp \frac{3(d-1)}{2^{d-1}-1}.$$

The spectral radius of  $\varphi'(\pm \vec{v}_k)$  is therefore given by

$$\rho(\varphi'(\pm \vec{v}_k)) = \frac{3(d-1)}{2^{d-1}-1},$$

which is strictly less than 1 if and only if  $d \geq 5$ .  $\square$

**Example 61.** In order to visualize the typical performance of the tensor power method in the two-dimensional case, let us proceed as follows. Each point on the unit sphere in  $\mathbb{R}^2$  is uniquely determined by its angle (i.e., its argument), and we assign a color to each possible angle. For each starting point  $\vec{v}^{(0)}$  on the unit sphere, we iterate the tensor power method given by Eq. (7.7) until convergence, and then paint  $\vec{v}^{(0)}$  with the color of its limit point. The resulting domains of attraction in the special case  $d = 7$  are visualized in Fig. 7.3, together with the spanning



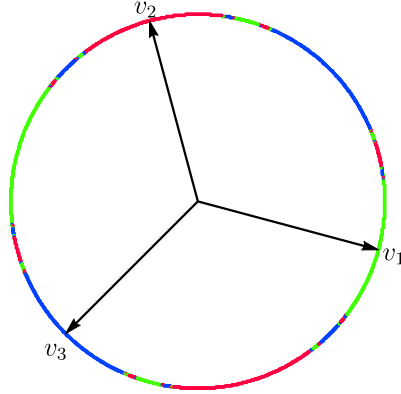


Figure 7.3: Regions of attraction of the tensor power method if  $n = 2$  and  $d = 7$ . The figure is taken from Ref. [H].

vectors  $\vec{v}_1$ ,  $\vec{v}_2$  and  $\vec{v}_3$  of the regular simplex tensor  $\mathcal{T} = \sum_{k=1}^3 \vec{v}_k^{\otimes d}$ . It can be clearly seen that there are six regions of attraction which moreover contain the vectors  $\pm\vec{v}_1, \pm\vec{v}_2, \pm\vec{v}_3$  as their respective centers. This observation is backed by Theorem 56 (3), which states that in this case  $d = 7$ , all normalized eigenvectors of  $\mathcal{T}$  are robust.

### 7.6.2 Local dimension $n = 3$

In the special case  $n = 3$ , we conduct two numerical experiments to assess the robustness of a normalized eigenvector.

The first experiment is set up in a similar way as in the two-dimensional case. We assign a unique color to each point on the unit sphere in  $\mathbb{R}^3$ , depending on the respective spherical angles. In Fig. 7.4, we visualize the results of the tensor power iteration in the cases  $d = 4$  and  $d = 5$ . In each scenario, we can observe several regions of convergence which correspond to the robust eigenvectors. The position of each eigenvector  $\vec{v}_k$  is indicated by a black dot. Each eigenvector  $\vec{v}_k$  corresponds to two regions, as  $-\vec{v}_k$  is also an eigenvector.

## 7.7 Discussion and Conclusion

In this Chapter, we have analyzed the eigenstructure of real symmetric tensors. In the special case of regular simplex tensors, where all weights  $\lambda_j$  in the symmetric decomposition in Eq. (7.1) are equal and the  $\vec{v}_j$  are induced by  $n + 1$  equiangular vectors in  $\mathbb{R}^n$ , we have seen that some normalized eigenpairs are attractive fixed points with respect to the tensor power iteration, whereas others are repelling. Therefore, as long as the tensor power iteration is used without modification, some eigenvectors

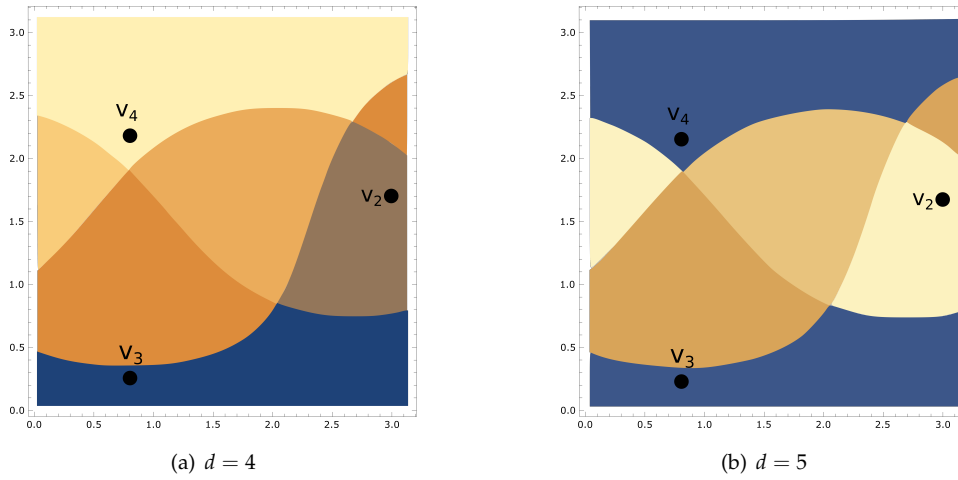


Figure 7.4: Regions of attraction of the tensor power method for  $n = 3$  and  $d \in \{4, 5\}$ . Each unit vector on upper half of the sphere is represented by its polar angle  $\varphi \in [0, \pi]$  ( $y$ -axis) and its azimuthal angle  $\vartheta \in [0, \pi]$  ( $x$ -axis). The figure is taken from Ref. [H].

may not be detectable numerically. These observations induce several directions of further research.

On the one hand, maintaining the viewpoint of the orthodox tensor power iteration, it would be a natural generalization to study symmetric tensors whose symmetric decomposition in Eq. (7.1) uses different weights  $\lambda_j$  and/or is induced by a set of more than  $n + 1$  equiangular vectors  $\vec{v}_j$  or even a generic tight frame. Here, one might aim at a complete characterization of those symmetric tensors which do have repelling eigenvectors, or whose normalized eigenvectors are given by the vectors  $\vec{v}_j$  alone. Partial answers to the latter question have been given in [338]. Furthermore, the case of complex-valued tensors seems to be open and nontrivial, because even the existence and construction of equiangular tight frames are delicate tasks in large dimensions (see Ref. [341, 342]).

On the other hand, algorithmic modifications could be employed to recover those eigenvectors that are non-robust under the orthodox tensor power iteration. Shifted tensor power iterations have been considered in [343], with good success, including a characterization of which normalized eigenpairs can and cannot be found numerically by such a scheme. However, it seems that further modifications are necessary to obtain a full numerical tensor eigenvalue solver.



# 8 Shadow tomography with generalized measurements

Recent advances in quantum technology require scalable techniques to efficiently extract classical information from a quantum system. However, traditional quantum state tomography is limited to a handful of qubits and shadow tomography has been suggested as a scalable replacement. While conventional shadow tomography is based on outcomes of ideal projective measurements, here we suggest a version formulated for generalized measurements. Based on the idea of the least-squares estimator, shadow tomography with generalized measurements is both more general and simpler compared to its original formulation. We provide a detailed study of the implication of symmetries in shadow tomography and demonstrate how the optimization of measurements for shadow tomography tailored towards a particular set of observables can be carried out. This Chapter is based on Project [E] and Project [A].

## 8.1 Motivation

Quantum technology is based on our ability to manipulate quantum mechanical states of well-isolated systems. More precisely, it is necessary to encode, to process and to extract information from the states of the system. Extracting information in this context means to design and perform measurements on the system such that observables or other properties of the system such as its entropy can be inferred. Naively, one may attempt to perform tomography of the state of the system. This amounts to making a sufficiently large number of different measurements on the system such that the density operator describing the state can be inferred [344–348]. However, when considering how the density operator is used later on, the entire information contained in the density operator is often not needed [349]. Even more, it is impractical to even write down the density operator when the number of qubits is large, since the dimension of the many-qubit system increases exponentially. In addition, the parameters appearing in the density operator cannot be directly accessed due to the probabilistic character of quantum theory and thus must be estimated by measuring the system. More precisely, as an informative quantum measurement is destructive, many identically prepared samples are required to estimate accurately even a single parameter of

the state. In practice, one is typically interested in certain properties of the quantum state, such as the mean values of certain observables or its entropy. Indeed, the prediction of many expectation values of a quantum state is indispensable for variational quantum algorithms [350,351] such as the variational quantum eigensolver [352]. Aiming at directly inferring the observables, bypassing the reconstruction of the density operator, shadow tomography has been theoretically proposed [349]. Based on this scheme, a practical procedure to realize that aim was suggested [353], which has attracted a lot of attention in the research of quantum information processing. The idea of the protocol is simple. Traditionally, quantum state tomography is thought to be only useful given accurate enough statistics of the measurements. However, state estimators such as the least-squares estimator can actually be carried out in principle for arbitrary diluted data [225], a fact well established in data science and machine learning [306,354]. Indeed, a single data point can contribute a noisy estimate of the state and the final estimated state is obtained by averaging over all the data points. As one would expect, when the data are diluted, the estimated quantum state can be highly noisy and far away from the targeted actual state in the high-dimensional state space. This noisy estimation is, however, sufficient to predict certain observables or properties of the quantum states accurately [349,353]. Crucially, estimation of observables and certain properties of the quantum states for single data points can also be processed without writing down the density operator explicitly [353]. In order to collect the data, it was suggested to perform random unitaries from a certain chosen set of unitaries on the system and perform a standard ideal measurement afterwards. Clearly, this is equivalent to randomly choosing a measurement from a predefined set. Various applications of this technique have been found in energy estimation [355,356], entanglement detection [357,358], metrology [359], analyzing scrambled data [360] and quantum chaos [361]. Further developments to improve the performance of the scheme [362–366] and generalization to channel shadow tomography have also been proposed [367,368].

In this Chapter we propose a general framework for shadow tomography with generalized measurements. In Section 8.2 we formulate our more general scheme and demonstrate that it contains the theoretical framework of the randomization of unitaries as a special case. So far, there is a single proposed procedure for shadow tomography with generalized measurements [369]. The suggested protocol in Ref. [369] is, however, based on an application of the original construction of classical shadows upon manually synthesizing the postmeasurement states for generalized measurements. On the contrary, we show that classical shadows for generalized measurements can be derived from the least-squares estimator, which requires no further assumptions on the post-measurement states. In Section 8.3 we provide a detailed study of the implication of symmetries in shadow tomography. We proceed in Section 8.4 by demonstrating how the optimization of measurements for shadow tomography tai-

lored towards a particular set of observables can be carried out. Afterwards, we prove in Section 8.5 that the so-called octahedron measurement is optimal if the targeted observables are all the projections on arbitrary pure states of a qubit. Finally, we show in Section 8.6 how one can take the effect of noise in the measurements into account.

## 8.2 Formulating shadow tomography with POVMs

In this Section we will introduce a formulation of shadow tomography which is based on generalized measurements, thus extending the original proposal in Ref. [353]. Further we will demonstrate its connection to the least-squares estimator and explain how expectation values of measurements can be obtained.

### 8.2.1 Shadow tomography with generalized measurements

Consider a quantum system of dimension  $D$ , which can be either a single qudit or many qubits,  $D = 2^n$ . Suppose that a generalized measurement  $E = (E_1, \dots, E_N)$  with  $N \in \mathbb{N}$  outcomes is performed on the system. By virtue of the Born rule, each generalized measurement defines a map  $\Phi_E$ , which maps a density operator  $\rho \in \mathcal{B}(\mathbb{C}^D)$  to a probability distribution over the set of measurement outcomes. More precisely,

$$\Phi_E : M_D(\mathbb{C}) \rightarrow \mathbb{R}^N, \quad \rho \mapsto \Phi_E(\rho) = (\text{Tr}[\rho E_j])_{j=1}^N. \quad (8.1)$$

When the measurement  $E$  is performed on the quantum system, an outcome  $j \in \{1, \dots, N\}$  is obtained according to this distribution. Typical measurements in quantum mechanics are generalized measurements whose effects  $E_j$  are rank-1 projections, referred to as ideal measurements, see also Section 1.1.2. For example, the measurement of the Pauli operator  $\sigma_1 = X$  is an ideal measurement, whose effects are projections on the spin states in the  $x$  direction, that is,  $\{|x^+\rangle\langle x^+|, |x^-\rangle\langle x^-|\}$ . On the other hand, randomizing three Pauli measurements  $\sigma_1, \sigma_2, \sigma_3$  is equivalent to a generalized measurement with effects proportional to the projections on the spin states in the  $x, y$  and  $z$  direction, i.e.,  $(1/3)\{|x^\pm\rangle\langle x^\pm|, |y^\pm\rangle\langle y^\pm|, |z^\pm\rangle\langle z^\pm|\}$ . Since these effects form an octahedron on the Bloch sphere, we refer to them as the octahedron measurement.

### 8.2.2 Shadows from the least-squares estimator

An unbiased linear estimator for a quantum state  $\rho$  is a map  $\chi : \mathbb{R}^N \rightarrow M_D(\mathbb{C})$  which maps a probability distribution to a quantum state such that  $(\chi \circ \Phi_E) : M_D(\mathbb{C}) \rightarrow M_D(\mathbb{C})$  acts as the identity on  $M_D(\mathbb{C})$ . In particular, this means that as long as the statistics  $\Phi_E(\rho)$  of a quantum state  $\rho$  with respect to  $E$  can be exactly measured,  $\chi$  allows for an exact construction of the underlying quantum state  $\rho$ . In the case that the generalized measurement  $E = (E_j)_{j=1}^N$  forms a basis for the operator space  $M_D(\mathbb{C})$ , that

is,  $E$  is informationally complete but not overcomplete, then  $\Phi_E$  is invertible and one can choose  $\chi = \Phi_E^{-1}$  as an unbiased linear estimator. In the case that the generalized measurement  $E$  is overcomplete, i.e., the effects span the operator space but it is not a minimal generating set, the map  $\chi$  is not uniquely defined. More precisely, in this case  $\Phi_E$  is not surjective, such that there exist probability distributions in  $\mathbb{R}^N$  that give rise to no state  $\varrho$ . In this situation, it is a natural choice to use the least-squares estimator as a replacement for the inverse. For a distribution  $\vec{p} \in \mathbb{R}^N$  the assigned state is given by

$$\chi_{\text{LS}}(\vec{p}) = \arg \min_{\tau} L(\tau), \quad \text{where} \quad L(\tau) = \sum_{j=1}^N (\text{Tr}[\tau E_j] - p_j)^2. \quad (8.2)$$

In order to carry out the minimization in Eq. (8.2), we consider the expansion of  $\delta L = L(\tau + \delta\tau) - L(\tau)$  in the first order of the small variation  $\delta\tau$ . This yields

$$\delta L = 2 \sum_{j=1}^N (\text{Tr}[\tau E_j] - p_j) \text{Tr}[E_j \delta\tau]. \quad (8.3)$$

A necessary condition for the minimality of  $L$  is that  $\delta L = 0$  for all choices of  $\delta\tau$ . Consequently, it must then hold that

$$\sum_{j=1}^N (\text{Tr}[\tau E_j] - p_j) E_j = 0. \quad (8.4)$$

Using the definition of  $\Phi_E$  in Eq. (8.1) one can rewrite Eq. (8.4) as

$$\Phi_E^\dagger [\Phi_E(\tau) - \vec{p}] = 0 \quad \Rightarrow \quad \chi_{\text{LS}} = (\Phi_E^\dagger \Phi_E)^{-1} \Phi_E^\dagger, \quad (8.5)$$

where  $\Phi_E^\dagger : \mathbb{R}^N \rightarrow \text{M}_D(\mathbb{C})$  is the adjoint map of  $\Phi_E$ . Note that in the case of  $\Phi_E$  invertible, one has  $\chi_{\text{LS}} = \Phi_E^{-1}$ . Remarkably this estimator is linear, which implies that the estimated state over the whole data set can be split into the sum of the estimated states for the single data points. Indeed, a single observation of an outcome  $k$  can be associated with an elementary statistics vector denoted by  $\vec{q}_k := (\delta_{kj})_{j=1}^N \in \mathbb{R}^N$ . Note that in the following we always label deterministic distribution by  $\vec{q}$  and general distributions by  $\vec{p}$ . Repeating the measurement  $M$  times on the system results in a string of outcomes  $\{k_j\}_{j=1}^M$ . The statistics of the whole data set of a string of outcomes is given by

$$\vec{p} = \frac{1}{M} \sum_{j=1}^M \vec{q}_{k_j} \quad \Rightarrow \quad \chi_{\text{LS}}(\vec{p}) = \frac{1}{M} \sum_{j=1}^M \chi_{\text{LS}}(\vec{q}_{k_j}). \quad (8.6)$$

Further, one finds that  $\Phi_E^\dagger(\vec{q}_k) = E_k$  and if we define the *measurement channel* as  $C_E(\varrho) = \Phi_E^\dagger \Phi_E \varrho$  one obtains that  $C_E(\varrho) = \sum_{j=1}^N \text{Tr}[\varrho E_j] E_j$ . Following the notation in Ref. [353], given a single data point  $\vec{q}_k$  one can use this point to obtain a noisy estimate

of  $\rho$ , which is called *classical shadow* and is given by  $\hat{\rho}_k := \chi(\vec{q}_k)$ . Given observation  $k$ , the classical shadow can be written as

$$\hat{\rho}_k = \chi_{\text{LS}}(\vec{q}_k) = [(\Phi_E^\dagger \Phi_E)^{-1}](\Phi_E^\dagger(\vec{q}_k)) = C_E^{-1}(E_k). \quad (8.7)$$

It is easy to see that for an infinite number of runs of the measurement the average of the classical shadows converges to the underlying quantum state  $\rho$ . More precisely, if  $k_j \in \{1, \dots, N\}$  denotes the measurement outcome of the  $k$ th repeat and  $f(j, M)$  is the number of observations of outcome  $j \in \{1, \dots, N\}$  given  $M$  repetitions of the measurement in total, we have

$$\frac{1}{M} \sum_{j=1}^M \hat{\rho}_{k_j} = C^{-1} \left( \sum_{j=1}^N \frac{f(j, M)}{M} E_j \right) \xrightarrow{M \rightarrow \infty} C^{-1} \left( \sum_{j=1}^N \text{Tr}[E_j \rho] E_j \right) = C^{-1}(C(\rho)) = \rho. \quad (8.8)$$

However, it is important to emphasize that the convergence in Eq. (8.8) does not guarantee that  $C_E^{-1}$  is a good estimator. More precisely, the convergence of the estimator holds true even if  $C_E$  is replaced by any other invertible channel  $D_E(\rho) = \sum_{j=1}^N \text{Tr}[\rho E_j] \tau_j$  for arbitrary operators  $\tau_j$ . This highlights that one should not associate the effect  $E_j$  with the state of the system after the measurement. That  $C_E^{-1}$  is a good estimator is supported by the fact that it is actually the least-squares estimator. It should also be noted that, despite being unbiased, in general,  $\hat{\rho}_k$  is not unit trace. The estimator  $\hat{\rho}_k$  is only unit trace if all the effects  $E_j$  that appear in the generalized measurement  $E$  share the same trace. The linearity of the estimator in Eq. (8.7) is crucial to shadow tomography. For instance, this allows one to estimate linear observables with single data points and later on average the whole data set.

### 8.2.3 Estimation of observables and sample complexities

Each of the classical shadows given by Eq. (8.7) serves as an intermediate processed data point for further computation of observables. Given an observable  $X$ , each of the classical shadows  $\hat{\rho}_k$  gives an estimate of the mean value  $\langle X \rangle$  via  $\hat{x} = \text{Tr}[\hat{\rho}_k X]$ . With the whole data set of observations  $\{k_j\}_{j=1}^M$ , we obtain from the convergence of the classical shadows in Eq. (8.8) that  $(1/M) \sum_{j=1}^M \hat{x}_{k_j}$  converges to  $\langle X \rangle$  for  $M \rightarrow \infty$ . In this way, the mean value  $\langle X \rangle$  can be estimated. For further refinement using the median-of-means estimation and estimation of polynomial functions of the density operator, see Ref. [353] and Chapter 9. In total, we end up with the following protocol for shadow tomography with a generalized measurement  $E = \{E_j\}_{j=1}^N$  for the estimation of a set of observables  $\mathfrak{X} = \{X_j\}_{j=1}^m$ .

- (1) Given the measurement  $E$ , the classical shadows  $\{\hat{\rho}_j\}_{j=1}^N$  are classically computed using Eq. (8.7) or Eq. (8.14) depending on whether or not  $E$  admits some symmetry.



- (2) The quantum system is prepared in the designed state for investigation and the measurement  $E$  is carried out. This is repeated on the system  $M$  times and the corresponding string of outcomes  $\{k_j\}_{j=1}^M$  is recorded.
- (3) The mean values of the observables in the targeted set  $\mathfrak{X}$  are estimated via  $\langle X_i \rangle \approx (1/M) \sum_{j=1}^M \text{Tr}[\hat{\rho}_{k_j} X_i]$  using the string of observations  $\{k_j\}_{j=1}^M$  obtained in the second step.

As noted in Ref. [353], the asymptotic rate of convergence of the estimation is related to the variance of the estimator. For an observable  $X$  the variance of the estimator can be computed as  $\text{Var}[\hat{x}_k] = \sum_{j=1}^N \text{Tr}[\hat{\rho}_j X]^2 \text{Tr}[\rho E_j] - \langle X \rangle^2$ . Ignoring the second term results in an upper bound for the variance, and finally assuming the worst case scenario, i.e., a maximisation over  $\rho$ , one arrives at the definition of the shadow norm of  $X$  [353]. More precisely,

$$\begin{aligned} \text{Var}[\hat{x}_k] &\leq \sum_{j=1}^N \text{Tr}[\hat{\rho}_j X]^2 \text{Tr}[\rho E_j] \leq \max_{\rho} \text{Tr} \left[ \rho \sum_{j=1}^N \text{Tr}[\hat{\rho}_j X]^2 E_j \right] \\ &= \lambda_{\max} \left[ \sum_{j=1}^N \text{Tr}[\hat{\rho}_j X]^2 E_j \right] =: \|X\|_E, \end{aligned} \quad (8.9)$$

where  $\lambda_{\max}$  denotes the maximal eigenvalue of the corresponding operator. The estimation procedure applies not only to an observable, but equally well to a set of observables  $\mathfrak{X}$ . Assuming that the observables by a certain normalisation all have the same physical unit, the quality of shadow tomography with a generalised measurement  $E$  can be characterised by the maximal shadow norm,

$$\kappa_E^2(\mathfrak{X}) := \max \{ \|X\|_E^2 : X \in \mathfrak{X} \}. \quad (8.10)$$

Being an upper bound of the variance of the estimator, the smaller  $\kappa_E^2(\mathfrak{X})$ , the better the estimator accuracy. However, in practice the targeted state could be very different from the worst case scenario assumed in obtaining the shadow norm. Therefore, it might also be informative to consider the average of the variance with respect to a certain ensemble of states.

## 8.2.4 Relation to randomized ideal measurements

As already pointed out, the often-used scheme of randomized ideal projective measurements can be considered as a special realization of the generalized measurement framework presented in Section 8.2.1. Here we will discuss how the corresponding generalized measurement can be constructed explicitly given a randomized measurement scheme. Let  $U$  be a random unitary drawn from an ensemble  $\mathcal{U}$  of unitaries, which is implemented on the physical system before making a measurement in the

computational basis  $\{|b\rangle\}_{b=1}^D$ . This scenario corresponds to a random projective measurement in the basis  $\{U^\dagger|b\rangle\}_{b=1}^D$ . The whole procedure effectively simulates a generalized measurement with effects

$$E = \left\{ \frac{1}{|\mathcal{U}|} U^\dagger |b\rangle \langle b| U : U \in \mathcal{U}, b = 1, \dots, D \right\}, \quad (8.11)$$

where  $|\mathcal{U}|$  is the total number of unitaries in  $\mathcal{U}$ . Mathematically, the resulted generalized measurement is a convex combination of the chosen ideal measurements. In general, this is a well-known method for simulating generalized measurements with ideal ones, known as preprocessing [39]. However, it should be noted that not every generalized measurement admits such a decomposition [370].

The suggested scheme of shadow tomography using a single generalized measurement brings several interesting new perspectives. Indeed, the randomized unitary description is heavily over-parametrised, i.e., having much more parameters than necessary. For instance, the unitary ensemble of the Clifford group on a qubit contains 24 elements, and yet is equivalent to a single generalized measurement with 6 outcomes corresponding to the six directions of the three standard axes. This over-parametrisation makes it difficult to keep track over the parameters and prevents optimisation. Using generalized measurements also brings a new perspective on the implementation of the procedure. A generalized measurement can be implemented by a measurement on the ancillary system after an appropriate coupling to the objective system. While this can still be challenging in practice, it avoids changing measurement settings, which requires frequent (re)calibration of the setup.

### 8.3 Symmetries and generalized measurements

It has been observed that, for certain classes of measurements, the inverse of the measurement channel  $C_E^{-1}$  is particularly simple [353, 371]. We now show that this simplicity originates from the symmetry of the corresponding generalized measurement [372–374]. More precisely, we will show how the formula for the classical shadows in Eq. (8.7) can be drawn just from consideration of symmetry. In order to get an intuition of the argument, we will first present the example of the octahedron generalized measurement over a qubit, see Fig. 8.1. Picking a vertex of the octahedron which corresponds to the effect  $E_j$ , we consider the symmetry rotations of the octahedron that leave this vertex invariant. These are rotations by multiples of  $\pi/2$  around the axis going through the chosen vertex. Noticeably, there is a single projection, together with its complement, that is invariant under these rotations. This projection is the operator corresponding to the state of the spin pointing to the vertex itself. More precisely, this means that the effect  $E_j$  is uniquely specified by the symmetry. It then follows that also the corresponding classical shadow  $\hat{Q}_j$  is invariant under these rotations, implying that

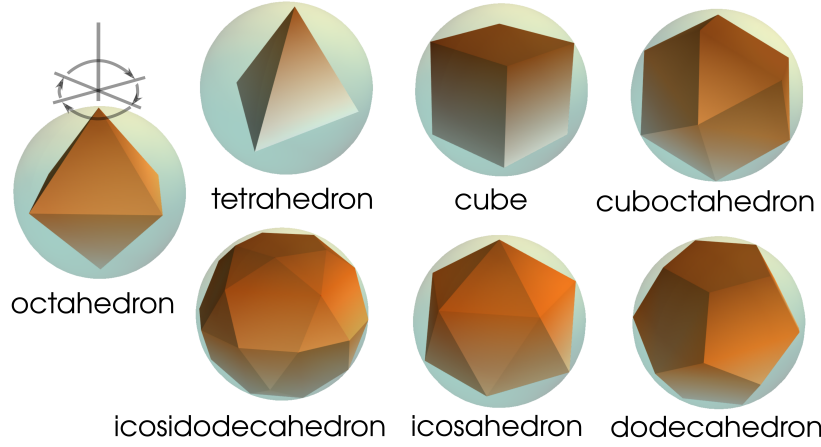


Figure 8.1: Example of generalized measurements defined by polytopes on the Bloch sphere: octahedron ( $N = 6$ ), tetrahedron ( $N = 4$ ), cube ( $N = 8$ ), cuboctahedron ( $N = 12$ ), icosahedron ( $N = 12$ ), dodecahedron ( $N = 20$ ), icosidodecahedron ( $N = 30$ ). The figure is taken from Ref. [E].

it must be a linear combination of  $E_j$  and the identity operator  $\mathbb{1}$ , i.e.,  $\hat{q}_j = aE_j + b\mathbb{1}$  for some coefficients  $a, b$ .

Let us now consider the general case of an arbitrary generalized measurement  $E = (E_1, \dots, E_N)$ , where  $E_j \in \mathcal{B}(\mathbb{C}^D)$ . The unitary group acting on the space  $\mathbb{C}^D$  will be denoted by  $U_D(\mathbb{C})$ . We say that  $E$  is symmetric [372] if there exists a subgroup  $G$  of the permutation group over  $\{1, \dots, N\}$ , also called symmetric group and denoted by  $\mathfrak{S}_N$ , and a unitary representation  $U : G \rightarrow U_D(\mathbb{C})$  such that  $E_{g(k)} = U_g E_k U_{g^{-1}}$ . If  $E$  is symmetric under  $G$ , then the associated map  $\Phi_E : M_D(\mathbb{C}) \rightarrow \mathbb{R}^N$  is covariant under the action of  $G$ , that is,  $\Phi_E(U_g \rho U_{g^{-1}}) = g[\Phi_E(\rho)]$ , where  $g$  acts on a probability distribution  $\vec{p}$  by  $[g(\vec{p})]_j = p_{g^{-1}(j)}$ . Indeed, one has

$$\begin{aligned} \Phi_E(U_g \rho U_{g^{-1}}) &= \{\text{Tr}[U_g \rho U_{g^{-1}} E_j]\}_{j=1}^N = \{\text{Tr}[\rho U_{g^{-1}} E_j U_g]\}_{j=1}^N \\ &= \{\text{Tr}[\rho E_{g^{-1}(j)}]\}_{j=1}^N = g[\Phi_E(\rho)]. \end{aligned} \quad (8.12)$$

As a consequence,  $\Phi_E^\dagger$  is also covariant under  $G$ , i.e., one has  $\Phi_E^\dagger(g(\vec{p})) = U_g \Phi_E^\dagger(\rho) U_{g^{-1}}$  for all distributions  $\vec{p} \in \mathbb{R}^N$ . This in turn implies that  $C_E = \Phi_E^\dagger \Phi_E$  and in particular its inverse  $C_E^{-1}$  are also covariant under the action of the group  $G$ ,

$$C_E^{-1}(U_g X U_{g^{-1}}) = U_g C_E^{-1}(X) U_{g^{-1}}, \quad \forall X \in \mathcal{B}(\mathbb{C}^D) \text{ hermitian.} \quad (8.13)$$

For the octahedron measurement any outcome can be related to any other by a symmetry transformation. Such measurements are called uniform, which means that  $G$  acts transitively on the outcomes. In particular, this implies that  $\text{Tr}[E_j]$  is independent of the index  $j$  and we denote  $\alpha = \text{Tr}[E_j]$ . Following the procedure introduced in

Ref. [372], we consider the stabilizer subgroup  $G_j$  over  $j$ , that is, the subgroup of  $G$  that leaves the label  $j$  invariant. It directly follows from the definition of a symmetric measurement that the corresponding effect  $E_j$  commutes with all the unitary operators of the stabilizer group  $U(G_j) := \{U_g \mid g \in G_j\}$ . Because  $C_E^{-1}$  is covariant under  $G$ , it is clear that  $\hat{q}_j = C_E^{-1}(E_j)$  also commutes with  $U(G_j)$ .

Further, we have seen that for the octahedron the only projections commuting with all of the unitary operators from the stabilizer group at a vertex are the spin projection in the direction of the vertex and its complement. In general, if for all outcomes  $j$  of the measurement the set of all operators that commute with the stabilizer  $U(G_j)$  is spanned by a single projection  $\Pi_j$  and its complement  $\mathbb{1} - \Pi_j$ , one says that  $E$  is rigidly symmetric [372]. In other words,  $E$  is rigidly symmetric if the representation  $U$  restricted to any stabilizer subgroup over  $j$ ,  $G_j$ , has exactly two irreducible representations. This is the characterization of the property that an operator commuting with  $U(G_j)$  can only be a linear combination of  $E_j$  and  $\mathbb{1}$ . Hence, if  $E$  is rigidly symmetric, then

$$\hat{q}_j = aE_j + b\mathbb{1}. \quad (8.14)$$

To compute the coefficients  $a, b$  notice that

$$C_E(\hat{q}_k) = aC_E(E_k) + bC_E(\mathbb{1}) = a \sum_{j=1}^N \text{Tr}[E_k E_j] E_j + b\alpha\mathbb{1}. \quad (8.15)$$

Further, the expression  $\hat{q}_k = C_E^{-1}(E_k)$  implies  $C_E(\hat{q}_k) = E_k$  such that the left-hand side of Eq. (8.15) can be identified with  $E_k$ . Taking the trace of Eq. (8.15) and the trace of it after multiplying the two sides with  $E_k$  one obtains

$$\alpha = a\alpha^2 + b\alpha D, \quad \beta = a\gamma + b\alpha^2, \quad (8.16)$$

where  $\beta = \text{Tr}[E_k^2]$ ,  $\gamma = \sum_{j=1}^N \text{Tr}[E_k E_j]^2$  which are both independent of  $k$  due to the uniformity. The coefficients  $a, b$  can be explicitly computed as

$$a = \frac{D\beta - \alpha^2}{D\gamma - \alpha^3}, \quad b = \frac{\gamma - \alpha\beta}{D\gamma - \alpha^3}. \quad (8.17)$$

It should be noted that a large class of uniform and rigidly symmetric measurements beyond qubits exists and is studied in Ref. [372]. The corresponding parameters of the classical shadows are presented in Tab. 8.1.

## 8.4 Optimization of measurements

Given a set of observables  $\mathfrak{X}$ , one would like to find the generalized measurement  $E$  such that the maximal shadow norm is minimized. More precisely, one aims to

$d$	ST	$N$	Comments	$a$	$b$
2	8	6	Octahedron	9	-1
		8	Cube	12	-1
		12	Cuboctahedron	18	-1
	16	12	Icosahedron	18	-1
		20	Dodecahedron	30	-1
		30	Icosidodecahedron	45	-1
3	24	21		28	-1
	25	12	csMUB	16	-1
	27	45		60	-1
		60		80	-1
4	28	12	Real MUB	9	$-\frac{1}{2}$
	29	20	csMUB	25	-1
		40		50	-1
		80		100	-1
	30	300		225	$-\frac{1}{2}$
	31	60		75	-1
		480		600	-1

Table 8.1: Parameters for the inverse of the measurement channels for symmetric generalized measurements. Here  $d$  refers to the dimension of the system, ST is the Shephard–Todd number of the corresponding symmetry group and  $N$  is the number of outcomes of the measurement. The table is taken from Ref. [D].

compute

$$E^* := \arg \min_E \kappa_E^2(\mathfrak{X}). \quad (8.18)$$

Restricted to random unitaries, this optimization is impractical to carry out. Extending to all generalized measurements, this is simply an optimization over a convex domain. For the minimization in Eq. (8.18) we implemented simulated annealing and found that the obtained optima are highly reliable. We start with the discussion of the single qubit case, which, despite being simple, is also the basis to understand the case of many qubits.

#### 8.4.1 The single qubit case

Consider a single qubit. For the observables that can be contained in the set  $\mathfrak{X}$  we consider three different possibilities. First, we take the observables which correspond to the four projections given by the orange tetrahedron in Fig.8.2(a). The squared shadow

norm  $\kappa_E^2(\mathfrak{X})$  in this case is 2 for the tetrahedron generalized measurement defined exactly by these four projections, and 3/2 for the octahedron measurement. The optimizer suggests that the tetrahedron measurement plotted in violet in Fig. 8.2(a), obtained by centrally inverting the orange tetrahedron, is optimal with  $\kappa_E^2(\mathfrak{X}) = 1$ .

Second, we consider as observables the projections onto the eigenstates of the Pauli observables, see the orange octahedron in Fig. 8.2(b). Considering as a generalized measurement  $E$  the octahedron measurement itself, one obtains  $\kappa_E^2(\mathfrak{X}) = 3/2$ . Interestingly, the optimizer shows that  $\kappa_E^2(\mathfrak{X}) = 3/2$  can also be obtained with the tetrahedron generalized measurement of four outcomes indicated in violet in Fig. 8.2(b).

Third, we consider as observables random projections distributed according to the Haar measure on the Bloch sphere. Fig. 8.2(c) presents the shadow norms obtained by the optimizer with respect to the number of observables. For a small number of observables, e.g.,  $|\mathfrak{X}| \leq 15$ , the optimizer always finds measurements with a given number of outcomes significantly better than the standard tetrahedron ( $N = 4$ ) or the octahedron ( $N = 6$ ) measurements. It is interesting to see that if the number of outcomes is fixed to be 6 or 8, the shadow norm  $\kappa_E^2(\mathfrak{X})$  converges to the octahedron measurement with a value of 3/2. This indicates that the octahedron measurement is special. As we will prove in Section 8.5, it turns out that the octahedron measurement is optimal if  $\mathfrak{X}$  is the set of all projections on arbitrary pure states of the qubit.

### 8.4.2 Tensoring construction and the multi-qubit case

Shadow tomography is especially designed for the cases where the system is large. Consider the case where the system consists of  $n$  qubits, corresponding to the total dimension of  $D = 2^n$ . In this case, shadow tomography can be performed by making possibly different generalized measurements  $\{E^{(1)}, \dots, E^{(n)}\}$  on each of the qubits, each described by a collection of  $N_j$  effects, that is,  $E^{(j)} = \{E_k^{(j)}\}_{k=1}^{N_j}$ . Theoretically, this corresponds to a measurement of a generalized measurement  $E^{\text{tot}}$  on the whole  $n$ -qubit system with each effect labeled by a string of outcomes  $\vec{k} = (k^{(1)}, \dots, k^{(n)})$  such that

$$E_{\vec{k}}^{\text{tot}} = E_{k^{(1)}}^{(1)} \otimes E_{k^{(2)}}^{(2)} \otimes \dots \otimes E_{k^{(n)}}^{(n)}. \quad (8.19)$$

We can now apply our previous analysis to the measurement defined by the effects in Eq. (8.19). In fact, such a string  $\vec{k}$  of outcomes corresponds simply to the classical shadow

$$\hat{\rho}_{\vec{k}}^{\text{tot}} = \hat{\rho}_{k^{(1)}}^{(1)} \otimes \hat{\rho}_{k^{(2)}}^{(2)} \otimes \dots \otimes \hat{\rho}_{k^{(n)}}^{(n)}, \quad (8.20)$$

where  $\hat{\rho}_{k^{(j)}}^{(j)}$  is the classical shadow corresponding to the measurement  $E^{(j)}$  on the  $j$ th qubit. Crucially, the typical observables of the system can be easily estimated without explicitly computing the classical shadows in the form of a  $D \times D$  matrix, which would

be impractical. Indeed, an observable  $X$  on the system is often of the form  $X = X^{(1)} \otimes X^{(2)} \otimes \dots \otimes X^{(n)}$ . Then, a single string of outcomes  $\vec{k}$  gives rise to a single estimate of  $\langle X \rangle$  as

$$\mathrm{Tr} \left[ \hat{\rho}_{k^{(1)}}^{(1)} X^{(1)} \right] \mathrm{Tr} \left[ \hat{\rho}_{k^{(2)}}^{(2)} X^{(2)} \right] \dots \mathrm{Tr} \left[ \hat{\rho}_{k^{(n)}}^{(n)} X^{(n)} \right]. \quad (8.21)$$

The final estimate of  $\langle X \rangle$  is as usual obtained by averaging over all data points. Observe that it is not necessary to construct the large density operator of the whole system. Moreover, the shadow norm of such a factorized observable also factorizes, i.e.,

$$\|X\|_E = \|X^{(1)}\|_{E^{(1)}} \|X^{(2)}\|_{E^{(2)}} \dots \|X^{(n)}\|_{E^{(n)}}. \quad (8.22)$$

We can also use our approach to optimize the generalized measurements for many-body systems. For many qubits, the number of parameters to be optimized in Eq. (8.18) increases exponentially. To simplify, one can assume that for a many-qubit system, the generalized measurement is factorized as a tensor product over the qubits as discussed above. Moreover, if there is no preference among the qubits, one can also assume that  $E^{(1)} = E^{(2)} = \dots = E^{(n)}$ .

The complexity of the computation under these assumptions is only linear in the number of qubits and the number of observables. As an example, we consider a system of up to  $n = 64$  qubits. We choose  $|\mathcal{X}| = n$  observables which are products of different component observables on single qubits. The component observables on single qubits are randomly distributed according to the Haar measure. As the qubits are equivalent, one might anticipate that the optimal factorizing measurement for the qubits is similar to those that are optimized separately for each qubit. The simulation confirms this expectation. In Fig. 8.2(d), for small number of qubits ( $n \leq 10$ ), the optimizer with  $N = 6$  and  $N = 8$  gives significantly lower shadow norms for the choice of tetrahedron or octahedron measurements. On the other hand, observe that as the number of qubits increases, the obtained optimal shadow norm converges to that given by the octahedron measurement. This points to the speciality of the octahedron measurement on qubit-based platforms.

## 8.5 Optimality of measurements

We have already mentioned that the octahedron measurement plays a distinguished role and we have seen in Section 8.4.1 that the squared shadow norm with respect to that measurement for any projection is identically  $3/2$ . Using this, we will now prove that if the targeted observables are *all* the projections on arbitrary pure states of the qubit, the optimal measurement would be the octahedron measurement assuming equal trace of the effects. For a pure state  $|\lambda\rangle \in \mathbb{C}^2$  we denote by  $\Pi_\lambda$  the associated projection operator, that is,  $\Pi_\lambda = |\lambda\rangle\langle\lambda|$ . We consider the problem of predicting all

the expectation values of these projections based on classical shadows generated by a measurement  $E$ . We characterize the predicting power of the shadow tomography by the maximal shadow norm among all these projections,

$$\max_{\lambda} \|\Pi_{\lambda}\|_E^2 = \max_{\lambda} \max_{\varrho} \sum_{k=1}^N \langle \lambda | \hat{\varrho}_k | \lambda \rangle^2 \text{Tr}[\varrho E_k]. \quad (8.23)$$

To show that the octahedron measurement is optimal, we assume that the measurement  $E$  has uniform trace. This is typically not a restrictive assumption, since starting with a measurement with effects of non-uniform traces, by virtually splitting each effect in an appropriate number of identical smaller effects, a measurement with uniform traces can be achieved. Indeed, suppose that  $E$  is a measurement with different traces for each effect,  $\alpha_j = \text{Tr}[E_j]$ . Then, one can always approximate  $\alpha_j$  by a rational number. Consequently, one can choose a sufficient small number  $\epsilon$  such that  $\alpha_j/\epsilon = N_j$  are all integers. Then one splits an effect  $E_j$  into  $N_j$  identical effects ( $E_j/N_j, \dots, E_j/N_j$ ) and obtains a new measurement of  $\sum_{j=1}^N N_j$  effects. By construction, this new measurement is uniform. If the measurement  $E$  is uniform, this implies that the classical shadows  $\hat{\varrho}_k$  are of unit trace. In fact, one can derive an explicit formula for the classical shadows in the Bloch representation. Recall that any hermitian operator  $\mathcal{B}(\mathbb{C}^2) \ni X$  can be identified with a real vector  $\vec{x} \in \mathbb{R}^4$  such that  $X = (1/2) \sum_{j=0}^3 x_j \sigma_j$  with  $x_j = \text{Tr}[\varrho \sigma_j]$ . Notice that for two operators  $X, Y$  represented by  $\vec{x}$  and  $\vec{y}$  respectively one has  $\text{Tr}[XY] = (1/2) \sum_{j=0}^3 x_j y_j$ . By the uniformity of the effects, each effect  $E_j$  can be represented by a vector of the form  $(2/N)(1, \vec{r}_j)$ , where  $\vec{r}_j \in \mathbb{R}^3$  with  $\|\vec{r}_j\| \leq 1$  and  $\sum_{j=1}^N \vec{r}_j = \vec{0}$ . An explicit calculation shows that

$$\chi = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ H^{-1} \vec{r}_1 & H^{-1} \vec{r}_2 & \cdots & H^{-1} \vec{r}_N \end{pmatrix} \quad \text{with} \quad H = \frac{1}{N} \sum_{j=1}^N \vec{r}_j \vec{r}_j^{\top}. \quad (8.24)$$

Here it is important to note that the columns of  $\chi$  are exactly the classical shadows  $\hat{\varrho}_k$  in the Bloch representation. This explicit formula for the classical shadow allows us to explicitly compute the shadow norm. For maximizing the shadow norm in Eq. (8.23), take the particular choice of  $\varrho = |\lambda\rangle\langle\lambda|$  as an ansatz, which yields the lower bound

$$\max_{\lambda} \|\Pi_{\lambda}\|_E^2 \geq \max_{\lambda} \sum_{j=1}^N \langle \lambda | \hat{\varrho}_j | \lambda \rangle^2 \langle \lambda | E_j | \lambda \rangle. \quad (8.25)$$

Then, by replacing the maximum over the pure states  $\lambda$  by the average with respect to the Haar measure  $\mu$  on the Bloch sphere  $\mathcal{S}$ , one obtains the further lower bound

$$\max_{\lambda} \|\Pi_{\lambda}\|_E^2 \geq \int_{\mathcal{S}} \sum_{j=1}^N \text{Tr}[(\hat{\varrho}_j \otimes \hat{\varrho}_j \otimes E_j) \Pi_{\lambda}^{\otimes 3}] d\mu(\lambda). \quad (8.26)$$

By employing Schur-Weyl duality for the unitary group [375], one can show that

$$\int_{\mathcal{S}} \Pi_{\lambda}^{\otimes 3} d\mu(\lambda) = \frac{1}{12} [P_{(1,2)} + P_{(2,3)} + P_{(3,1)}], \quad (8.27)$$



where  $P_{(1,2)}$ ,  $P_{(2,3)}$  and  $P_{(3,1)}$  denote the operators that permute the corresponding tensor terms in  $(\mathbb{C}^2)^{\otimes 3}$ . Combining Eq. (8.26) and Eq. (8.27) one arrives at

$$\max_{\lambda} \|\Pi_{\lambda}\|_E^2 \geq \frac{1}{12} \sum_{j=1}^N (\text{Tr}[\hat{\varrho}_j^2] \text{Tr}[E_j] + 2 \text{Tr}[\hat{\varrho}_j] \text{Tr}[\hat{\varrho}_j E_j]), \quad (8.28)$$

where we have used the identities

$$\begin{aligned} \text{Tr}[(\hat{\varrho}_j \otimes \hat{\varrho}_j \otimes E_j)P_{(1,2)}] &= \text{Tr}[\hat{\varrho}_j^2] \text{Tr}[E_j], \\ \text{Tr}[(\hat{\varrho}_j \otimes \hat{\varrho}_j \otimes E_j)P_{(2,3)}] &= \text{Tr}[\hat{\varrho}_j] \text{Tr}[\hat{\varrho}_j E_j], \\ \text{Tr}[(\hat{\varrho}_j \otimes \hat{\varrho}_j \otimes E_j)P_{(3,1)}] &= \text{Tr}[\hat{\varrho}_j] \text{Tr}[\hat{\varrho}_j E_j]. \end{aligned} \quad (8.29)$$

Using the explicit Bloch representation for  $E_j$  and  $\hat{\varrho}_j$ , one obtains

$$\max_{\lambda} \|\Pi_{\lambda}\|_E^2 \geq \frac{1}{12} \left[ 3 + \frac{1}{N} \sum_{j=1}^N (\vec{r}_j^{\top} H^{-2} \vec{r}_j + 2 \vec{r}_j^{\top} H^{-1} \vec{r}_j) \right]. \quad (8.30)$$

Further, it is important to note that

$$\frac{1}{N} \sum_{j=1}^N \vec{r}_j^{\top} H^{-1} \vec{r}_j = \frac{1}{N} \sum_{j=1}^N \text{Tr}[H^{-1} \vec{r}_j \vec{r}_j^{\top}] = \text{Tr}[H^{-1} H] = 3 \quad (8.31)$$

and by a similar reasoning one obtains

$$\frac{1}{N} \sum_{j=1}^N \vec{r}_j^{\top} H^{-2} \vec{r}_j = \frac{1}{N} \sum_{j=1}^N \text{Tr}[H^{-2} \vec{r}_j \vec{r}_j^{\top}] = \text{Tr}[H^{-2} H] = \text{Tr}[H^{-1}]. \quad (8.32)$$

Taking Eq. (8.30), Eq. (8.31) and Eq. (8.32) together one ends up with

$$\max_{\lambda} \|\Pi_{\lambda}\|_E^2 \geq \frac{1}{12} (9 + \text{Tr}[H^{-1}]). \quad (8.33)$$

Notice that  $H$  is positive and denote its positive eigenvalues by  $\zeta_1, \zeta_2, \zeta_3$ . Then one has  $\text{Tr}[H^{-1}] = \zeta^{-1} + \zeta^{-2} + \zeta^{-3} \geq 9/(\zeta_1 + \zeta_2 + \zeta_3) = 9/\text{Tr}[H]$ . Moreover,  $\text{Tr}[H] = (1/N) \sum_{j=1}^N \vec{r}_j^{\top} \vec{r}_j \leq 1$ . Finally we arrive at

$$\max_{\lambda} \|\Pi_{\lambda}\|_E^2 \geq \frac{3}{2}. \quad (8.34)$$

We have already seen that the inequality in Eq. (8.34) is saturated for the octahedron measurement, in which case,  $H = (1/3)\mathbb{1}$ . This demonstrates that the octahedron measurement is an optimal choice for the chosen set of observables.

## 8.6 Mitigation of measurement noise

Measurements that appear in realistic experimental setups are not ideal. The imperfection is due to various sources of noise in setting up the parameters of the measurement

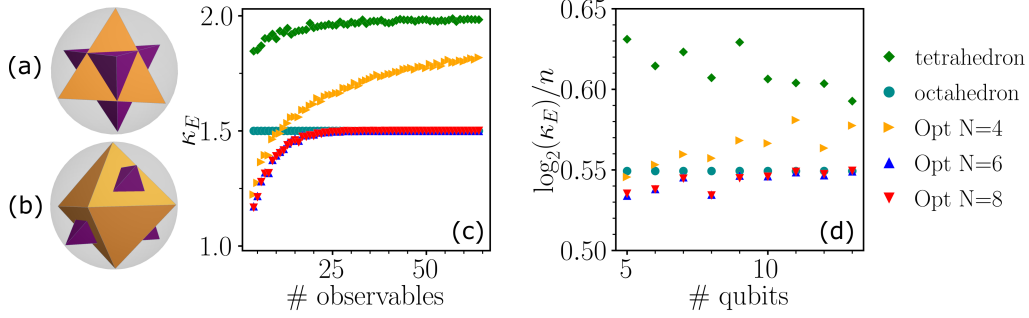


Figure 8.2: Targeted observables and optimal generalized measurements. (a) For observables corresponding to four projections defined by the orange tetrahedron, the measurement corresponding to the inverted tetrahedron measurement (violet) is optimal. (b) For observables corresponding to eigenprojections of the Pauli observables  $\sigma_1, \sigma_2$  and  $\sigma_3$  (orange octahedron), the violet tetrahedron measurement is optimal. (c) Optimal shadow norms given by the optimizer (labeled Opt with the number of measurement outcomes) as a function of the number of single-qubit projection observables randomly distributed according to the Haar measure. (d) Similarly, optimal shadow norms given by the optimizer as a function of the number of qubits. The observables are tensor products of single-qubit projections distributed according to the Haar measure. In (c) and (d), the shadow norms for the tetrahedron and octahedron measurements are also shown. The figure is taken from Ref. [E].

devices, or the resolution and the accuracy of readout signals [28, 376]. For example, suppose that the measurement  $E$  is not perfectly implemented and the device fails to couple to the system with probability  $p$ . In this case, the device generates its output at complete random. This can be modeled by the effects that are depolarized as

$$E_j \mapsto (1 - p)E_j + p \frac{\mathbb{1}}{N}, \quad (8.35)$$

where  $E = (E_1, \dots, E_N)$ . Another example is the readout error, which is particularly important for superconducting qubits [377–379]. As the name suggests, readout noise refers to errors that occur due to a misreading of the outcomes in the computational basis. Typically, such errors result from decoherence during the measurement process and from overlapping support between the measured physical quantities that correspond to the  $|0\rangle$  and the  $|1\rangle$  state. Errors of that kind have recently attracted a lot of attention and different error mitigation schemes were developed [378–381]. One common model for readout noise is the tensor product noise, where one assumes that the noise acts independently on each qubit. Note that this model does not take into account cross-talk during the readout, i.e., correlated noise between multiple qubits. We

define  $q_+$  to be the probability that outcome 0 in the computational basis is misread as 1 and  $q_-$  to be the probability that outcome 1 in the computational basis is misread as 0. In the case of one qubit, this can be summarised in a matrix of the form

$$A = \begin{pmatrix} 1 - q_+ & q_- \\ q_+ & 1 - q_- \end{pmatrix} \quad (8.36)$$

and  $A_{ij}$  is the probability that outcome  $j$  is correctly read as  $i$ . For a 2-outcome measurement  $E = (E_0, E_1)$  intended, the noise  $A$  will result in a new actual measurement  $\tilde{E}$  with effects of the form  $\tilde{E}_0 = A_{00}E_0 + A_{01}E_1$  and  $\tilde{E}_1 = A_{10}E_0 + A_{11}E_1$ . For the case of randomisation of three Pauli observables, i.e., the octahedron measurement, each pair of effects corresponding to the same Pauli observable suffers from this modification due to noise, and the effects of the generalized measurement become  $1/3\{(1 - q_{\pm})|t^{\pm}\rangle\langle t^{\pm}| + q_{\mp}|t^{\mp}\rangle\langle t^{\mp}|, t = x, y, z\}$ . The error rate averaged over the two bases is denoted by  $\bar{q} := (q_+ + q_-)/2$  and the asymmetry between them is characterized by  $\epsilon := (q_+ - q_-)/(q_+ + q_-)$ . Our formalism directly takes measurement error correction into account, once the noisy effects with an appropriate model are used instead of the ideal ones. To access the quality of the shadow tomography after error correction, we choose  $|\mathfrak{X}| = 128$  pure state projections distributed according to the Haar measure as observables. The dependence of the maximal shadow norm  $\kappa_E^2(\mathfrak{X})$  on the noise parameters for the tetrahedron and the octahedron measurements is shown in Fig. 8.3. It is interesting to see that in either case the maximal shadow norm  $\kappa_E^2(\mathfrak{X})$  depends only weakly on the small error rate, showing the robustness of shadow tomography against noise. However, an increasing error results in an increased shadow norm and this in a higher number of necessary samples. This is a typical behavior for error mitigation protocols where one obtains an unbiased estimator even in the presence of noise at the cost of an increased variance [32]. In general, the model depends on the chosen implementation, i.e., on the dilation of the measurement. More precisely, often the generalized measurement is implemented by an ideal measurement on  $n$  ancillary qubits after appropriately coupling to the system. For  $n$  ancillary qubits, each can be affected by a different noise and hence the error matrix for the tensor product model is given by  $A = A_1 \otimes \cdots \otimes A_n$  where  $A_k$  depends on the parameters  $q_+^{(k)}$  and  $q_-^{(k)}$ . The precise effects for the noisy generalized measurement however depend on how precisely the outcomes on the ancillary qubits represent the outcomes of the generalized measurement.

## 8.7 Conclusion and discussion

In this Chapter we have presented a formulation of shadow tomography based generalized measurements which is more general and simpler than the original formulation. We showed how expectation values of observables can be estimated without the need

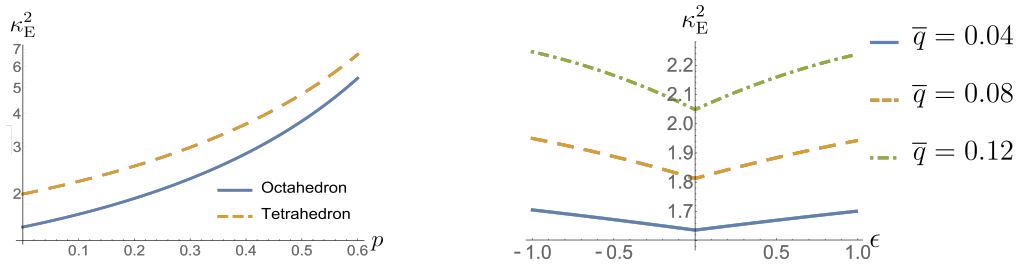


Figure 8.3: Effects of depolarizing noise and simple readout error noise on the maximal shadow norm of 128 pure state projections distributed according to the Haar measure. The figure is taken from Ref. [E].

of reconstructing the full density operator and explained how the symmetry of the generalized measurement simplifies the calculation of the classical shadows. In addition, we proved the optimality of the octahedron measurement with respect to a particular class of targeted observables. Finally, we discussed how noise can affect the shadow reconstruction and how its effect on the estimated expectation values can be mitigated.

The new formulation via generalized measurements sheds light on various aspects of shadow tomography itself, opening a range of interesting questions for further research. An extension of this framework to channel tomography would be of direct interest. Further, it would also be important to see whether the technique of derandomization could be incorporated into this framework. Such a protocol would then start with some generalized measurement  $E$ , corresponding to a randomized measurement scheme, which is then adaptively modified, yielding a new measurement  $\tilde{E}$  which performs at least as well as the initial one [362]. The optimality of the octahedron measurement for shadow tomography for a qubit-based system suggests a connection between geometry and shadow tomography. Investigation of this connection and extension for higher dimensional systems would be an interesting direction. Also the construction of optimal measurements for nonlinear functions of the density operator, or shadow tomography of a specific set of density operators, is in demand for further applications of shadow tomography.



## 9 Error mitigated classical shadows

Near-term and early fault-tolerant quantum computers are only able to prepare noisy quantum states deviating from the targeted ideal, noise-free state. However, one often aims to predict properties of the ideal state while just having access to a noisy device. Here we consider error mitigation techniques such as probabilistic error cancellation, zero-noise extrapolation and symmetry verification, which have been developed for mitigating errors in single expected value measurements, and generalize them for mitigating errors in classical shadows. These classical shadows provide a description of the quantum state that can be efficiently stored and processed. The probabilistic error cancellation approach turns out to be most natural and we develop a thorough theoretical framework including rigorous sample complexities. Naturally, the sample complexity for simultaneously predicting many linear properties of the ideal state turns out to be a combination of those two. Further, we showcase in numerical simulations a broad range of useful practical applications of our approach. This Chapter is based on Project [B]. The main contribution of the author of this thesis to this work are the proofs of the theorems and lemmas.

### 9.1 Motivation

Quantum computers and simulators are developing rapidly and can already be said to perform certain demonstration tasks that are very difficult even with the largest supercomputers [28, 382–384]. It is, however, still to be seen whether the technology can achieve true practical quantum advantage, i.e., the point when these machines can solve an otherwise impossible computational task that is relevant in fields like quantum field theory [350], condensed matter physics [385] or material science [22, 386–388].

Quantum computers turn out to be highly vulnerable to noise. While quantum error correction provides a comprehensive solution, its implementation poses an extreme engineering challenge [31]. It is generally expected that in the near future some form of early practical quantum advantage just beyond the reach of classical computing could be achieved even with noisy quantum computers [32–34, 389]. This prospect has motivated the development of a broad range of quantum error mitigation protocols, which has grown into an entire subfield [32]. While the range of error mitigation techniques is very diverse, they collectively aim to mitigate the effects of gate errors in

measuring the expectation value of observables, which is a key subroutine in quantum computing [390].

Another major challenge is that near-term quantum algorithms typically require an extremely high number of circuit repetitions in order to suppress shot noise [351, 390–392]. To overcome this problem, the concept of classical shadows was introduced [349, 353], representing another promising approach in achieving quantum advantage. These classical shadows allow one to extract many properties of the quantum state without having to repeat the measurement many times. This is achieved by performing measurements in random bases. The measurement outcomes as bit strings, along with the indices of the measurement bases, form a classical shadow, which is an efficient classical representation of the entire quantum state. Using the technique of classical shadows, various promising applications have been proposed [393, 394] that greatly benefit from the rich information one can access via classical shadows. For instance, in shadow spectroscopy [394], one aims to estimate many time-dependent expected values from time-evolved quantum states, which then allows to reveal accurate spectra through the use of efficient classical post-processing.

In this Chapter we aim to amalgamate quantum error mitigation techniques with classical shadows. It is worth noting that prior works have considered fruitful connections between quantum error mitigation and classical shadows. For instance [395], one can use classical shadows obtained from a noisy quantum state to perform purification-based error mitigation [396] offline, with only access to a single copy of the state but at an exponential complexity in the number of qubits. Further, the mitigation of errors in the measurement phase has similarly been addressed [366, 397]. However, these methods assume that the task involves extracting information from a predetermined quantum state  $\rho$ , such as the output of a quantum device. Here we consider the practically more relevant scenario where the state  $\rho$  is prepared by a noisy quantum circuit, and one aims to mitigate the impact of errors induced by the noisy quantum gates. Our focus is thus to extract properties of an ideal state  $\rho_{\text{id}}$ , which would be generated by a noise-free quantum computer. This approach can be seen as a generalization of quantum error mitigation techniques which generally aim to extract an ideal expectation value  $\text{Tr}[O\rho_{\text{id}}]$  when only noisy measurements  $\text{Tr}[A\rho]$  are available. In contrast, our techniques are not restricted to a single expectation value but instead provide efficient classical representations of the ideal quantum state  $\rho_{\text{id}}$  through classical shadows, see Fig. 9.1. While we cover most classes of conventional error mitigation techniques including probabilistic error cancellation (PEC), zero-noise extrapolation (ZNE) and symmetry verification (SV), we find that PEC is the most natural to be used in combination with classical shadows. Further, we showcase in numerical simulations a broad range of useful practical applications that will play a crucial role in both near-term and in the early fault-tolerance era. These examples comprise: (a) determining error mitigated energies in variational quantum circuits, which constitutes a fundamental subroutine

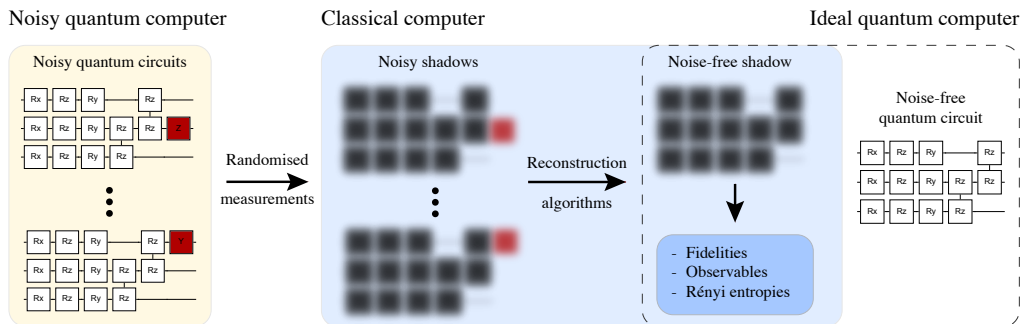


Figure 9.1: In this Chapter we assume that we only have access to a noisy quantum computer (left) such that every circuit run (left, yellow area) gets corrupted by gate noise, represented by the red gate elements. We aim to extract properties of a state that would be prepared by an ideal quantum computer (right) with the use of powerful error mitigation techniques. Our approach allows us to obtain classical shadows of the ideal quantum state (noise-free shadow) from which we can predict many ideal properties in a classical post-processing step (middle blue area, classical computer). In our formalism, we run a series of distinct quantum circuit variants (left, yellow area) that cast different classical shadows (noisy shadows) due to the gate noise and due to our intentional recovery operations. Under the assumption that the device’s error characteristics have been appropriately learned, we can estimate the noise free shadow (middle) via classical post-processing. The figure is taken from Ref. [B].

in near-term applications, (b) predicting many properties simultaneously in ground state preparations to extract two-point correlators or to accelerate the training of circuit parameters, and (c) extracting error mitigated local entanglement entropies of a ground state that is prepared by a noisy quantum circuit.

This Chapter is organized as follows. We begin by introducing our notation and recapitulating the concepts of classical shadows and probabilistic error cancellation in Section 9.2. In Section 9.3 we introduce probabilistic error canceled shadows and provide a rigorous analysis of its performance. In addition, we discuss classical post-processing algorithms for the pivotal scenario of Pauli basis measurements. We proceed in Section 9.4 by combining further error mitigation techniques with classical shadows, including zero-noise extrapolation and symmetry verification. Finally, in Section 9.5, we demonstrate a broad range of useful practical applications, that will play a crucial role in the near-term era.



## 9.2 Introduction and notation

### 9.2.1 Classical shadows

The original idea of classical shadow tomography is to apply to the quantum system of  $N$  qubits, prepared in a specific state  $\rho$ , a unitary  $Q_j$  randomly sampled from a certain ensemble  $\mathcal{Q}$ . Typically the ensemble corresponds to just rotating the individual qubits with single-qubit unitaries (Pauli basis measurements) or applying Clifford rotations. This is followed by a measurement in the computational basis, yielding a bit string  $b \in \{0,1\}^N$  as the outcome. This bit string is logged along with the measurement basis forming the index  $l = (j, b)$ . The collection of these indices from many independent runs of the protocol then allows us to construct a classical shadow of the state. A classical shadow provides a description of the quantum state that can be classically efficiently stored and manipulated, bypassing the computationally-expensive reconstruction of the full density matrix [353].

Mathematically, we describe a particular measurement outcome  $l = (j, b)$  by a positive operator as  $E_l = p_j Q_j^\dagger |b\rangle\langle b| Q_j$ . The probability  $q_l = \text{Tr}[\rho E_l]$  of this outcome is a product of a (classical) probability  $p_j$  of choosing a unitary  $Q_j$  and the probability of observing the bit string  $b$  given the rotated measurement basis. The shadow protocol can therefore be compactly described by a set  $E$  of  $N_E = 2^N |\mathcal{Q}|$  positive operators given by

$$E = \{E_l = p_j Q_j^\dagger |b\rangle\langle b| Q_j, \text{ with } Q_j \in \mathcal{Q}, b \in \{0,1\}^N\}. \quad (9.1)$$

The collection  $E$  of positive operators  $E_l$  sums up to the identity and therefore constitutes a valid generalized measurement [29]. It has been shown that formulating shadow tomography using POVMs brings various advantages [398], see also Section 8. Particularly relevant to our purpose, this formulation allows one to automatically account for errors in the measurement phase, which include both read-out errors and gate errors in the implementation of the random unitaries  $Q_j$  [28, 31, 366, 376, 397]. This is carried out by simply adjusting the effects  $E_l$  appropriately [398, 399]. As the errors in the measurement phase are subsumed into the POVM, we need only focus on mitigating errors in the state-preparation phase.

Given the above generalised measurement, a single outcome  $l = (j, b)$  can be used to construct a snapshot as  $\hat{\rho}_l = C_E^{-1}(E_l)$  where the channel  $C_E$  is defined by

$$C_E(\rho) = \sum_{l=1}^{N_E} \text{Tr}[\rho E_l] E_l, \quad (9.2)$$

which is invertible if  $E$  spans the whole space of observables [353, 398]. The snapshot can be thought of as single-shot estimator of the prepared state  $\rho$ . In an experiment one repeats the above single-shot procedure  $N_s$  times, which produces a collection of

outcomes  $\{l_1, l_2, \dots, l_{N_s}\}$ . Accordingly, a collection of snapshots can be constructed

$$S(\rho, N_s) = \{\hat{\rho}_{l_1}, \hat{\rho}_{l_2}, \dots, \hat{\rho}_{l_{N_s}}\},$$

which is called a *classical shadow* of  $\rho$ . The classical shadow allows us to obtain an unbiased estimate for the density operator in the sense  $\rho = \mathbb{E}_l[\hat{\rho}_l]$ .

Crucial to the advantage of shadow tomography is that when the measurement  $E$  consists of independent measurements on individual qubits, the snapshots  $\hat{\rho}_l$  also factorize into a tensor product over the qubits. It is therefore sufficient to store single-qubit tensoring factors of  $\hat{\rho}_l$ , instead of the exponentially large matrix itself [353]. Functions of the density operator with appropriate locality, such as correlation functions or the Rényi entropy, can also be efficiently estimated [353]. As an example, for the experimentally-friendly case of randomized noiseless Pauli basis measurements on the qubits, the snapshot corresponding to  $l = (j, b)$  is given explicitly by

$$\hat{\rho}_l = \bigotimes_{i=1}^N \left[ 3(Q_j^{(i)})^\dagger |b^{(i)}\rangle \langle b^{(i)}| Q_j^{(i)} - \mathbb{1} \right]. \quad (9.3)$$

Above,  $b^{(i)}$  is the  $i^{\text{th}}$  bit of the  $N$ -qubit measurement outcome bit string  $b$ , and  $Q_j^{(i)}$  is the  $i^{\text{th}}$  single-qubit basis transformation in the applied  $N$ -qubit Pauli basis transformation  $Q_j$ . In the following, we focus on this practically-pivotal randomized Pauli-measurement scheme. However, our general formalism can immediately be applied to other unitary ensembles such as matchgates [400], Clifford circuits [353] and beyond [401].

### 9.2.2 Probabilistic error cancellation

PEC is one of the most broadly studied error mitigation techniques [32, 381, 402] and indeed has been implemented experimentally [33, 34]. It allows us to remove errors in an expected value measurement under the assumption that we have a precise knowledge of all the quantum gates' error characteristics. In particular, one proceeds by decomposing the channel  $\mathcal{U}$  of an ideal unitary gate into a linear combination of noisy physical gate operations  $\mathcal{G}_k$  as  $\mathcal{U} = \sum_k \gamma_k \mathcal{G}_k$ . Negative quasiprobabilities  $\gamma_k < 0$  are required to formally implement the inverse of a noise channel. Thus, the above operation is nonphysical, similarly as the inverse measurement channels of the shadow protocols are nonphysical operations. For this reason, PEC only applies the decomposition in classical post-processing, at the level of expected values and allows us to compute ideal expected values of an observable  $O$  as a linear combination of noisy ones as  $\sum_k \gamma_k \text{Tr}[O \mathcal{G}_k |0\rangle \langle 0|]$ .

Efficient methods have been developed for accurately learning approximate noise models in practice [403–406]. The simplest such approach exploits the fact that noise models are approximately local and one can thus efficiently characterize the local noise

channel of each gate and invert them classically. Furthermore, non-local noise models of the form  $\Lambda^G = e^{\mathcal{L}}$  can be efficiently learned for the case of sparse Pauli operations  $\mathcal{L}$  which allows us to invert the channel trivially as  $\Lambda^{G^{-1}}$  [405]. One then randomly applies circuit variants that implement the operations  $\mathcal{G}_k$  in the inverse noise channel. Here we assume that an exact decomposition for each gate is known in advance and thus  $\mathcal{U}$  is supported only on the space spanned by the noisy operations  $\mathcal{G}_k$ .

Assume now that an ideal quantum state  $\rho_{id} := \mathcal{U}_{circ}|0\rangle\langle 0|$  is prepared by an ideal circuit  $\mathcal{U}_{circ} = \mathcal{U}_v \circ \dots \circ \mathcal{U}_2 \circ \mathcal{U}_1$  of  $v$  gates. By introducing the vector notation  $\underline{k} = (k_1, k_2, \dots, k_v)$ , we can compactly represent the decomposition of this circuit into noisy gate sequences as  $\mathcal{U}_{circ} = \sum_{\underline{k}} g_{\underline{k}} \mathcal{G}_{\underline{k}}$ . Here the index  $\underline{k}$  indexes all possible gate sequences as

$$\mathcal{G}_{\underline{k}} = \mathcal{G}_{k_1}^{(1)} \mathcal{G}_{k_2}^{(2)} \dots \mathcal{G}_{k_v}^{(v)}, \quad g_{\underline{k}} = \gamma_{k_1}^{(1)} \gamma_{k_2}^{(2)} \dots \gamma_{k_v}^{(v)}, \quad (9.4)$$

and as shown above the corresponding quasiprobabilities  $g_{\underline{k}}$  factorize (the superscript indexes individual gates, e.g.,  $\mathcal{G}_{k_1}^{(1)}$  stands for the decomposition of  $\mathcal{U}_1$ ). We now define the quasiprobability decomposition of a quantum circuit.

**Definition 62.** *The quasiprobability decomposition of an ideal circuit  $\mathcal{U}_{circ}$  is the set  $G := \{(g_{\underline{k}}, \mathcal{G}_{\underline{k}})\}_{\underline{k}}$ . The associated probability distribution is  $p(\underline{k}) := |g_{\underline{k}}| / \|g\|_1$  and here the norm factorizes as  $\|g\|_1 = \prod_{k=1}^v \|\gamma^{(k)}\|_1$  when the circuits are of a product form as in Eq. (9.4).*

The above quasiprobability decomposition has been used for estimating the ideal expected value of an observable  $O$ ,  $\text{Tr}[O\mathcal{U}_{circ}|0\rangle\langle 0|]$ , by randomly sampling the noisy expected values  $\text{sign}(g_{\underline{k}}) \text{Tr}[O\mathcal{G}_{\underline{k}}|0\rangle\langle 0|]$  according to the probability distribution  $p(\underline{k})$  and linearly combining them in a classical post-processing step [381, 402].

### 9.3 Probabilistic error canceled shadows

While PEC has been used in the literature to remove the bias in expectation value measurements [32], here we apply it to classical shadows in order to obtain an efficient, classical representation of the entire, ideal and noise-free state  $\rho_{id}$ . While this procedure even allows us to estimate the full density matrix  $\rho_{id}$ , we will focus mostly on efficient practical applications such as simultaneously predicting many properties of  $\rho_{id}$ . At a technical level, PEC shadows is a combination of two random processes, i.e., sampling circuit variants  $\mathcal{G}_{\underline{k}}$  and sampling the bit strings and the basis transformations that form a shadow. We start by generalizing the PEC technique such that the quasiprobability decomposition allows us to obtain an unbiased estimator of the full density matrix.

### 9.3.1 The protocol

**Lemma 63.** *Given a quasiprobability decomposition  $G$  from Definition 62, by sampling the noisy circuits  $\mathcal{G}_{\underline{k}}$  according to the probability distribution  $p(\underline{k})$  we obtain an unbiased estimator of the ideal density matrix  $\varrho_{\text{id}} := \mathcal{U}_{\text{circ}}|0\rangle\langle 0|$  as*

$$\hat{\varrho}_{\text{id}} = \|g\|_1 \text{sign}(g_{\underline{k}}) \mathcal{G}_{\underline{k}} |0\rangle\langle 0| \quad (9.5)$$

in the sense that  $\mathbb{E}_{\underline{k}}[\hat{\varrho}_{\text{id}}] = \varrho_{\text{id}}$ .

*Proof.* The statement directly follows from the fact that  $\|g\|_1 \text{sign}(g_{\underline{k}}) \mathcal{G}_{\underline{k}}$  is an unbiased estimator for the ideal operation  $\mathcal{U}_{\text{circ}}$ . In particular, as we sample  $\underline{k}$  according to the probability distribution  $p(\underline{k})$ , we obtain the expectation as

$$\begin{aligned} \mathbb{E}_{\underline{k}}[\hat{\varrho}] &= \mathbb{E}_{\underline{k}}[\|g\|_1 \text{sign}(g_{\underline{k}}) \mathcal{G}_{\underline{k}}(|0\rangle\langle 0|)] \\ &= \sum_{\underline{k}} p(\underline{k}) \|g\|_1 \text{sign}(g_{\underline{k}}) \mathcal{G}_{\underline{k}}(|0\rangle\langle 0|). \end{aligned} \quad (9.6)$$

The above expression can be further simplified by collecting the constant factors as  $p(\underline{k}) \|g\|_1 \text{sign}(g_{\underline{k}}) = \text{sign}(g_{\underline{k}}) |g_{\underline{k}}| = g_{\underline{k}}$  and thus we obtain the quasiprobability decomposition

$$\mathbb{E}_{\underline{k}}[\hat{\varrho}] = \sum_{\underline{k}} g_{\underline{k}} \mathcal{G}_{\underline{k}}(|0\rangle\langle 0|) = \mathcal{U}_{\text{circ}}(|0\rangle\langle 0|) = \varrho_{\text{id}}. \quad (9.7)$$

□

The above estimator has a clear operational meaning. First, choose a multi-index  $\underline{k}$  randomly according to the probability distribution  $p(\underline{k})$  and run the noisy quantum circuit  $\mathcal{G}_{\underline{k}}$ . Second, the output state  $\mathcal{G}_{\underline{k}}|0\rangle\langle 0|$  is a density matrix that we multiply by  $\text{sign}(g_{\underline{k}})$  and with the norm  $\|g\|_1$ . Finally, from a formal perspective, the mean of these matrices is an estimate of the ideal density matrix  $\varrho_{\text{id}}$ .

Regrettably, the above protocol is purely formal as the multiplication with negative quasiprobabilities is non physical and could only be achieved in post-processing, e.g., after fully reconstructing the density matrix. We thus exploit classical shadows as a powerful tool for obtaining an efficient classical description of the states which can then be naturally assigned negative quasiprobabilities in classical post-processing. Indeed, snapshots are not physical density matrices either, as it is apparent in Eq. (9.3). We now state our protocol that serves as an unbiased estimator of the ideal state.

**Theorem 64.** *Given a quasiprobability decomposition  $G$  of the ideal circuit  $\mathcal{U}_{\text{circ}}$  from Definition 62, and a classical shadow protocol with the POVM measurement  $E$  from (9.1), we define PEC shadows as the set  $H := \{(g_{\underline{k}}, \mathcal{G}_{\underline{k}}, E_l)\}_{\underline{k}, l}$  and define the corresponding PEC snapshot as*

$$\hat{\varrho}_{\text{id}} := \hat{\varrho}_{k,l} = \|g\|_1 \text{sign}(g_{\underline{k}}) C_E^{-1}(E_l). \quad (9.8)$$

We will often use the notation  $\hat{q}_{id}$  to abbreviate  $\hat{q}_{\underline{k},l}$  as it is an unbiased estimator of the ideal density matrix  $q_{id}$  such that  $\mathbb{E}[\hat{q}_{id}] = \mathbb{E}_{\underline{k},l}[\hat{q}_{\underline{k},l}] = q_{id}$ .

*Proof.* Using the abbreviation  $\hat{q}_{id} \equiv \hat{q}_{\underline{k},l}$  we calculate the expected value as

$$\mathbb{E}_{\underline{k},l}[\hat{q}_{id}] = \sum_{\underline{k},l} p_{\underline{k}} q_l \hat{q}_{\underline{k},l}, \quad (9.9)$$

where  $p_{\underline{k}} = |g_{\underline{k}}|/||g||_1$  is the probability of choosing the circuit variant  $\mathcal{G}_{\underline{k}}$  from Definition 62 and we also use the probability  $q_l = \text{Tr}[\mathcal{G}_{\underline{k}}(|0\rangle\langle 0|)E_l]$  of observing the POVM outcome  $l$  given that the circuit variant  $\mathcal{G}_{\underline{k}}$  was implemented. We obtain the expected value by substituting these in Eq. (9.9) as

$$\mathbb{E}_{\underline{k},l}[\hat{q}_{id}] = \sum_{\underline{k},l} \frac{|g_{\underline{k}}|}{||g||_1} \text{Tr}[\mathcal{G}_{\underline{k}}(|0\rangle\langle 0|)E_l] ||g||_1 \text{sign}(g_{\underline{k}}) C_E^{-1}(E_l). \quad (9.10)$$

Here we can collect and simplify all constant factors as  $||g||_1 \frac{|g_{\underline{k}}|}{||g||_1} \text{sign}(g_{\underline{k}}) = g_{\underline{k}}$  and simplify the expected value as

$$\begin{aligned} \mathbb{E}_{\underline{k},l}[\hat{q}_{id}] &= \sum_l \text{tr} \left[ \left( \sum_{\underline{k}} g_{\underline{k}} \mathcal{G}_{\underline{k}} \right) (|0\rangle\langle 0|) E_l \right] C_E^{-1}(E_l) \\ &= \sum_l \text{Tr}[q_{id} E_l] C_E^{-1}(E_l) = C_E^{-1} \left( \sum_l \text{Tr}[q_{id} E_l] E_l \right) \\ &= (C_E^{-1} \circ C_E)(q_{id}) = q_{id}. \end{aligned} \quad (9.11)$$

Above in the first equality we simply used the linearity of the trace operation while in the second equality we used that by definition  $\sum_{\underline{k}} g_{\underline{k}} \mathcal{G}_{\underline{k}} |0\rangle\langle 0| = q_{id}$ . We finally substituted the definition of the measurement channel  $C_E(\cdot)$  given by Eq. (9.2).  $\square$

The averaging in Theorem 64 happens not only over the effects  $E_l$  indexed by  $l$  (all basis transformations and measurement outcomes), but additionally we average over all circuit variants indexed by  $\underline{k}$ . The reason is that the measurement  $E = \{E_l\}_l$  is not performed on a fixed input density matrix  $\rho$  as in conventional shadow tomography but rather on the quasiprobability decomposition of the ideal state  $q_{id} \propto \mathcal{G}_{\underline{k}} |0\rangle\langle 0|$ . Let us now summarize the resulting experimental protocol.

- (1) randomly choose a multi-index  $\underline{k}$  according to the probabilities  $p(\underline{k})$  and store the  $\text{sign}(g_{\underline{k}})$
- (2) uniformly and randomly choose a unitary rotation  $Q_j \in \mathcal{Q}$  and store its index  $j$
- (3) execute in a quantum computer the gate sequence  $\mathcal{G}_{\underline{k}}$ , the unitary rotation  $Q_j$ , perform a measurement in the standard basis and finally register its outcome  $b$
- (4) each stored index  $(\text{sign}(g_{\underline{k}}), j, b)$  uniquely identifies a classical snapshot  $\hat{q}_{\underline{k},l} = ||g||_1 \text{sign}(g_{\underline{k}}) C_E^{-1}(E_l)$  where we recall that  $E_l$  is a POVM effect with the index  $l = (j, b)$  from Eq. (9.1)

- (5) repeat the procedure and collect  $N_s$  classical snapshots to build a classical shadow of the ideal state  $S(\rho_{\text{id}}, N_s) = \{(\hat{Q}_{\text{id}})_1, (\hat{Q}_{\text{id}})_2, \dots, (\hat{Q}_{\text{id}})_{N_s}\}$

The classical data set  $S(\rho_{\text{id}}, N_s)$  can then be classically post-processed offline and we detail explicit algorithms for predicting local properties in Section 9.3.3. Note that PEC shadows produce a distribution of snapshots that is different from directly applying conventional shadow tomography to a noise-free state  $\rho_{\text{id}}$ , albeit with an identical mean. The reason is that each circuit variant  $\mathcal{G}_k$  in (9.5) yields a different distribution of classical snapshots.

### 9.3.2 Rigorous performance guarantees

We first consider the pivotal practical application of predicting error mitigated expected values of observables  $O$  via the estimator  $\hat{o} = \text{Tr}[O\hat{\rho}_{\text{id}}]$ . A key observation is that in error mitigation techniques the ability to predict noise-free expected values comes at the cost of an increased statistical variance which implies an increased number of circuit repetitions.

**Lemma 65.** *We define the shadow norm with respect to the generalized measurement  $E$  as*

$$\|O\|_E^2 := \left\| \sum_{l=1}^{N_s} \text{Tr}[\hat{Q}_l O]^2 E_l \right\|_{\infty}, \quad (9.12)$$

where  $\|\cdot\|_{\infty}$  denotes the maximal eigenvalue of the corresponding operator and  $\hat{Q}_l = C_E^{-1}(E_l)$ . For the specific case of Pauli-basis measurements and observables that are  $q$ -local Pauli strings, the squared shadow norm is given as  $3^q$ .

*Proof.* When formulating shadow tomography with generalized measurements, the case of uniformly sampled Pauli-basis measurements corresponds to the so-called octahedron POVM [398], where the effects on a single qubit are given by

$$E_j = \frac{1}{3} Q_j^\dagger |b\rangle\langle b| Q_j, \quad (9.13)$$

where  $b \in \{0, 1\}$  is a single bit and  $Q_j$  is one of the three basis transformation unitaries that allow us to measure in the bases of the Pauli  $X$ ,  $Y$  and  $Z$  operators. More precisely,  $Q_j$  is a rotation mapping  $|0\rangle, |1\rangle$  to  $|t^-\rangle, |t^+\rangle$ , e.g.,  $E_1 = \frac{1}{3} Q_1^\dagger |0\rangle\langle 0| Q_1$ ,  $E_2 = \frac{1}{3} Q_1^\dagger |1\rangle\langle 1| Q_1$  with  $Q_1 = \exp(-i(\pi/4)Y)$  a rotation around the  $Y$ -axis. Thus the effect is equivalent to  $\frac{1}{3} |t^\pm\rangle\langle t^\pm|$  for  $t \in \{x, y, z\}$  where  $|t^\pm\rangle$  denotes the eigenvector corresponding to eigenvalue  $\pm 1$  of the single-qubit Pauli- $t$  operator. It follows from the symmetry of the measurement [398] that the shadows can be computed directly from the effects as

$$\hat{Q}_l = 9E_l - \mathbb{1}. \quad (9.14)$$

For the case of a system consisting of  $n$  qubits where one aims to estimate local observables of the form  $O = O_1 \otimes \cdots \otimes O_n$ , and the measurement is given by the tensor product of local measurements  $E_{j_1}^{(1)} \otimes \cdots \otimes E_{j_n}^{(n)}$ , with  $E^{(j)}$  the POVM acting on the  $j$ th qubit, the shadow norm similarly factorizes as  $\|O\|_E^2 = \prod_j \|O_j\|_{E^{(j)}}^2$ . We now consider the case when the single-qubit operator  $O_j$  acting on the  $j$ th qubit is a Pauli operator  $X$ ,  $Y$  or  $Z$  and thus  $\text{tr}[O_j] = 0$ . By the previous discussion, it is sufficient to only consider a single qubit, thus we will suppress the index  $j$ . This yields the shadow norm

$$\begin{aligned} \|O\|_E^2 &= \left\| \sum_{l=1}^6 \text{Tr}[\hat{\rho}_l O]^2 E_l \right\|_\infty \\ &= \left\| \sum_{t^\pm} \frac{1}{3} \text{Tr}[(3|t^\pm\rangle\langle t^\pm|)O]^2 |t^\pm\rangle\langle t^\pm| \right\|_\infty \\ &= 3 \left\| \sum_{t^\pm} \langle t^\pm | O | t^\pm \rangle^2 |t^\pm\rangle\langle t^\pm| \right\|_\infty. \end{aligned} \quad (9.15)$$

Now observe that if  $O, T \in \{X, Y, Z\}$  with  $|t^\pm\rangle$  the normalized eigenvectors of  $T$  to eigenvalues  $\pm 1$ , we have due to the anticommutation relation  $\delta_{O,T} = \frac{1}{2} \langle t^\pm | \{O, T\} | t^\pm \rangle = \pm \langle t^\pm | O | t^\pm \rangle$ . This implies that the sum in Eq. (9.15) collapses to the identity  $\mathbb{1}$ . Hence we obtain  $\|O\|_E^2 = 3$ . When the single-qubit observable is the identity  $O = \mathbb{1}$  we obtain the shadow norm  $\|O\|_E^2 = 1$ . Consequently, for  $q$ -local Pauli strings acting on  $n$  qubits the squared shadow norm is  $\|O\|_E^2 = 3^q$ , thus independent of  $n$ .  $\square$

In practice it is often the case that the set of targeted observables possess a certain structure. If this is the case, small variations to the classical shadow protocol in which the measurement basis is sampled uniformly at random can yield a substantial improvement with respect to sample complexity [356]. For instance, in electronic structure problems where one aims to, e.g., determine the ground-state of molecules using a quantum algorithm, one typically starts by transforming the molecular Hamiltonian into a qubit Hamiltonian as a sum of Pauli observables by means of an appropriate *mapping*. Common types of such mappings are Jordan-Wigner (JW), Bravyi-Kitaev (BK) and the parity (P) transformation [22]. Here it is important to note that depending on the encoding, the different Pauli operators  $X, Y$  and  $Z$  appear with different frequencies in the corresponding qubit observable. For instance, in the BK encoding, the appearance of Pauli- $Y$  operators is suppressed compared to  $X$  and  $Z$ . Consequently, measuring the different Pauli bases uniformly on each qubit, i.e., using the octahedron measurement, would be very wasteful. A similar statement concerning sample complexity as in Lemma 65 can be made for the case of locally biased shadows [356, 362]. Let us assume that the bias is  $p_x, p_y, p_z$ , where  $p_t$  is the probability for performing the measurement in Pauli- $t$  basis. The corresponding POVM would be  $E_{t^\pm} = p_t |t^\pm\rangle\langle t^\pm|$ . Then the classical shadow based on the measurement outcome  $\pm 1$  would be

$$\hat{\rho}_{t^\pm} = p_t^{-2} E_{t^\pm} - \frac{\mu - p_t^2}{2p_t\mu} \mathbb{1}, \quad (9.16)$$

where  $\mu = p_x^2 + p_y^2 + p_z^2$ . With this, given a Pauli string, one can directly calculate the shadow norm.

**Lemma 66.** *Given an observable  $O$  and the PEC snapshot  $\hat{q}_{id}$  from Theorem 64, the variance of the estimator  $\hat{\delta} = \text{Tr}[O\hat{q}_{id}]$  can be upper bounded as*

$$\text{Var}[\hat{\delta}] \leq \|g\|_1^2 \|O\|_E^2, \quad (9.17)$$

where  $\|\cdot\|_E^2$  is the shadow norm of the observable  $O$  with respect to the measurement  $E$  as defined in Lemma 65. When  $O$  is a  $q$ -local Pauli string and we use Pauli basis measurements, then  $\|O\|_E^2 = 3^q$  as explained in Lemma 65.

*Proof.* Note that  $\text{Var}[\hat{\delta}] = \mathbb{E}[(\hat{\delta} - \mathbb{E}[\hat{\delta}])^2]$ . As  $\hat{q}_{id}$  is an unbiased estimator for the ideal state, we have  $\mathbb{E}[\text{Tr}(O\hat{q}_{id})^2] = \langle O \rangle^2$  and thus  $\text{Var}[\hat{\delta}] = \mathbb{E}[\text{Tr}(O\hat{q}_{id})^2] - \langle O \rangle^2 \geq \mathbb{E}[\text{Tr}(O\hat{q}_{id})^2]$ . Hence it remains to bound the term

$$\mathbb{E}_{k,l} \left[ \text{Tr}(O\hat{q}_{id})^2 \right] = \mathbb{E}_{k,l} \left[ \text{Tr} \left[ O \|g\|_1 \text{sign}(g_k) C_E^{-1}(E_l) \right]^2 \right] \quad (9.18)$$

$$= \|g\|_1^2 \mathbb{E}_{k,l} \left[ \text{Tr} \left[ OC_E^{-1}(E_l) \right]^2 \right]. \quad (9.19)$$

We can now calculate the expectation by recalling that  $p_k = |g_k|/\|g\|_1$  is the probability from Definition 62 of choosing the circuit variant  $\mathcal{G}_k$  and  $q_l = \text{Tr}[\mathcal{G}_k(|0\rangle\langle 0|)E_l]$  is the probability of observing the POVM outcome  $l$ . Thus the above expectation is calculated as

$$\begin{aligned} & \|g\|_1^2 \sum_{k,l} \frac{|g_k|}{\|g\|_1} \times \text{Tr} \left[ \mathcal{G}_k(|0\rangle\langle 0|)E_l \right] \times \text{Tr} \left[ OC_E^{-1}(E_l) \right]^2 \\ &= \|g\|_1^2 \sum_l \text{Tr} \left[ \Omega(|0\rangle\langle 0|)E_l \right] \times \text{Tr} \left[ OC_E^{-1}(E_l) \right]^2. \end{aligned}$$

Above we introduced  $\Omega := \|g\|_1^{-1} \sum_k |g_k| \mathcal{G}_k$  which is actually a permissible quantum channel [29], i.e., a CPTP map, since by its definition it is a convex combination of CPTP maps  $\mathcal{G}_k$ . This expression is similar to the one in Ref. [353].

The above expression can be upper bounded by replacing the initial state  $|0\rangle\langle 0|$  by a maximization over all states  $\sigma$ . Thus we obtain the upper bound

$$\begin{aligned} & \mathbb{E}_{k,l} \left[ \text{Tr}(O\hat{q}_{id})^2 \right] \\ & \leq \|g\|_1^2 \max_{\sigma} \sum_l \text{Tr} \left[ \Omega(\sigma)E_l \right] \times \text{Tr} \left[ OC_E^{-1}(E_l) \right]^2 \\ & = \|g\|_1^2 \max_{\sigma} \text{Tr} \left[ \Omega(\sigma) \sum_l \left( \text{Tr} \left[ O\hat{q}_l \right]^2 E_l \right) \right], \end{aligned} \quad (9.20)$$

where we moved the summation inside the trace. By introducing the abbreviation



$\Gamma = \sum_l \text{Tr}[O\hat{q}_l]E_l$  we obtain the upper bound as

$$\begin{aligned} \mathbb{E}_{k,l} \left[ \text{Tr}(O\hat{q}_{\text{id}})^2 \right] &\leq \|g\|_1^2 \max_{\sigma} \text{Tr} \left[ \Omega(\sigma)\Gamma \right] \\ &\leq \|g\|_1^2 \max_{\sigma} \text{Tr} [\sigma\Gamma] = \|g\|_1^2 \|\Gamma\|_{\infty} = \|g\|_1^2 \|O\|_E^2. \end{aligned}$$

Above we used that  $\Omega(\sigma)$  is a valid density matrix and thus one can upper bound the trace by the operator norm  $\text{Tr} [\sigma\Gamma] \leq \|\Gamma\|_{\infty}$  as the largest singular value of  $\Gamma$ , which is by definition the shadow norm from Lemma 65. Since  $\langle O \rangle^2 \geq 0$ , we obtain  $\text{Var}[\hat{o}] \leq \|g\|_1^2 \|O\|_E^2 - \langle O \rangle^2 \leq \|g\|_1^2 \|O\|_E^2$ .  $\square$

Observe that the above variance depends on two factors. The first one is the squared shadow norm  $\|O\|_E^2$  which determines the sample complexity of conventional shadows [353]. The second factor is a multiplicative term  $\|g\|_1^2$  which accounts for the well-known measurement overhead associated with the conventional PEC protocol [32, 381, 402]. Some further comments are in order. First, recall that conventional classical shadows make no assumption about the input state  $\rho$  [353]. In contrast, in our case, a circuit description  $\mathcal{U}_{\text{circ}}|0\rangle\langle 0|$  of the “input state”  $\rho_{\text{id}}$  is actually part of the protocol. Of course, knowing such a description of the input state does not allow one to classically efficiently predict its properties without using classical shadows unless the circuit  $\mathcal{U}_{\text{circ}}$  has some special properties permitting efficient classical simulation, such as Clifford circuits. Second, the proof in Lemma 66 involves a maximization over density matrices such that our bounds are independent of the particular quasiprobability decomposition and thus depends only on the norm  $\|g\|_1^2$ . Third, it can be expected that the upper bound in Lemma 66 is very pessimistic. Similar, constant factor looseness of the bounds was already observed for conventional shadows [353], however the discrepancy is strongly expected to be even larger for PEC shadows. This is due to the maximization, as we do not take into account properties of the individual circuits in the quasiprobability decomposition but rather apply a pessimistic global bound. While we only state explicitly the shadow norm for the practically most important ensemble of Pauli basis measurements, we note that bounds for other ensembles are immediately available in the literature [353, 400, 401].

Following the approach of [353] we use concentration properties of the median of means estimator to derive rigorous sample complexities. For the simultaneous prediction of many observables  $O_1, \dots, O_M$  we exponentially suppress statistical outliers by splitting the PEC shadows  $S(\rho_{\text{id}}, N_s)$  into independent batches and then computing a median of the means. The resulting bounds depend on two performance metrics: The accuracy  $\epsilon$  and the success probability  $\delta$ . More precisely, in order to predict expectation values of  $M$  independent observables  $\{O_1, \dots, O_M\}$ , we group  $N_s = N_{\text{batch}}K$  independent snapshots onto  $K$  batches  $\mathcal{B}_1, \dots, \mathcal{B}_K$  each of size  $N_{\text{batch}}$ . Then, for each subset  $\mathcal{B}_j$

one uses the empirical mean as

$$\hat{\mu}_i(O_j) = \frac{1}{N_{batch}} \sum_{l \in \mathcal{B}_i} \text{Tr}[O_j(\hat{\rho}_{id})_l]. \quad (9.21)$$

The final estimate for the expectation value of  $O_j$  is then obtained by the median of the individual empirical means, that is,

$$\hat{\mu}_{K,b}(O_j) := \text{median} \{ \hat{\mu}_1(O_j), \dots, \hat{\mu}_K(O_j) \}. \quad (9.22)$$

Even though this method requires an increased number  $N_{batch}K$  of independent classical shadows, it is much more robust against outlier corruption. The idea is that if  $\hat{\mu}_{K,b}(O_j)$  deviates by more than  $\epsilon$  from  $\text{Tr}[O_j \rho_{id}]$ , more than  $K/2$  of the individual empirical mean values must deviate by more than  $\epsilon$ . This is in fact an exponentially suppressed event. This can be made more formal by the following concentration inequality of the median of means estimator [407, 408],

$$\text{Prob} \left[ |\hat{\mu}_{K,b}(O_j) - \langle O_j \rangle| \geq \frac{2\sigma}{\sqrt{N_{batch}}} \right] \leq \exp\left(-\frac{K}{8}\right), \quad (9.23)$$

where  $\sigma$  denotes the standard deviation.

**Theorem 67.** *Let  $\mathcal{U}_{circ}$  be the ideal quantum circuit producing the ideal output state  $\rho_{id}$  from Definition 62. Suppose that we want to predict  $M$  linear properties  $O_1, \dots, O_M$  of the ideal state, i.e.,  $\langle O_j \rangle = \text{Tr}[O_j \rho_{id}]$ . For fixed performance metrics  $\epsilon, \delta \in [0, 1]$  set*

$$N_{batch} = \frac{4\|g\|_1^2}{\epsilon^2} \max_{1 \leq j \leq M} \|O_j\|_E^2 \quad \text{and} \quad K = 8 \log\left(\frac{M}{\delta}\right). \quad (9.24)$$

*Then a collection of  $N = KN_{batch}$  independent classical shadows allows for accurately predicting all ideal expectation values via median of means estimation such that*

$$\text{Prob}[|\hat{\mu}_{K,b}(O_j) - \langle O_j \rangle| \leq \epsilon] \geq 1 - \delta. \quad (9.25)$$

*Proof.* This is a direct consequence of the concentration property of the median of means estimator together with the bound on the variance from Lemma 66. Because  $\text{Var}[\hat{\mu}] \leq \|g\|_1^2 \|O\|_E^2$  and if the accuracy is  $\epsilon$ , we have  $N_{batch} \geq 4\|g\|_1^2 \|O\|_E^2 \geq 4\sigma^2/\epsilon^2$ . Further, as we have  $M$  measurements that we want to accurately predict with at most failure probability  $\delta$ , we need for each individual measurement  $\exp(-K/8) \leq \delta/M$ . Thus the choice  $K = 8 \log(M/\delta)$  yields the desired bound. In total we have

$$\begin{aligned} & \text{Prob}[|\hat{\mu}_{K,b}(O_j) - \mu_j| \geq \epsilon \ \forall j] \\ &= \text{Prob}\left[\bigcup_{j=1}^M \{|\hat{\mu}_{K,b}(O_j) - \mu_j| \geq \epsilon\}\right] \\ &\leq \sum_{j=1}^M \text{Prob}[|\hat{\mu}_{K,b}(O_j) - \mu_j| \geq \epsilon] \leq M \frac{\delta}{M} = \delta. \end{aligned}$$

□

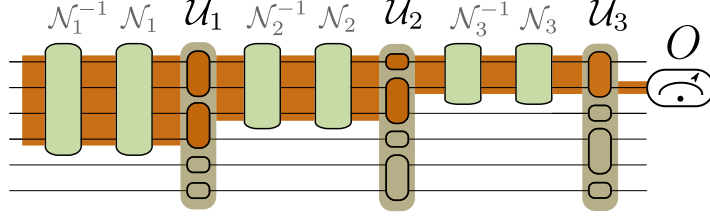


Figure 9.2: Illustration of the light cone of an observable  $O$  which is represented by the measurement apparatus and only acts nontrivially on the second (from top) qubit. The orange area indicates the qubits which are contained in the light cone  $\mathcal{I}$  of the observable with respect to the ideal quantum circuit (orange boxes)  $\mathcal{U}_3\mathcal{U}_2\mathcal{U}_1$ . To simplify derivations we assume the gate noise channels  $\mathcal{N}_k$  are local (light green boxes) such that they are contained within the lightcone  $\mathcal{I}$  but our results can be extended to non-local models via [409]. The figure is taken from Ref. [B].

Finally, we consider the problem of predicting non-linear properties of the state, which are of the form  $\text{Tr}[O(\varrho_{\text{id}})^m]$ . We can bound the variance of any non-linear property as follows.

**Theorem 68.** *Given the PEC snapshots  $\hat{q}_{\text{id}}$  from Theorem 64 we can estimate polynomial properties of degree  $m$  of the ideal state  $\varrho_{\text{id}}$  via  $U$ -statistics of tensor products of all distinct snapshots. The number of samples required to predict the non-linear property scales as  $N_s \in \mathcal{O}(\|g\|_1^{2m}/\epsilon^2)$  for a desired accuracy  $\epsilon$ .*

*Proof.* In order to obtain rigorous performance guarantees of the shadow estimator, two ingredients are needed. First, note that any polynomial function in the quantum state can be written as a linear function in tensor products of the quantum state. More precisely, suppose we want to estimate a polynomial function of degree  $m$  of the quantum state  $\varrho$ , e.g.,  $\tilde{f} : B(\mathcal{H}) \rightarrow \mathbb{R}$  with  $\tilde{f}(\varrho) := \text{Tr}[\tilde{A}\varrho^m]$ , where  $\varrho, \tilde{A} \in B(\mathcal{H})$ . If  $C^{(m)}$  denotes the cyclic permutation operator acting on  $m$  copies, that is,

$$C^{(m)} : B(\mathcal{H}^{\otimes m}) \rightarrow B(\mathcal{H}^{\otimes m}), \quad C^{(m)}(|\phi_1\rangle|\phi_2\rangle \cdots |\phi_m\rangle) = |\phi_m\rangle|\phi_1\rangle \cdots |\phi_{m-1}\rangle, \quad (9.26)$$

we can associate to  $\tilde{f}$  a function  $f$  and an operator  $A \in B(\mathcal{H}^{\otimes m})$  such that

$$f(\varrho) = \text{Tr}[A\varrho^{\otimes m}], \quad A = \text{Tr}_1[C^{(m+1)}\tilde{A} \otimes \mathbb{1}^{\otimes m}] \quad (9.27)$$

and  $f(\varrho) = \tilde{f}(\varrho)$ . The second tool needed is the so called  $U$ -statistics, which often provides a uniformly minimum variance unbiased estimator for nonlinear polynomial functions. Suppose we have access to  $N$  independent snapshots  $\hat{\rho}_1, \dots, \hat{\rho}_N$  which are generated by an underlying state  $\rho$  and that  $f(\hat{\rho}_1, \dots, \hat{\rho}_m)$  is a polynomial function in the

shadows such that  $\theta$ , what is our parameter of interest, is given by  $\theta = \mathbb{E}[f(\hat{\rho}_1, \dots, \hat{\rho}_m)]$ . The  $U$ -statistic [309] of order  $m$  is defined as

$$U_N := \binom{N}{m}^{-1} \sum_{\mathcal{C}_{N,m}} f(\hat{\rho}_{i_1}, \dots, \hat{\rho}_{i_m}), \quad (9.28)$$

where  $\mathcal{C}_{N,m}$  is the set of all combinations of  $m$  distinct elements that one can build out of  $N$  different snapshots. The variance of this estimator has a closed form in dependence on the function  $f$ . For a  $U$ -statistic  $U_N$  given by Eq. (9.28) the variance obeys [309]

$$\text{Var}[U_N] = \frac{1}{\binom{N}{m}} \sum_{d=1}^m \binom{m}{d} \binom{N-m}{m-d} \text{Var}[f^{(d)}(\hat{\rho}_1, \dots, \hat{\rho}_d)], \quad (9.29)$$

where

$$f^{(d)}(\hat{\rho}_1, \dots, \hat{\rho}_d) := \mathbb{E}_{\hat{\rho}_{d+1}, \dots, \hat{\rho}_m} [f(\hat{\rho}_1, \dots, \hat{\rho}_d, \hat{\rho}_{d+1}, \dots, \hat{\rho}_m)].$$

In order to understand the scaling of  $\text{Var}[U_N]$  it is sufficient to consider a particular instance  $\text{Var}[f^{(d)}(\hat{\rho}_1, \dots, \hat{\rho}_d)]$ . Notice that for  $A \in \mathcal{B}(\mathcal{H}^{\otimes m})$  as defined in Eq. (9.27) one has

$$f^{(d)}(\hat{\rho}_1, \dots, \hat{\rho}_d) = \text{Tr}[A \hat{\rho}_1 \otimes \dots \otimes \hat{\rho}_d \otimes \rho^{\otimes p}], \quad (9.30)$$

where  $p = m - d$  with the convention that  $\rho^{\otimes 0} = 1 \in \mathbb{C}$ . Further define  $\hat{\rho}_l = \hat{\rho}_{l_1} \otimes \dots \otimes \hat{\rho}_{l_d} \otimes \rho^p$  using the abbreviation  $\hat{\rho}_{l_1} \equiv (\hat{\rho}_{id})_{l_1}$ . For  $1 \leq d \leq m$  we have

$$\begin{aligned} \mathbb{E}[f^{(d)}(\hat{\rho}_1, \dots, \hat{\rho}_d)^2] &= \mathbb{E}[\text{Tr}(A \hat{\rho}_1 \otimes \dots \otimes \hat{\rho}_d \otimes \rho^{\otimes p})^2] \\ &= \|g\|_1^{2d} \mathbb{E}_{\substack{k_1, l_1 \\ \dots \\ k_d, l_d}} \text{Tr}[A \hat{\rho}_l]^2 = \|g\|_1^{2d} \sum_{k_1, \dots, k_d} \sum_{l_1}^d \left( \prod_{j=1}^d p(k_j) p(l_j | k_j) \right) \text{Tr}[A \hat{\rho}_l]^2 \\ &= \|g\|_1^{2d} \sum_{k_1, \dots, k_d} \sum_{l_1}^d \left( \prod_{j=1}^d \frac{|g_{k_j}|}{\|g\|_1} \right) \text{Tr}[\mathcal{G}_{k_j}(|0\rangle\langle 0|) E_{l_j}] \text{Tr}[A \hat{\rho}_l]^2. \end{aligned}$$

As in the proof of Lemma 66 we denote the operator  $\Omega := \|g\|^{-1} \sum_{k_j} |g_{k_j}| \mathcal{G}_{k_j}$  for all  $1 \leq j \leq d$  and write  $E_l = E_{l_1} \otimes \dots \otimes E_{l_d}$ . Then

$$\begin{aligned} &\|g\|_1^{2d} \sum_{l_1}^d \prod_{j=1}^d \text{Tr}[\Omega(|0\rangle\langle 0|) E_j] \text{Tr}[A \hat{\rho}_l]^2 \\ &= \|g\|_1^{2d} \sum_{l_1}^d \text{Tr}[\Omega(|0\rangle\langle 0|)^{\otimes d} E_l] \text{Tr}[A \hat{\rho}_l]^2 \\ &\leq \|g\|_1^{2d} \max_{\sigma} \text{Tr} \left[ \Omega(\sigma)^{\otimes d} \sum_{l_1}^d \text{Tr}[A \hat{\rho}_l]^2 E_l \right] \\ &\leq \|g\|_1^{2d} \max_{\sigma} \text{Tr}[\Omega(\sigma)^{\otimes d} \Gamma] = \|g\|_1^{2d} \|\Gamma\|_{\infty}, \end{aligned} \quad (9.31)$$

where  $\Gamma = \sum_L \text{Tr}[A\hat{\rho}_L]^2 E_L$ . Finally, we can evaluate an upper bound of the summation in Eq. (9.29) analytically as

$$\begin{aligned} \text{Var}[U_N] &\leq \|g\|_1^{2m} \|\Gamma\|_\infty \frac{1}{\binom{N}{m}} \sum_{d=1}^m \binom{m}{d} \binom{N-m}{m-d} \\ &= \|g\|_1^{2m} \|\Gamma\|_\infty \left[1 - \frac{((N-m)!)^2}{N!(N-2m)!}\right]. \end{aligned}$$

Given the factorial formula is of  $\mathcal{O}(1/N)$  with  $N \equiv N_s$ , the above bound implies that the number of samples needed to predict polynomial functions of degree  $m$  with accuracy  $\epsilon > 0$  scales as  $\mathcal{O}(\|g\|_1^{2m}/\epsilon^2)$ .  $\square$

One can similarly apply a median of means estimator to enable simultaneous prediction of many non-linear properties. We detail an explicit protocol in Section 9.3.3 for simultaneously estimating many local Rényi entropies. Note that the measurement overhead  $\|g\|_1^{2m}$  grows with the  $2m^{\text{th}}$  power of the quasiprobability norm. This is consistent with our effective construction of  $m$  copies of the original noisy circuit which leads to an effective  $m$ -fold increase in the number of noisy gates.

### 9.3.3 Classical post-processing algorithms

In this section we summarize reconstruction algorithms for the practically pivotal scenario of Pauli basis measurements.

#### Algorithm 1: Estimating local observables via Pauli strings

For an arbitrary observable  $O$  we can calculate the estimator from Eq. (9.8) as

$$\text{Tr}[O\hat{\rho}_{k,l}] = \|g\|_1 \text{sign}(g_k) \text{Tr}\left[OC_E^{-1}(E_l)\right]. \quad (9.32)$$

In the idealized measurement case,  $E$  simplifies as detailed in Eq. (9.1) and our classical shadow is then a collection of  $N_s$  measurement outcomes  $b \in \{0,1\}^N$  and the corresponding single-qubit Pauli measurement basis defined by the single-qubit rotations  $Q_k \in \mathcal{Q}$ .

For the case  $O = P$  with  $P$  a  $q$ -local Pauli string, it admits the product form  $P = \otimes_{i \in Q} P^{(i)}$ , while the snapshot  $C_E^{-1}(E_l)$  similarly is of a product form via Eq. (9.3).

Note also that we use the index set  $Q$  to abbreviate the set of qubits on which  $P$  acts non-trivially and  $|Q| = q$ . Thus we obtain the trace as

$$\text{Tr}\left[PC_E^{-1}(E_l)\right] = \prod_{i \in Q} \text{Tr}\left[P^{(i)}\left(3(Q_l^{(i)})^\dagger |b^{(i)}\rangle \langle b^{(i)}| Q_l^{(i)} - \mathbb{1}\right)\right]. \quad (9.33)$$

Above we have used that the trace of a tensor product simplifies to a product of traces and that on every qubit  $i$  for which  $P^{(i)} \equiv \mathbb{1}$  the single-qubit expression evaluates to

$$\text{Tr}\left[P^{(i)}\left(3(Q_l^{(i)})^\dagger |b^{(i)}\rangle \langle b^{(i)}| Q_l^{(i)} - \mathbb{1}\right)\right] = 1, \quad (9.34)$$

as the Pauli operators are traceless. The expression in Eq. (9.33) evaluates to  $\{\pm 3^q, 0\}$  as we explain now. The expression evaluates to  $\pm 3^q$  if the measurement bases defined by  $Q_l^{(i)}$  are the same as the single qubit Pauli matrices  $P^{(i)}$  on the qubits  $i \in Q$ . Indeed, it can be directly seen that for one factor in Eq. (9.33) one has

$$\text{Tr}\left[P^{(i)}(3(Q_l^{(i)})^\dagger|b^{(i)}\rangle\langle b^{(i)}|Q_l^{(i)} - \mathbb{1})\right] = 3\langle b^{(i)}|Q_l^{(i)}P^{(i)}(Q_l^{(i)})^\dagger|b^{(i)}\rangle. \quad (9.35)$$

The sign is then determined by the bits  $b^{(i)}$  in the bit string  $b$ , i.e., it is negative if the Hamming weight of the bit string is odd on the qubits in  $Q$ . Otherwise, if the measurement bases appearing in  $Q_l$  are not compatible with  $P$  on the qubits in  $Q$ , then the above expression evaluates to zero. Thus we obtain the simplified estimator as

$$\text{Tr}[P\hat{\rho}_{k,l}] = \|g\|_1 3^q \text{sign}(g_k) f(b, Q_l), \quad (9.36)$$

where  $f(b, Q_l) \in \{\pm 1, 0\}$ . More precisely, the function  $f(b, Q_l)$  will result in 0 if the measurement bases in  $Q_l$  are incompatible with  $P$  and  $\pm 1$  if the measurement bases are compatible with  $P$  while the sign is determined by the bit string  $b$ . The reconstruction algorithm thus takes the classical shadow data as the collection of bit strings and Pauli measurement bases  $\{b_k, Q_k\}_{k=1}^{N_s}$ , as well as the Pauli observable  $P$ , and calculates the values of  $f(b, Q_k)$ . The algorithm has runtime  $\mathcal{O}(qN_s)$ . As we reconstruct  $q$ -local Pauli observables, we can significantly reduce the sample variance by using so-called light-cone arguments [409, 410]. In Fig. 9.2 we illustrate the light cone that an observable creates with respect to the ideal unitary circuit  $\mathcal{U}_{\text{circ}}$ . To simplify our arguments, we will assume local noise models to guarantee the same light cone is valid for all gate sequences  $\mathcal{G}_k$ . However, the extension to non-local noise models is straightforward [409].

### Algorithm 2: Improved estimation of local observables via light cones

The idea is that for each gate that is not within the light cone of the observable  $P$  we can “turn off” PEC, thereby not wasting the measurement budget on mitigating noisy gates that do not affect our observable of interest. Given a  $q$ -local Pauli string  $P$  we define the set of indices of all gates in the light cone of the observable as

$$\mathcal{I} = \{l \mid \mathcal{U}_l \text{ is in the light cone of } P\}. \quad (9.37)$$

Then, one can simply use Algorithm 1 with a modified set of quasiprobabilities from Definition 62 as

$$\|\tilde{g}\|_1 = \prod_{l \in \mathcal{I}} \|\gamma^{(l)}\|_1, \quad \text{and} \quad \text{sign}(\tilde{g}_k) = \prod_{l \in \mathcal{I}} \text{sign}(\gamma_{k_l}^{(l)}). \quad (9.38)$$

The algorithm has the same asymptotic runtime  $\mathcal{O}(qN_s)$  as Algorithm 1 and only incurs a negligible preprocessing time to determine the index set  $\mathcal{I}$  specifically for

each observable  $P$ . The measurement cost  $\|\tilde{g}\|_1$  is thus determined by the number of gates in the light cone of  $P$  rather than by the total number of gates  $\nu$ . Imagine, for example, noisy quantum gates with  $\|\gamma^{(l)}\|_1 = 1 + p$ . The measurement cost is determined by  $(1 + p)^{|Z|}$  as opposed to the worst case  $\|\tilde{g}\|_1 = (1 + p)^\nu$  where  $\nu$  is the total number of noisy gates as detailed in Ref. [409]. A significant advantage of this procedure is that it does not require one to modify the experimental protocol, i.e., the noise in all gates can be mitigated in the shadows.

### Algorithm 3: Local purities

Here we consider the problem of estimating Rényi entropies via the purities  $\text{Tr}(\varrho_Q^2)$  as  $R_Q := -\log \text{Tr}(\varrho_Q^2)$ , where  $\varrho_Q$  is the reduced density matrix of the subsystem  $Q$ . In order to simplify the notation, we abbreviate the indices of PEC snapshots as  $\hat{\varrho}_i := \hat{\varrho}_{k,l}$  with  $i = (k, l)$ . Given a subsystem as a set of qubits  $Q = \{q_1, \dots, q_m\}$  we obtain an unbiased estimator for the respective purity as

$$\text{Tr}[\hat{\varrho}_Q^2] := \text{Tr}[\text{SWAP}_{Q,Q'} \hat{\varrho}_i \otimes \hat{\varrho}_j] = \|g\|_1^2 \text{sign}(g_i) \text{sign}(g_j) f(i, j, Q), \quad (9.39)$$

where  $i \neq j$ . This induces an unbiased estimator for the Rényi entropies via  $\hat{R}_Q := -\log \text{Tr}(\hat{\varrho}_Q^2)$ . Here  $\text{SWAP}_{Q,Q'}$  swaps all pairs of qubits  $q_k$  and  $q_{k+N}$  in the system of  $2N$  qubits in  $\hat{\varrho}_i \otimes \hat{\varrho}_j$ . The factor  $f(i, j, Q)$  can be computed analytically using that the snapshots are of product form  $\hat{\varrho}_i = \bigotimes_{q=1}^N \hat{\varrho}_i^{(q)}$  as

$$f(i, j, Q) = \prod_{q \in Q} \text{Tr}[\text{SWAP}_{\hat{\varrho}_i^{(q)} \otimes \hat{\varrho}_j^{(q)}}], \quad (9.40)$$

where SWAP is the standard 2-qubit SWAP operator. Here we have used that traces for qubits not in subsystem  $Q$  evaluate to 1 and that the trace of a tensor product simplifies to a product of traces. Further, we can evaluate analytically the expression

$$\text{Tr}[\text{SWAP}_{\hat{\varrho}_i^{(q_k)} \otimes \hat{\varrho}_j^{(q_k)}}], \quad (9.41)$$

as it only involves two qubits. If the single-qubit measurement bases  $Q_k^{(q)}$  and  $Q_l^{(q)}$  in the snapshots are not identical, then the expression will evaluate to  $\frac{1}{2}$ . If the measurement bases are identical, then the expression will evaluate to 5 given the measurement outcome bits are identical. Otherwise it will be  $-4$  for not identical measurement outcome bits. The function  $f(i, j, Q)$  is then just a product of these values evaluated for all qubits in the set  $Q$ . The algorithm simply iterates over all distinct pairs of snapshots and evaluates  $f(i, j, Q)$ . We further multiply each snapshot outcome by the corresponding signs  $\text{sign}(g_i) \text{sign}(g_j)$  and with the squared norm  $\|g\|_1^2$ . Finally, we compute the median of means of these individual outcomes. The algorithm has a runtime of  $\mathcal{O}(|Q|N_s^2)$ . Note that the runtime is linear in the subsystem size  $|Q|$  and quadratic in the number of shots  $N_s$ . For sufficiently small subsystems and large number of shots it might be preferred to use the exponentially  $\mathcal{O}(4^{|Q|}N_s)$  scaling algorithm of Ref. [353].

## 9.4 Further error mitigation techniques

### 9.4.1 Error extrapolated shadows

The key idea behind zero-noise extrapolation (ZNE) resides in the possibility of increasing the noise in the circuit and extrapolating expected values back to the case of zero noise. The approach is intuitive to use, requires less resources than PEC but yields a biased estimator. A non-trivial aspect, however, is choosing the correct model function for the extrapolation which has been extensively discussed in the literature [32,33,402]. Typical models include a linear function, an exponential function or a linear combination of multiple exponentials. We consider extrapolation as a means for mitigating errors in properties extracted from classical shadows. The key ingredient we require is the ability to generate a collection of shadows at different noise strengths  $S(q_{p_0}, N_s), \dots, S(q_{p_n}, N_s)$  such that  $p_k \geq p_0$  and  $p_0$  is the device's lowest possible noise strength. These shadows enable us to extract the expected values  $f_m(p) = \text{Tr}[O_m q_p]$  at a given noise level  $p$ . By fitting a suitable model function  $\tilde{f}_m(p)$ , e.g., a linear model, to this data set, we can approximate ideal properties of the state using an extrapolation via the limit

$$\text{Tr}[O_m q_{\text{id}}] \approx \lim_{p \rightarrow 0} \tilde{f}_m(p). \quad (9.42)$$

While we could certainly leverage existing techniques for physically increasing noise rates in a circuit to obtain  $S(q_p, N_s)$  [32, 33], we can also exploit the power and flexibility of the previously derived PEC shadows approach. Instead of considering the quasiprobability representation of the ideal circuit in Definition 62 we can rather decompose the noise-boosted circuits as  $\mathcal{U}_{\text{circ}}(p) = \sum_{\underline{k}} g_{\underline{k}}(p) \mathcal{G}_{\underline{k}}$  with non-negative probabilities  $g_{\underline{k}}(p)$ . For example, in the case of local depolarising noise, the circuit variants  $\mathcal{G}_{\underline{k}}$  are simply obtained by randomly inserting Pauli  $X$ ,  $Y$  or  $Z$  operations after each noisy gate with probabilities  $p - p_0$ . Furthermore, Lindblad-Pauli learning directly gives access to the continuous set of circuits  $\mathcal{U}_{\text{circ}}(p)$  [405].

Let us now state a corollary to Theorem 64 that allows us to obtain the shadows of error boosted states  $q_p := \mathcal{U}_{\text{circ}}(p)|0\rangle\langle 0|$ .

**Corollary 69.** *Consider the parametric quasiprobability decomposition  $G$  as noise-boosted circuits  $\mathcal{U}_{\text{circ}}(p)$  with  $p \geq p_0$ . The PEC shadows from Theorem 64 which are given by  $H := \{(g_{\underline{k}}(p), \mathcal{G}_{\underline{k}}, E_l)\}_{\underline{k}, l}$  result in the simplified snapshots as  $\hat{q}_p := \hat{q}_{p,(\underline{k}, l)} = C_E^{-1}(E_l)$  due to  $\text{sign}(g_{\underline{k}}(p)) = +1$  and  $\|g(p)\|_1 = 1$ . It follows that  $\hat{q}_p$  is an unbiased estimator of the noise-boosted density matrix  $q_p$  such that  $\mathbb{E}[\hat{q}_p] = \mathbb{E}_{\underline{k}, l}[\hat{q}_{p,(\underline{k}, l)}] = q_p$ .*

A significant advantage in boosting the noise via  $p \geq p_0$  rather than reducing it is that now every quasiprobability is non-negative  $g_{\underline{k}}(p) \geq 0$  and thus we do not incur a measurement overhead as described in Lemma 66 via  $\|g(p)\|_1 = 1$ . Nevertheless,



the extrapolated value indeed suffers from an increased variance which implies an increased number of samples. A more detailed discussion can be found in Ref. [32].

Note that the above scheme can be applied beyond the estimation of expectation values. For instance, one can in principle use shadows to reconstruct partial density matrices  $\hat{\rho}_p$  at different noise strengths  $p$  and apply ZNE to individual matrix entries. However, note that ZNE might require different kinds of model functions  $f(p)$  for different properties, e.g., non-linear models for predicting non-linear properties of the state. In contrast, the great advantage of PEC shadows is that it provides an unbiased estimator for the entire quantum state.

### 9.4.2 Symmetry verified shadows

Symmetry verification is another leading quantum error mitigation technique [411, 412]. It exploits that often the ideal states to be prepared  $\rho_{\text{id}}$  are pure states that obey certain problem-specific symmetry group operations which are described by  $S \in \mathcal{S}$ . The fact that the ideal state is symmetric then implies that it “lives in” the subspace defined by the projection operator

$$\Pi_{\mathcal{S}} = \frac{1}{|\mathcal{S}|} \sum_{S \in \mathcal{S}} S,$$

which satisfies  $\Pi_{\mathcal{S}}^2 = \Pi_{\mathcal{S}}$ .

Given a noisy state  $\rho$ , one might be able to measure the above symmetries (in fact their generators are sufficient) via, e.g., Hadamard-test circuits, and retain only circuit runs that produce the correct symmetry outcomes [32,413]. Such post-selection projects the noisy state back into this symmetry subspace producing the effective output state as  $\rho_{\text{sym}} = \Pi_{\mathcal{S}} \rho \Pi_{\mathcal{S}} / \text{Tr}(\Pi_{\mathcal{S}} \rho)$ . We can apply conventional shadow tomography to this symmetry-verified state  $\rho_{\text{sym}}$  thereby effectively obtaining error mitigated shadows, i.e., an unbiased estimator of  $\rho_{\text{sym}}$ . The sampling overhead of this post-selection technique is  $\text{Tr}(\Pi_{\mathcal{S}} \rho)^{-1}$  the inverse of the fraction of circuit runs that pass the symmetry verification process.

Instead of post-selection, we can also perform symmetry verification at the post-processing stage. Suppose we are interested in the expectation value of the ideal state with respect to the target observable  $O$ . The target expectation value can be written as

$$\text{Tr}(O \rho_{\text{sym}}) = \frac{\text{Tr}(O \Pi_{\mathcal{S}} \rho \Pi_{\mathcal{S}})}{\text{Tr}(\Pi_{\mathcal{S}} \rho)} = \frac{1}{|\mathcal{S}|} \frac{\sum_{S, S' \in \mathcal{S}} \text{Tr}(S O S' \rho)}{\sum_{S \in \mathcal{S}} \text{Tr}(S \rho)}.$$

Conventional shadow tomography can be used well in practice for estimating  $S O S'$  and  $S$  for all  $S, S' \in \mathcal{S}$  when the symmetries are sufficiently local, i.e., they are supported on at most weight- $s$  Pauli operators. Then, given a Pauli observable  $O$  of weight at most  $q$ , the effective observable  $S O S'$  is then at most of weight- $(2s+q)$ . However, the

sample complexity of conventional shadows with Pauli measurements grows exponentially with the weight of the Pauli string and it is thus crucial that the total weight  $2s+q$  is reasonably small.

For example, a typically used symmetry in fermionic simulation is the fermionic particle number parity which is, however, usually a high-weight operator for standard encodings such as the Jordan-Wigner encoding. Nevertheless, one can use encodings that come with inherent local symmetry generators like Majorana loop encodings [414], or even implement the circuit using some small quantum codes with local stabilisers [415]. However, even if these symmetry *generators* are local, the number of generators will scale with the number of qubits. Consequently some symmetries they generate are still of high weight. Hence, in order to efficiently use shadow techniques, we can apply verification using a constant number of local symmetry generators, such that the highest-weight symmetry that can be generated is upper-bounded by some constant.

We also note that the sampling cost can be reduced when the target observable  $O$  commutes with the symmetry projector  $\Pi_S$  which is often the case in typical applications. In such a scenario,  $\Pi_S O \Pi_S = O \Pi_S$  and thus

$$\mathrm{Tr}(O \rho_{\mathrm{sym}}) = \frac{\mathrm{Tr}(O \Pi_S \rho)}{\mathrm{Tr}(\Pi_S \rho)} = \frac{\sum_{S \in \mathcal{S}} \mathrm{Tr}(S O \rho)}{\sum_{S \in \mathcal{S}} \mathrm{Tr}(S \rho)}. \quad (9.43)$$

This way, the effective observables we need to estimate from shadows are  $SO$  and  $S$  for all  $S \in \mathcal{S}$  which have a reduced weight  $s + q$  compared to the previous  $2s + q$ .

## 9.5 Applications

In this Section we showcase how our approach can effectively extend the reach of noisy quantum computers and explore its practical applications. For instance, noisy quantum computers in either the late NISQ era or in the early fault-tolerance era will enable us to simulate the time evolution of quantum states or to prepare ground or eigenstates [351,390,391,393,394,416,417]. Our approach can then be used to accurately and efficiently extract a large number of properties of these states provided that the noise rates are reasonable, i.e., the sample overhead  $\|g\|_1$  is moderate.

### 9.5.1 Ground-state preparation

We first consider a spin-ring Hamiltonian as

$$\mathcal{H} = \sum_{k \in \mathrm{ring}(N)} \omega_k Z_k + J \vec{\sigma}_k \cdot \vec{\sigma}_{k+1}, \quad (9.44)$$

with coupling  $J = 0.3$  and on-site interaction strengths uniformly randomly generated in the range  $-1 \leq \omega_k \leq 1$ . This spin problem is relevant in condensed-matter physics

in understanding many-body localisation [418] but is challenging to simulate classically for large  $N$  [419, 420]. A broad range of techniques are available in the literature for finding eigenstates of such quantum Hamiltonians using near-term or early fault-tolerant quantum computers [351, 390, 391, 393]. Here we prepare the ground state of this model using a variational Hamiltonian ansatz in Fig. 9.3 (left) of  $l = 5$  layers on 12 qubits and assume local depolarising noise for two-qubit gates. More precisely, we assume that two-qubit gates undergo depolarising noise with a probability of  $p$  while single qubit gates, including recovery operations of PEC, can be implemented with negligible error. The single-qubit depolarising noise channel is given by

$$\Phi_p(\varrho) := (1 - p)\varrho + \frac{p}{3}(X\varrho X + Y\varrho Y + Z\varrho Z). \quad (9.45)$$

The inverse of this channel can be calculated as

$$\Phi_p^{-1}(\varrho) = \gamma_0\varrho + \gamma_1 X\varrho X + \gamma_2 Y\varrho Y + \gamma_3 Z\varrho Z, \quad (9.46)$$

where the coefficients are given by

$$\underline{\gamma} = \|\gamma\|_1 \left( \frac{3-p}{2p+3}, \frac{-p}{2p+3}, \frac{-p}{2p+3}, \frac{-p}{2p+3} \right), \quad (9.47)$$

and  $\|\gamma\|_1 = (3 + 2p)/(3 - 4p)$ .

### Ground state energies via PEC

Fig. 9.3 (middle) shows the error in the ground state energy estimated using conventional shadows (dashed blue and dashed red) and PEC shadows (solid blue, solid red) for an increasing number of shots  $N_s$ . Our ansatz circuit would ideally prepare the ground state  $\varrho_{\text{id}}$  but due to gate noise we actually prepare the noisy state  $\varrho$ . Thus, conventional shadows (dashed blue, dashed red) converge to a plateau corresponding to the biased energy  $\text{Tr}(\varrho\mathcal{H})$  (solid gray). This bias is significantly increased as we increase the circuit error rate from  $\zeta = 0.12$  to  $\zeta = 0.24$  (dashed blue vs. dashed red), where  $\zeta = 2p\nu$  is the per gate error rate  $p$  times the number of noisy entangling gates  $\nu = 60$ . In contrast, PEC shadows converge to the true energy  $\text{Tr}(\varrho_{\text{id}}\mathcal{H})$  in standard shot-noise scaling  $O(1/\sqrt{N_s})$ .

### Local properties via PEC

Besides Hamiltonian energy estimation, which is one of the typical subroutines in quantum computing, there is also significant value in simultaneously determining many local observables' expectation values. For example, the rich information from classical shadows can be used to significantly improve parameter training or to directly estimate Hamiltonian energy gaps by the use of efficient classical post-processing [393, 394]. In Fig. 9.3 (right), we plot errors when simultaneously estimating all 3-local Pauli

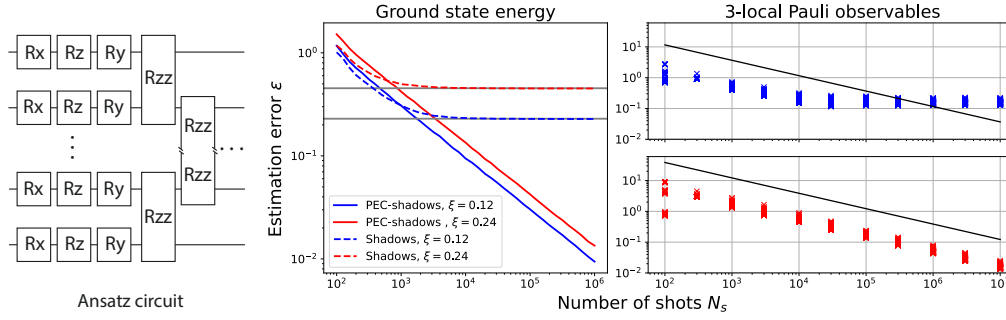


Figure 9.3: (left) A noisy variational Hamiltonian ansatz is used to prepare the ground state given by Eq. (9.44) but our aim is to learn properties of the noise-free state. (middle) Energy estimation errors for different noise strengths with conventional shadows (dashed blue, dashed red) and with PEC shadows (solid blue, solid red). A bias (gray solid lines) is introduced when the ground state energy  $\text{Tr}(\rho\mathcal{H})$  is directly estimated from the noisy quantum state  $\rho$ . Error mitigated shadows are unbiased as they estimate  $\text{Tr}(\rho_{\text{id}}\mathcal{H})$ . Increasing the circuit error rate  $\zeta$  (blue vs. red) increases the bias in standard shadows (dashed blue vs. dashed red) and increases the variance of the error mitigated shadows (solid blue vs. solid red). Each data point is an average over  $10^4$  experiments of a fixed shot budget  $N_s$ . (right) Error in simultaneously estimating all 3-local Pauli strings without (blue) and with (red) error mitigation – only the 200 observables with the highest estimation error are shown and a circuit error rate  $\zeta = 0.6$  is assumed. Errors are significantly below our rigorous bounds from Lemma 66 for PEC shadows but the errors for conventional shadows can be above their respective bounds from Ref. [353] due to the bias (right end of blue). The figure is taken from Ref. [B].

operators for an increasing number of shots  $N_s$ . Fig. 9.3 (red) shows that the errors in PEC shadows are always significantly below the theoretical bounds (black line) from Lemma 66 confirming looseness of the bounds (assuming success probability  $\delta = 10^{-3}$ , and  $M = 3^3 \binom{12}{3} = 5940$ ). Fig. 9.3 (blue) shows that the errors in conventional shadow tomography are below their bounds (with  $\|g\|_1 = 1$ ) only for a small number of shots but then asymptotically reach a plateau due to circuit noise.

### Local properties via extrapolation

We now consider the same task of simultaneously estimating expectation values of Pauli operators but we use error extrapolation. Here we start by generating shadows  $S(\rho_{p_1}, N_s), \dots, S(\rho_{p_n}, N_s)$  at different noise strengths which are subsequently used to

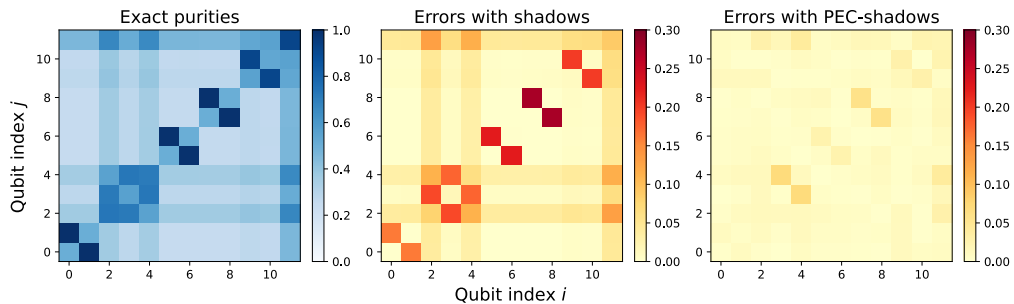


Figure 9.4: A noisy variational Hamiltonian ansatz is used to prepare the ground state of Eq. (9.48) whose ideal, noise-free Rényi entropies  $R_Q$  can be learned with PEC shadows. We plot purities  $\text{Tr}(q_Q^2)$  as a proxy for  $R_Q := -\log \text{Tr}(q_Q^2)$ . (left) Purity heat map in the noiseless case and infinite shot limit. An increasing value indicates that the subsystem  $Q$  is less entangled with the remaining qubits. (middle-right) Absolute error in the purities due to gate noise for a circuit error rate  $\zeta = 0.6$  and due to finite repetition using  $N_s = 10^5$ . (middle) Although the entanglement pattern is approximately recovered with conventional shadows, in some instances we observe substantial errors, i.e., the largest error is 0.27. (right) Absolute errors with PEC shadows are significantly smaller, i.e., the largest error is  $7 \times 10^{-2}$  but this figure could be further reduced by increasing  $N_s$ . The figure is taken from Ref. [B].

compute the noisy Pauli expectation values. Fig. 9.5 shows 10 examples of expected values (crosses) as a function of noise strength and the respective linear models we fit (dashed lines). The intercept of the fitted model (dashed lines) is our estimate of the exact expected value (disks) and is indeed reasonably close in the example. While ZNE has been very effective and typically has a lower measurement overhead than PEC, it is generally biased.

### 9.5.2 Error mitigated estimation of entanglement entropies

Finally, we consider an application for which classical shadows are a primary enabler but for which error mitigation techniques have been less explored [32]. As opposed to studying entanglement properties or verifying the presence thereof in mixed quantum states [290, 421–423], here our primary goal is to extend the reach of noisy quantum computers. We aim to study entanglement properties of ideally pure states which are prepared by quantum algorithms, such as phase estimation or the variational quantum eigensolver. For example, near-term quantum computers will enable us to prepare eigenstates [351, 390, 391, 393] of quantum Hamiltonians and error mitigated entangle-

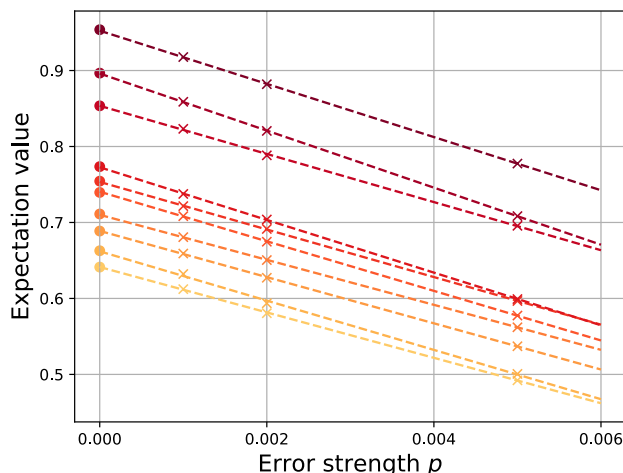


Figure 9.5: Simultaneously estimating all 3-local Pauli operators using error extrapolated shadows. We estimated noisy expected values (crosses) from shadows of size  $N_s = 10^7$  by increasing the native depolarising error rate  $p = 10^{-3}$  to higher levels  $\{2 \times 10^{-3}, 5 \times 10^{-3}\}$  by randomly sampling noisy circuit variants. Using a linear model function we then extrapolate to zero noise in order to obtain an error mitigated expectation value close to the ideal ones (disks). The figure is taken from Ref. [B].

ment measures can be used for, e.g., characterizing phase transitions. Similarly, one could simulate the time evolution of a collision of two molecules with an early fault-tolerant quantum computer and investigate how entanglement builds up across the individual subsystems. Furthermore, efficiently characterising many local correlations in a state can be used to train models for density functional theory in order to obtain accurate classical simulations [424].

Consider the Heisenberg chain

$$\mathcal{H} = \sum_k J_k \vec{\sigma}_k \cdot \vec{\sigma}_{k+1}, \quad (9.48)$$

with uniformly random  $-1 \leq J_k \leq 1$ . Its ground state is prepared with a variational Hamiltonian ansatz of  $l = 8$  layers on 12 qubits. This system was used in Ref. [353] to illustrate the power of classical shadows in predicting entanglement entropies. However, the ground state was approximated by a set of noise-free singlet states [425, 426] whereas we assume a noisy quantum computer is used for state preparation. We use PEC shadows to extract purities  $\text{Tr}(\rho_Q^2)$  for all single and two-qubit subsystems  $Q$ . These purities then define Rényi entropies as  $R_Q := -\log \text{Tr}(\rho_Q^2)$ . In Fig. 9.4, we plot the exact purities in the noiseless case. Here disjoint blocks involving two qubits

confirm that the ground state could be approximated by a tensor product of noise-free singlet states. Fig. 9.4 (middle) shows the errors in estimating local purities using shadows of size  $N_s = 10^5$  for a circuit error rate  $\zeta = 0.6$ . Even for this moderate error rate conventional shadows are significantly impacted by imperfections and result in errors as large as 0.27. Furthermore, for an increasing noise rate all purities converge to a constant value of  $1/d$  where  $d$  is the dimension of the subsystem. In contrast, PEC shadows drastically improve the accuracy in Fig. 9.4 (right) and the largest error is approximately  $7 \times 10^{-2}$  at a number of samples  $N_s = 10^5$ .

### 9.5.3 Further applications

The techniques presented in this Chapter enable us to approximate an unbiased estimator of an ideal noise free state  $\rho_{\text{id}}$ , which can be enabling for a broad range of further practical applications. For example, Ref. [353] proposed that classical shadows with randomised Clifford measurements can be used to predict fidelities, such as the fidelity of  $\rho$  with respect to a known state  $|\psi\rangle$ . One can imagine applications where the fidelity  $\langle\psi|\rho|\psi\rangle$  is not a relevant indicator due to the impact of noise on  $\rho$  and one rather aims to predict  $\langle\psi|\rho_{\text{id}}|\psi\rangle$ , e.g., to quantify how well a variational quantum circuit or phase estimation can prepare a known ground state thereby verifying a circuit structure under the presence of gate noise. Furthermore, the quantum Fisher information (QFI), which is a key quantity in quantum metrology, can be bounded and approximated using classical shadows via techniques of Ref. [359]. Indeed, in certain applications the relevant quantity might not be the QFI of the noisy state  $\rho$  but rather the QFI of the noise-free state  $\rho_{\text{id}}$  which can be approximated with the presented techniques [427].

## 9.6 Discussion and conclusion

In this Chapter we considered the powerful classical shadows methodology which allowed us to obtain an efficient classical representation of a quantum state  $\rho$  and thus to simultaneously predict many of its properties in classical post-processing. A major difficulty concerning near-term and early fault-tolerant quantum computers is that they can only prepare noisy quantum states  $\rho$  from which we would estimate corrupted properties. This challenge motivated the development of quantum error mitigation techniques that allow us to estimate expected values  $\text{Tr}[O\rho_{\text{id}}]$  of observables  $O$  in an ideal noise-free state  $\rho_{\text{id}}$  but with having access only to noisy expected values.

We consider a range of typical quantum error mitigation techniques and generalise them from single expected-value measurements to the case of mitigating errors in classical shadows. We find that PEC is the most well-suited candidate which motivates us to develop a thorough theory of PEC shadows. In the conventional PEC approach

one learns error characteristics of the device and counters them by a probabilistic implementation of the inverse noise channel. Therefore the only source of noise is due to a possibly imperfect knowledge of gate-error characteristics and due to finite circuit repetition. Under the assumption that the error model of the quantum device has been appropriately learned such that a quasiprobability representation is known, we prove that PEC shadows are an unbiased estimator of the ideal state  $\rho_{\text{id}}$ . We additionally prove the following rigorous performance guarantees. First, we prove bounds on the number of samples required to simultaneously predict many linear properties of the ideal quantum state  $\rho_{\text{id}}$ . Second, the fact that we use noisy quantum circuits to predict ideal properties manifests in a multiplicative measurement overhead. This overhead is identical to the cost of the conventional PEC approach and grows exponentially with the number of noisy gates. Third, we prove rigorous sample complexities for predicting non-linear properties of the ideal states. We note that our results are completely general and apply to any shadow ensemble  $E$  via Eq. (9.1) and to any linear or non-linear property of the quantum state. Furthermore, we provide practical post-processing protocols for the pivotal scenario of randomised measurements in Pauli bases. Finally, we demonstrate in numerical simulations the usefulness of PEC shadows and error extrapolated shadows, and conclude that these techniques may be instrumental in practical applications of near-term and early fault-tolerant machines.





# Summary and Outlook

This thesis was concerned with different aspects of quantum information science from the perspective of foundational problems as well as the perspective of possible applications for quantum technology.

To summarize, we have first discussed the structure of quantum measurements from the viewpoint of simulability. We introduced a minimal scheme that allows for the certification of irreducible measurements by only taking into account the observed correlations. For this semi-device independent scenario, we derived a family of correlations that allow for the certification of an irreducible three-outcome measurement on a qubit and analyzed their noise robustness. As in this case irreducibility implies that the measurement was nonprojective, our scheme could be of interest for quantum random number generation. More generally, it would be highly desirable to find a systemic approach to construct families of distributions in a minimal scenario that are able to certify the irreducibility of a measurement.

Second, we considered the quantum measurement problem and asked for the compatibility of the universality of the unitary time evolution with the realization of a partially observed measurement. We introduced the concept of "relative event by incomplete information" and showed that, if combined with locality and no superdeterminism, it is in conflict with the predictions of quantum theory. An interesting extension would be to consider this quantum correlation sets with incomplete information beyond bipartite cases. The existence of a multipartite all-versus-nothing-like proof would be an interesting open question for future research.

We proceeded by deriving Bell inequalities together with quantum correlations that allow for an extremely low detection efficiency and high robustness to noise. These graph-based Bell inequalities were further optimized by using a symmetrized variant of the Gilbert's algorithm. Our analysis relied on optimizing a Bell inequality with respect to a particular quantum point, that was constructed from a state-independent contextuality set. Starting with a quantum point that is not associated to a state-independent contextuality set might yield Bell tests that allow for even smaller detection efficiencies.

Then, we developed methods that allow for rigorous statements on the statistical significants of experiments that demonstrate the phenomenon of activation of nonlocal correlations. We presented a technique for the construction of a suitable confidence polytope and an efficient algorithm to determine the correlation class of a quantum

state. Their combination allowed us to compute the number of state preparations that is necessary to demonstrate that the targeted initial state, which is intended to be activated, is Bell-local. Here it would be interesting to know, whether this number is in reach for near-future experimental setups, such that a concrete experimental implementation could be devised.

Furthermore, we presented an iterative method for the computation of maximally resourceful quantum states. We illustrated our approach for the special case of the geometric measure, allowing us to identify interesting quantum states, discover novel absolutely maximally entangled states, and characterize highly entangled subspaces, which may be useful for information processing. We further demonstrated the universality of the algorithm for various other quantifiers, yielding novel forms of correlations in the triangle network. Concerning future projects, the algorithm could be used to find new absolutely maximally entangled states for cases where the existence is still open, e.g., for systems consisting of more than five quhex.

Subsequently, we discussed the real eigenstructure of regular simplex tensors. We gave a full characterization for the case of an arbitrary number of modes and local dimension 2. In addition, we analyzed the robustness of the eigenvectors with respect to the tensor power iteration. Our findings may shed light on the problem of characterizing those symmetric tensors, which do have repelling eigenvectors, or whose normalized eigenvectors are given by the vectors of the underlying frame.

Finally, we introduced scalable methods for simultaneously predicting many expectation values of a multi-qubit system. We presented a formulation of shadow tomography with generalized measurements, which offers an interpretation as the least-squares estimator. In addition, we described how symmetries can be incorporated, measurements can be optimized towards a targeted set of observables, and errors in the measurement phase can be mitigated. Moreover, we also explained how errors during the preparation phase can be taken into account. For this, we devised a generalization of error mitigation protocols and applied them to generalized classical shadows. For the error mitigation scheme of probabilistic error cancellation, we developed a thorough theoretical framework and derived corresponding sample complexities. Further, we considered the error mitigation techniques of zero-noise extrapolation and symmetry verification and accompanied our theoretical complexity bounds by numerical simulations.

# Acknowledgments

First and foremost, I am grateful to my supervisor Professor Otfried Ghne, for his patience, his encouragement and especially for the freedom to work on whatever it was that caught my interest. This particularly includes opportunities to work on interesting projects and to meet with many interesting scientists all over the world.

A special thanks goes to Dr. Chau Nguyen with whom I collaborated in most of the projects that are contained in this thesis. Thank you for taking all the time to discuss with me and to give me various advices.

A special thanks also goes to Dr. Matthias Kleinmann who sparked my interest in foundational problems of quantum mechanics and philosophy. Our inspiring discussions during my bachelor studies have certainly contributed to my enthusiasm in research.

I would like to thank the current members in the group for providing a great working environment. Thanks also go to Jan-Lennard Bnzel, Sophia Denker, Carlos de Gois, Kiara Hansenne, Satoya Imai, Ties Ohst, Lisa Weinbrenner and Benjamin Yadin, who have proof-read large parts of this thesis.

I am deeply grateful to my parents, to my grandparents and to my family for always giving me the support that I needed. Here I also have to thank my aunt Heike, who had the unpleasant task to check the whole thesis for grammar mistakes.

In addition, I acknowledge various support from the House of Young Talents Siegen and their scholarship program.



# List of publications

- [A] J. Steinberg, H. C. Nguyen and O. Gühne  
*Mitigating measurement errors in classical shadows*  
in preparation
- [B] H. Inane, J. Steinberg, S. Cai, H. C. Nguyen and B. Koczor  
*Quantum Error Mitigated Classical Shadows*  
arXiv: 2305.04956 (2023)
- [C] J. Steinberg, H. C. Nguyen and M. Kleinmann  
*Certifying activation of quantum correlations with finite data*  
arXiv: 2305.03748 (2023)
- [D] J. Steinberg and O. Gühne  
*Maximizing the geometric measure of entanglement*  
arXiv: 2210.13475 (2022)
- [E] H. C. Nguyen, J. L. Bönsel, J. Steinberg and O. Gühne  
*Optimising shadow tomography with generalised measurements*  
Phys. Rev. Lett. **129**, 220502 (2022), arXiv: 2205.08990
- [F] Z.-P. Xu, J. Steinberg, J. Singh, A. J. López-Tarrida,  
J. R. Portillo and A. Cabello  
*Graph-theoretic approach to Bell experiments with low detection efficiency*  
Quantum **7**, 922 (2023), arXiv: 2205.05098
- [G] Z.-P. Xu, J. Steinberg, H. C. Nguyen and O. Gühne  
*No-go theorem based on incomplete information of Wigner about his friend*  
Phys. Rev. A **107**, 022424 (2023), arXiv: 2111.15010
- [H] A. Czaplinski, T. Raasch and J. Steinberg  
*Real eigenstructure of regular simplex tensors*  
Adv. Appl. Math. **148**, 102521 (2023), arXiv:2203.01865
- [I] J. Steinberg, H. C. Nguyen and M. Kleinmann  
*Minimal scheme for certifying three-outcome qubit measurements in the prepare-and-measure scenario*  
Phys. Rev. A **122**, 130404 (2019), arXiv:1812.09216

- [J] J. Steinberg  
*Extensions and Restrictions of Generalized Probabilistic Theories*  
Springer Spektrum BestMaster, ISBN:978-3-658-37580-5 (2021) <sup>1</sup>
- [K] J. Steinberg, H. C. Nguyen and M. Kleinmann  
*Quaternionic quantum theory admits universal dynamics only for two-level systems*  
J. Phys. A **120**, 060502 (2018), arXiv:1707.01050<sup>2</sup>

---

<sup>1</sup>These results were obtained in the authors master thesis

<sup>2</sup>These results were obtained in the authors bachelor thesis.

# Bibliography

- [1] A. Peres. *Quantum Theory: Concepts and Methods*. Fundamental Theories of Physics. Springer, Dordrecht (1995).
- [2] H. J. Leavitt and T. L. Whisler. *Management in the 1980's*. Harvard Business Review **11** (1958).
- [3] A. Einstein, B. Podolsky, and N. Rosen. *Can Quantum-Mechanical Description of Physical Reality Be Considered Complete?* Phys. Rev. **47**, 777 (1935).
- [4] J. S. Bell. *On the Einstein Podolsky Rosen paradox*. Physics **1**, 195–200 (1964).
- [5] H. Wiseman. *The two Bell's theorems of John Bell*. J. Phys. A **47**, 424001 (2014).
- [6] J. F. Clauser, M. A. Horne, A. Shimony, and R. A. Holt. *Proposed Experiment to Test Local Hidden-Variable Theories*. Phys. Rev. Lett. **23**, 880 (1969).
- [7] J. F. Clauser and M. A. Horne. *Experimental consequences of objective local theories*. Phys. Rev. D **10**, 526 (1974).
- [8] A. Aspect, P. Grangier, and G. Roger. *Experimental Realization of Einstein-Podolsky-Rosen-Bohm Gedankenexperiment: A New Violation of Bell's Inequalities*. Phys. Rev. Lett. **49**, 91 (1982).
- [9] G. Weihs, T. Jennewein, C. Simon, H. Weinfurter, and A. Zeilinger. *Violation of Bell's Inequality under Strict Einstein Locality Conditions*. Phys. Rev. Lett. **81**, 5039 (1998).
- [10] D. N. Matsukevich, P. Maunz, D. L. Moehring, S. Olmschenk, and C. Monroe. *Bell Inequality Violation with Two Remote Atomic Qubits*. Phys. Rev. Lett. **100**, 150404 (2008).
- [11] M. Ansmann et al. *Violation of Bell's inequality in Josephson phase qubits*. Nature **461**, 504–506 (2009).
- [12] B. Hensen et al. *Loophole-free Bell inequality violation using electron spins separated by 1.3 kilometres*. Nature **526**, 682–686 (2015).
- [13] M. Giustina et al. *Significant-Loophole-Free Test of Bell's Theorem with Entangled Photons*. Phys. Rev. Lett. **115**, 250401 (2015).



- [14] S. Wiesner. *Conjugate coding*. ACM SIGACT News **15**, 78–88 (1983).
- [15] C. H. Bennett and G. Brassard. *Quantum cryptography: Public key distribution and coin tossing*. Theor. Comput. Sci. **560**, 7–11 (2014).
- [16] W. K. Wootters and W. H. Zurek. *A Single Quantum Cannot be Cloned*. Nature **299**, 802–803 (1982).
- [17] A. K. Ekert. *Quantum cryptography based on Bell's theorem*. Phys. Rev. Lett. **67**, 661 (1991).
- [18] V. Scarani, H. Bechmann-Pasquinucci, N. J. Cerf, M. Dušek, N. Lütkenhaus, and M. Peev. *The security of practical quantum key distribution*. Rev. Mod. Phys. **81**, 1301 (2009).
- [19] D. Mayers and A. Yao. *Quantum cryptography with imperfect apparatus*. Proceedings 39th Annual Symposium on Foundations of Computer Science (IEEE, Los Alamitos, CA) p. 503 (1998).
- [20] D. P. Nadlinger et al. *Experimental quantum key distribution certified by Bell's theorem*. Nature **607**, 682–686 (2022).
- [21] S.-K. Liao et al. *Satellite-to-ground quantum key distribution*. Nature **549**, 43–47 (2017).
- [22] S. McArdle, S. Endo, A. Aspuru-Guzik, S. C. Benjamin, and X. Yuan. *Quantum computational chemistry*. Rev. Mod. Phys. **92**, 015003 (2020).
- [23] R. W. Landauer. *Information Is Physical*. Phys. Today **44**, 23–29 (1991).
- [24] R. P. Feynman. *Simulating physics with computers*. Int. J. Theor. Phys. **21**, 467–488 (1982).
- [25] Y. I. Manin. *Computable and Uncomputable (in Russian)*. Sov. Radio (1980).
- [26] H. Bernien et al. *Probing many-body dynamics on a 51-atom quantum simulator*. Nature **551**, 579–584 (2017).
- [27] J. Zhang et al. *Observation of a many-body dynamical phase transition with a 53-qubit quantum simulator*. Nature **551**, 601–604 (2017).
- [28] F. Arute et al. *Quantum supremacy using a programmable superconducting processor*. Nature **574**, 505–510 (2019).
- [29] M. A. Nielsen and I. L. Chuang. *Quantum Computation and Quantum Information*. Cambridge University Press (2010).

- [30] P. W. Shor. *Polynomial-Time Algorithms for Prime Factorization and Discrete Logarithms on a Quantum Computer*. SIAM J. Sci. Statist. Comput. **26**, 303–332 (1997).
- [31] J. Preskill. *Quantum Computing in the NISQ era and beyond*. Quantum **2**, 79 (2018).
- [32] Z. Cai, R. Babbush, S. C. Benjamin, S. Endo, W. J. Huggins, Y. Li, J. R. McClean, and T. E. O’Brien. *Quantum Error Mitigation*. arXiv:2210.00921 (2022).
- [33] A. Kandala, K. Temme, A. D. Córcoles, A. Mezzacapo, J. M. Chow, and J. M. Gambetta. *Error mitigation extends the computational reach of a noisy quantum processor*. Nature **567**, 491–495 (2019).
- [34] Y. Kim, C. J. Wood, T. J. Yoder, S. T. Merkel, J. M. Gambetta, K. Temme, and A. Kandala. *Scalable error mitigation for noisy quantum circuits produces competitive expectation values*. Nat. Phys. **19**, 752–759 (2023).
- [35] Č. Brukner. *On the Quantum Measurement Problem*. In *Quantum [Un] Speakables II*, pp. 95–117. Springer (2017).
- [36] Č. Brukner. *A No-Go Theorem for Observer-Independent Facts*. Entropy **20**(5), 350 (2018).
- [37] K.-W. Bong, A. Utreras-Alarcón, F. Ghafari, Y.-C. Liang, N. Tischler, E. G. Cavalcanti, G. J. Pryde, and H. M. Wiseman. *A strong no-go theorem on the Wigner’s friend paradox*. Nat. Phys. **16**, 1–7 (2020).
- [38] A. S. Holevo. *Statistical Structure of Quantum Theory*. Lecture Notes in Physics Monographs. Springer, Berlin (2001).
- [39] T. Heinosaari and M. Ziman. *The Mathematical Language of Quantum Theory*. Cambridge University Press (2011).
- [40] P. Busch, M. Grabowski, and P. J. Lahti. *Operational Quantum Physics*. Lecture Notes in Physics Monographs. Springer, Berlin (1995).
- [41] J. von Neumann. *Mathematische Grundlagen der Quantenmechanik*. Springer, Berlin (1932).
- [42] G. Lüders. *Concerning the state-change due to the measurement process*. Ann. Phys. **8**, 322–328 (1950).
- [43] F. Pokorny, C. Zhang, G. Higgins, A. Cabello, M. Kleinmann, and M. Hennrich. *Tracking the Dynamics of an Ideal Quantum Measurement*. Phys. Rev. Lett. **124**, 080401 (2020).
- [44] C. W. Helstrom. *Quantum detection and estimation theory*. J. Stat. Phys. **1**, 231–252 (1969).

- [45] C. A. Fuchs. *Distinguishability and Accessible Information in Quantum Theory*. arXiv:9601020 (1996).
- [46] A. Gleason. *Measures on the Closed Subspaces of a Hilbert space*. Indiana Univ. Math. J. **6**, 885–893 (1957).
- [47] E. Wigner. *Gruppentheorie und ihre Anwendung auf die Quantenmechanik der Atom-spektren*. Vieweg+Teubner, Wiesbaden (1931).
- [48] W. F. Stinespring. *Positive Functions on C\*-Algebras*. Proc. Am. Math. Soc. **6**, 211–216 (1955).
- [49] D. A. Lidar and T. A. Brun. *Quantum Error Correction*. Cambridge University Press (2013).
- [50] M. Schlosshauer. *The Role of Decoherence in Interpretations of Quantum Mechanics*, chap. 8, p. 330. Springer, Berlin (2007).
- [51] H. D. Zeh. *On the Interpretation of Measurement in Quantum Theory*. Found. Phys. **1**, 69–76 (1970).
- [52] W. H. Zurek. *Environment-induced superselection rules*. Phys. Rev. D **26**, 1862 (1982).
- [53] W. H. Zurek. *Decoherence, einselection, and the quantum origins of the classical*. Rev. Mod. Phys. **75**, 715 (2003).
- [54] E. P. Wigner. *Remarks on the Mind-Body Question*. In I. J. Good, ed., *The Scientist Speculates*, pp. 284–302. Heinemann, London (1962).
- [55] D. Deutsch. *Quantum theory as a universal physical theory*. Int. J.Theor.Phys. **24**, 1–41 (1985).
- [56] J. S. Bell. *Against "measurement"*. Phys. World **3**, 33–40 (1990).
- [57] D. Bohm. *Quantum Theory*. Prentice-Hall (1951).
- [58] R. F. Werner. *Quantum states with Einstein-Podolsky-Rosen correlations admitting a hidden-variable model*. Phys. Rev A **40**, 4277 (1989).
- [59] L. Gurvits. *Classical complexity and quantum entanglement*. J. Comp. Syst. Sci. **69**, 448–484 (2004).
- [60] S. Gharibian. *Strong NP-hardness of the Quantum Separability Problem*. Quant. Inf. Comp. **10**, 343–360 (2010).
- [61] G. W. Stewart. *Matrix Algorithms*, vol. 2: Eigensystems. Society for Industrial and Applied Mathematics (2001).

- [62] N. Linden, S. Popescu, and S. Popescu. *On Multi-Particle Entanglement*. Fortsch. Phys. **46**, 567–578 (1999).
- [63] J. Schlienz and G. Mahler. *Description of entanglement*. Phys. Rev. A **52**, 4396 (1995).
- [64] M. Grassl, M. Rötteler, and T. Beth. *Computing local invariants of quantum-bit systems*. Phys. Rev. A **58**, 1833 (1998).
- [65] C. H. Bennett, S. Popescu, D. Rohrlich, J. A. Smolin, and A. V. Thapliyal. *Exact and asymptotic measures of multipartite pure-state entanglement*. Phys. Rev. A **63**, 012307 (2000).
- [66] R. Horodecki, P. Horodecki, M. Horodecki, and K. Horodecki. *Quantum entanglement*. Rev. Mod. Phys. **81**, 865 (2009).
- [67] W. Dür, G. Vidal, and J. I. Cirac. *Three qubits can be entangled in two inequivalent ways*. Phys. Rev. A **62**, 062314 (2000).
- [68] A. Acín, A. Andrianov, L. Costa, E. Jané, J. I. Latorre, and R. Tarrach. *Generalized Schmidt Decomposition and Classification of Three-Quantum-Bit States*. Phys. Rev. Lett. **85**, 1560 (2000).
- [69] V. Scarani and N. Gisin. *Spectral decomposition of Bell's operators for qubits*. J. Phys. A **34**, 6043–6053 (2001).
- [70] P. Horodecki. *Separability criterion and inseparable mixed states with positive partial transposition*. Phys. Lett. A **232**, 333–339 (1997).
- [71] Z.-H. Chen, Z.-H. Ma, O. Gühne, and S. Severini. *Estimating Entanglement Monotones with a Generalization of the Wootters Formula*. Phys. Rev. Lett. **109**, 200503 (2012).
- [72] M. B. Plenio and S. Virmani. *An introduction to entanglement measures*. Quant. Inf. Comput. **7**, 1–51 (2007).
- [73] O. Gühne and G. Tóth. *Entanglement detection*. Phys. Rep. **474**, 1–75 (2009).
- [74] S. A. Hill and W. K. Wootters. *Entanglement of a Pair of Quantum Bits*. Phys. Rev. Lett. **78**, 5022 (1997).
- [75] M. B. Hastings. *Superadditivity of communication capacity using entangled inputs*. Nat. Phys. **5**, 255–257 (2009).
- [76] M. Hayashi, D. Markham, M. Muraio, M. Owari, and S. Virmani. *Bounds on Multipartite Entangled Orthogonal State Discrimination Using Local Operations and Classical Communication*. Phys. Rev. Lett. **96**, 040501 (2006).

- [77] T. J. Osborne and F. Verstraete. *General Monogamy Inequality for Bipartite Qubit Entanglement*. Phys. Rev. Lett. **96**, 220503 (2006).
- [78] B. S. Cirel'son. *Quantum generalizations of Bell's inequality*. Lett. Math. Phys. **4**, 93–100 (1980).
- [79] B. S. Cirel'son. *Some results and problems on quantum Bell-type inequalities*. Hadronic Journal Supplement **8**, 329–345 (1993).
- [80] D. Collins and N. Gisin. *A Relevant Two Qubit Bell Inequality Inequivalent to the CHSH Inequality*. J. Phys. A **37**, 1775 (2004).
- [81] V. Scarani. *Bell Nonlocality*. Oxford Graduate Texts. Oxford University Press (2019).
- [82] A. Fine. *Hidden Variables, Joint Probability, and the Bell Inequalities*. Phys. Rev. Lett. **48**, 291 (1982).
- [83] S. Pironio. *Lifting Bell inequalities*. J. Math. Phys. **46**, 062112 (2005).
- [84] M. Froissart. *Constructive generalization of Bell's inequalities*. Nuov. Cim. B **64**, 241–251 (1981).
- [85] P. H. Eberhard. *Background level and counter efficiencies required for a loophole-free Einstein-Podolsky-Rosen experiment*. Phys. Rev. A **47**, R747(R) (1993).
- [86] L. A. Khalfi and B. S. Tsirelson. *Quantum and quasi-classical analogs of Bell inequalities*. In P. Lahti and P. Mittelstaedt, eds., *Symposium on the Foundations of Modern Physics*, pp. 441–460. World Scientific, Singapore (1985).
- [87] S. Popescu and D. Rohrlich. *Quantum nonlocality as an axiom*. Found. Phys. **24**(3), 379–385 (1994).
- [88] A. Coladangelo and J. Stark. *Unconditional separation of finite and infinite-dimensional quantum correlations*. arXiv:1804.05116 (2018).
- [89] Z. Ji, A. Natarajan, T. Vidick, J. Wright, and H. Yuen. *MIP\* = RE*. arXiv:2001.04383 (2022).
- [90] W. Slofstra. *Tsirelson's problem and an embedding theorem for groups arising from non-local games*. J. Amer. Math. Soc. **33**, 1–56 (2020).
- [91] R. Horodecki, P. Horodecki, and M. Horodecki. *Violating Bell inequality by mixed spin- $\frac{1}{2}$  states: necessary and sufficient condition*. Phys. Lett. A **200**, 340–344 (1995).
- [92] D. M. Greenberger, M. A. Horne, and A. Zeilinger. *Going Beyond Bell's Theorem*. In *Bell's Theorem, Quantum Theory and Conceptions of the Universe*, pp. 69–72. Springer (1989).

- [93] N. D. Mermin. *What's Wrong with these Elements of Reality?* Phys. Today **43**, 9–11 (1990).
- [94] E. Schrödinger. *Discussion of Probability Relations between Separated Systems*. Proc. Cambridge Philos. Soc. **31**, 555–563 (1935).
- [95] H. M. Wiseman, S. J. Jones, and A. C. Doherty. *Steering, Entanglement, Nonlocality, and the Einstein-Podolsky-Rosen Paradox*. Phys. Rev. Lett. **98**, 140402 (2007).
- [96] H. C. Nguyen and K. Luoma. *Pure steered states of Einstein-Podolsky-Rosen steering*. Phys. Rev. A **95**, 042117 (2017).
- [97] D. Cavalcanti and P. Skrzypczyk. *Quantum steering: a review with focus on semidefinite programming*. Rep. Prog. Phys. **80**, 024001 (2017).
- [98] R. Uola, A. C. S. Costa, H. C. Nguyen, and O. Gühne. *Quantum steering*. Rev. Mod. Phys. **92**, 015001 (2020).
- [99] C. Budroni, A. Cabello, O. Gühne, M. Kleinmann, and J.-A. Larsson. *Kochen-Specker contextuality*. Rev. Mod. Phys. **94**(4), 045007 (2022).
- [100] S. Kochen and E. P. Specker. *The Problem of Hidden Variables in Quantum Mechanics*. J. Math. Mech. **17**, 59–87 (1967).
- [101] A. Peres. *Unperformed experiments have no results*. Am. J. Phys. **46**, 745 (1978).
- [102] R. Diestel. *Graph Theory*. Graduate Texts in Mathematics. Springer, Berlin (2017).
- [103] M. Hein, W. Dür, J. Eisert, R. Raussendorf, M. V. den Nest, and H.-J. Briegel. *Entanglement in graph states and its applications*. In G. Casati, D. L. Shepelyansky, P. Zoller, and G. Benenti, eds., *Quantum Computers, Algorithms and Chaos*, vol. 162 of *Proceedings of the International School of Physics "Enrico Fermi"*, pp. 115–218. IOS Press (2006).
- [104] G. Pólya. *Kombinatorische Anzahlbestimmungen für Gruppen, Graphen und chemische Verbindungen*. Acta Math. **68**, 145–254 (1937).
- [105] R. M. Karp. *Reducibility among Combinatorial Problems*. In R. E. Miller, J. W. Thatcher, and J. D. Bohlinger, eds., *Complexity of Computer Computations*, pp. 85–103. Springer, Boston (1972).
- [106] F. Huber, O. Gühne, and J. Siewert. *Absolutely Maximally Entangled States of Seven Qubits Do Not Exist*. Phys. Rev. Lett. **118**, 200502 (2017).
- [107] M. Grötschel, L. Lovász, and A. Schrijver. *The ellipsoid method and its consequences in combinatorial optimization*. Combinatorica **1**, 169–197 (1981).

- [108] D. E. Knuth. *The sandwich theorem*. Electron. J. Combin. **1**, 1–48 (1994).
- [109] M. V. den Nest, J. Dehaene, and B. D. Moor. *Local unitary versus local Clifford equivalence of stabilizer states*. Phys. Rev. A **71**, 062323 (2005).
- [110] Z. Ji, J. Chen, Z. Wei, and M. Ying. *The LU-LC conjecture is false*. Quant. Inf. Comp. **10**, 97–108 (2010).
- [111] W. G. Unruh. *Maintaining coherence in quantum computers*. Phys. Rev. A **51**, 992 (1995).
- [112] S. Lloyd, M. Mohseni, and P. Rebentrost. *Quantum principal component analysis*. Nat. Phys. **10**, 631–633 (2014).
- [113] P. O. Boykin, T. Mor, M. Pulver, V. Roychowdhury, and F. Vatan. *A new universal and fault-tolerant quantum basis*. Inf. Process. Lett. **75**, 101–107 (2000).
- [114] A. Y. Kitaev. *Quantum computations: algorithms and error correction*. Russ. Math. Surv. **52**, 1191–1249 (1997).
- [115] A. Peres. *Reversible logic and quantum computers*. Phys. Rev. A **32**, 3266 (1985).
- [116] I. D. Kivlichan et al. *Improved Fault-Tolerant Quantum Simulation of Condensed-Phase Correlated Electrons via Trotterization*. Quantum **296**, 1–45 (2020).
- [117] Y. Nesterov. *Lectures on Convex Optimization*, chap. Nonlinear Optimization. Optimization and Its Applications. Springer (2018).
- [118] S. Boyd and L. Vandenberghe. *Convex Optimization*. Cambridge University Press (2004).
- [119] A. C. Doherty, P. A. Parrilo, and F. M. Spedalieri. *Complete family of separability criteria*. Phys. Rev. A **69**, 022308 (2004).
- [120] S. Wehner. *Tsirelson bounds for generalized Clauser-Horne-Shimony-Holt inequalities*. Phys. Rev. A **73**, 022110 (2006).
- [121] X. Wang and R. Duan. *Improved semidefinite programming upper bound on distillable entanglement*. Phys. Rev. A **94**, 050301(R) (2016).
- [122] M. Oszmaniec, L. Guerini, P. Wittek, and A. Acín. *Simulating Positive-Operator-Valued Measures with Projective Measurements*. Phys. Rev. Lett. **119**, 190501 (2017).
- [123] D. Bacon, A. M. Childs, and W. van Dam. *From optimal measurement to efficient quantum algorithms for the hidden subgroup problem over semidirect product groups*. In *46th Annual IEEE Symposium on Foundations of Computer Science (FOCS'05)*, pp. 469–478 (2005).

- [124] D. Bacon, A. M. Childs, and W. van Dam. *Optimal measurements for the dihedral hidden subgroup problem*. Chicago J. Theo. Comput. Sci. **1**(2) (2006).
- [125] C. H. Bennett. *Quantum cryptography using any two nonorthogonal states*. Phys. Rev. Lett. **68**, 3121 (1992).
- [126] J. M. Renes. *Spherical-code key-distribution protocols for qubits*. Phys. Rev. A **70**, 052314 (2004).
- [127] A. Acín, S. Pironio, T. Vértesi, and P. Wittek. *Optimal randomness certification from one entangled bit*. Phys. Rev. A **93**, 040102 (2016).
- [128] A. Bisio, G. Chiribella, G. M. D’Ariano, S. Facchini, and P. Perinotti. *Optimal Quantum Tomography*. IEEE J. Sel. Top. Quantum Electron. **15**, 1646–1660 (2009).
- [129] T. Decker, D. Janzing, and M. Rötteler. *Implementation of group-covariant positive operator valued measures by orthogonal measurements*. J. Math. Phys. **46**, 012104 (2005).
- [130] P. Mironowicz and M. Pawłowski. *Experimentally feasible semi-device-independent certification of four-outcome positive-operator-valued measurements*. Phys. Rev. A **100**, 030301 (2019).
- [131] A. Tavakoli. *Semi-Device-Independent Certification of Independent Quantum State and Measurement Devices*. Phys. Rev. Lett. **125**, 150503 (2020).
- [132] N. Miklin, J. J. Borkała, and M. Pawłowski. *Semi-device-independent self-testing of unsharp measurements*. Phys. Rev. Research **2**, 033014 (2020).
- [133] E. S. Gómez, S. Gómez, P. González, G. Cañas, J. F. Barra, A. Delgado, G. B. Xavier, A. Cabello, M. Kleinmann, T. Vértesi, and G. Lima. *Device-Independent Certification of a Nonprojective Qubit Measurement*. Phys. Rev. Lett. **117**, 260401 (2016).
- [134] A. T. M. Smania, T. Vértesi, N. Brunner, and M. Bourennane. *Self-testing non-projective quantum measurements in prepare-and-measure experiments*. Sci. Adv. **6**, eaaw6664 (2020).
- [135] M. Smania, P. Mironowicz, M. Nawareg, M. Pawłowski, A. Cabello, and M. Bourennane. *Experimental certification of an informationally complete quantum measurement in a device-independent protocol*. Optica **7**, 123–128 (2020).
- [136] D. Collins, N. Gisin, N. Linden, S. Massar, and S. Popescu. *Bell Inequalities for Arbitrarily High-Dimensional Systems*. Phys. Rev. Lett. **88**, 040404 (2002).



- [137] S. Zohren and R. D. Gill. *Maximal Violation of the Collins-Gisin-Linden-Massar-Popescu Inequality for Infinite Dimensional States*. Phys. Rev. Lett. **100**, 120406 (2008).
- [138] G. Sentis, B. Gendra, S. D. Bartlett, and A. C. Doherty. *Decomposition of any quantum measurement into extremals*. J. Phys. A **46**, 375302 (2013).
- [139] J. Hoffmann, C. Spee, O. Gühne, and C. Budroni. *Structure of temporal correlations of a qubit*. New J. Phys. **20**, 102001 (2018).
- [140] P. E. Frenkel and M. Weiner. *Classical Information Storage in an  $n$ -Level Quantum System*. Comm. Math. Phys. **340**, 563–574 (2015).
- [141] N. Brunner, M. Navascués, and T. Vértesi. *Dimension Witnesses and Quantum State Discrimination*. Phys. Rev. Lett. **110**, 150501 (2013).
- [142] M. D. Mazurek, M. F. Pusey, K. J. Resch, and R. W. Spekkens. *Experimentally Bounding Deviations From Quantum Theory in the Landscape of Generalized Probabilistic Theories*. PRX Quantum **2**, 020302 (2021).
- [143] M. J. Grabowecky, C. A. J. Pollack, A. R. Cameron, R. W. Spekkens, and K. J. Resch. *Experimentally bounding deviations from quantum theory for a photonic three-level system using theory-agnostic tomography*. Phys. Rev. A **105**, 032204 (2022).
- [144] C. A. Fuchs and C. M. Caves. *Ensemble-Dependent Bounds for Accessible Information in Quantum Mechanics*. Phys. Rev. Lett. **73**, 3047 (1994).
- [145] E. G. Cavalcanti. *Classical causal models for Bell and Kochen-Specker inequality violations require fine-tuning*. Phys. Rev. X **8**(2), 021018 (2018).
- [146] G. C. Ghirardi, A. Rimini, and T. Weber. *Unified dynamics for microscopic and macroscopic systems*. Phys. Rev. D **34**(2), 470–491 (1986).
- [147] P. Pearle. *Combining stochastic dynamical state-vector reduction with spontaneous localization*. Phys. Rev. A **39**(5), 2277–2289 (1989).
- [148] M. Arndt, O. Nairz, J. Vos-Andreae, C. Keller, G. van der Zouw, and A. Zeilinger. *Wave-particle duality of  $C^{60}$  molecules*. Nature **401**(6754), 680–682 (1999).
- [149] C. Monroe, D. M. Meekhof, B. E. King, and D. J. A. Wineland. *Schrödinger cat superposition state of an atom*. Science **272**, 1131–1136 (1996).
- [150] L. Davidovich, M. Brune, J. M. Raimond, and S. Haroche. *Mesoscopic quantum coherences in cavity QED: Preparation and decoherence monitoring schemes*. Phys. Rev. A **53**, 1295 (1996).

- [151] J. L. Anderson, G. C. Ghirardi, R. Grassi, P. Pearle, N. Gisin, D. Z. Albert, G. Feinberg, P. R. Holland, V. Ambegaokar, and K. J. Epstein. *Negotiating the tricky border between quantum and classical*. Phys. Today **46**(4), 13 (1993).
- [152] A. Shimony. *Controllable and uncontrollable non-locality*, chap. 10, pp. 130–139. Cambridge University Press (1993).
- [153] R. Healey. *Quantum Theory and the Limits of Objectivity*. Found. Phys. **48**, 1568–1589 (2018).
- [154] V. Baumann, F. D. Santo, and Ā. Brukner. *Comment on Healey’s “Quantum Theory and the Limits of Objectivity”*. Found. Phys. **49**, 741–749 (2019).
- [155] N. Brunner, D. Cavalcanti, S. Pironio, V. Scarani, and S. Wehner. *Bell nonlocality*. Rev. Mod. Phys. **86**(2), 419 (2014).
- [156] A. Peres. *Contextuality*, chap. 7, p. 187. Springer, Berlin (2002).
- [157] J. S. Bell. *Quantum mechanics for cosmologists*, chap. 15, pp. 117–138. Cambridge University Press, 2 ed. (2004).
- [158] P. M. Pearle. *Hidden-Variable Example Based upon Data Rejection*. Phys. Rev. D **2**, 1418 (1970).
- [159] S. Pironio, A. Acín, S. Massar, A. B. de la Giroday, D. N. Matsukevich, P. Maunz, S. Olmschenk, D. Hayes, L. Luo, T. A. Manning, and C. Monroe. *Random numbers certified by Bell’s theorem*. Nature **464**, 1021–1024 (2010).
- [160] W.-Z. Liu, M.-H. Li, S. Ragy, S.-R. Zhao, B. Bai, Y. Liu, P. J. Brown, J. Zhang, R. Colbeck, J. Fan, Q. Zhang, and J.-W. Pan. *Device-independent randomness expansion against quantum side information*. Nat. Phys. **17**, 448–451 (2021).
- [161] J. Barrett, L. Hardy, and A. Kent. *No Signaling and Quantum Key Distribution*. Phys. Rev. Lett. **95**, 010503 (2005).
- [162] A. Acín, N. Brunner, N. Gisin, S. Massar, S. Pironio, and V. Scarani. *Device-Independent Security of Quantum Cryptography against Collective Attacks*. Phys. Rev. Lett. **98**, 230501 (2007).
- [163] J.-A. Larsson. *Loopholes in Bell inequality tests of local realism*. J. Phys. A **47**, 424003 (2014).
- [164] N. Brunner and N. Gisin. *Partial list of bipartite bell inequalities with four binary settings*. Phys. Lett. A **372**, 3162 (2008).
- [165] E. Z. Cruzeiro and N. Gisin. *Complete list of tight bell inequalities for two parties with four binary settings*. Phys. Rev. A **99**, 022104 (2019).

- [166] W. Rosenfeld, D. Burchardt, R. Garthoff, K. Redeker, N. Ortegel, M. Rau, and H. Weinfurter. *Event-Ready Bell Test Using Entangled Atoms Simultaneously Closing Detection and Locality Loopholes*. Phys. Rev. Lett. **119**, 010402 (2017).
- [167] S. Massar. *Nonlocality, closing the detection loophole, and communication complexity*. Phys. Rev. A **65**, 032121 (2002).
- [168] T. Vértesi, S. Pironio, and N. Brunner. *Closing the Detection Loophole in Bell Experiments Using Qudits*. Phys. Rev. Lett. **104**, 060401 (2010).
- [169] I. Márton, E. Bene, and T. Vértesi. *Bounding the detection efficiency threshold in bell tests using multiple copies of the maximally entangled two-qubit state carried by a single pair of particles*. Phys. Rev. A **107**, 022205 (2023).
- [170] N. Miklin, A. Chaturvedi, M. Bourennane, M. Pawłowski, and A. Cabello. *Exponentially Decreasing Critical Detection Efficiency for Any Bell Inequality*. Phys. Rev. Lett. **129**, 230403 (2022).
- [171] J. Hofmann, M. Krug, N. Ortegel, L. Gérard, M. Weber, W. Rosenfeld, and H. Weinfurter. *Heralded Entanglement Between Widely Separated Atoms*. Science **337**, 72–75 (2012).
- [172] T. C. Ralph and A. P. Lund. *Nondeterministic Noiseless Linear Amplification of Quantum Systems*. AIP Conference Proceedings **1110**, 155–160 (2009).
- [173] N. Gisin, S. Pironio, and N. Sangouard. *Proposal for Implementing Device-Independent Quantum Key Distribution Based on a Heralded Qubit Amplifier*. Phys. Rev. Lett. **105**, 070501 (2010).
- [174] C. Branciard. *Detection loophole in Bell experiments: How postselection modifies the requirements to observe nonlocality*. Phys. Rev. A **83**, 032123 (2011).
- [175] A. Cabello and F. Sciarrino. *Loophole-Free Bell Test Based on Local Precertification of Photon's Presence*. Phys. Rev. X **2**, 021010 (2012).
- [176] H.-X. Meng, J. Zhou, Z.-P. Xu, H.-Y. Su, T. Gao, F.-L. Yan, and J.-L. Chen. *Hardy's paradox for multisetting high-dimensional systems*. Phys. Rev. A **98**, 062103 (2018).
- [177] M. Planat. *Quantum States Arising from the Pauli Groups, Symmetries and Paradoxes*. The XXIXth International Colloquium on Group-Theoretical Methods in Physics pp. 295–300 (2012).
- [178] D. Gross. *Hudson's theorem for finite-dimensional quantum systems*. J. Math. Phys. **47**, 122107 (2006).
- [179] N. Ito. *Hadamard graphs I*. Graphs Comb. **1**, 57–64 (1985).

- [180] N. Ito. *Hadamard graphs II*. Graphs Comb. **1**, 331–337 (1985).
- [181] G. Brassard, R. Cleve, and A. Tapp. *Cost of Exactly Simulating Quantum Entanglement with Classical Communication*. Phys. Rev. Lett. **83**, 1874 (1999).
- [182] V. Galliard, A. Tapp, and S. Wolf. *Deterministic quantum non-locality and graph colorings*. Theor. Comput. Sci. **486**, 20–26 (2013).
- [183] P. J. Cameron, A. Montanaro, M. W. Newman, S. Severini, and A. Winter. *On the quantum chromatic number of a graph*. Electron. J. Comb. **14**, 81 (2007).
- [184] G. Scarpa and S. Severini. *Kochen–Specker Sets and the Rank-1 Quantum Chromatic Number*. IEEE Trans. Inf. Theory **58**, 2524–2529 (2012).
- [185] M. W. Newman. *Independent Sets and Eigenspaces*. PhD thesis (2004).
- [186] L. Mančinska and D. E. Roberson. *Quantum homomorphisms*. J. Combin. Theory Ser. B **118**, 228–267 (2016).
- [187] P. Wočjan and C. Elphick. *Spectral lower bounds for the orthogonal and projective ranks of a graph*. Electron. J. Comb. **26**, 45 (2019).
- [188] P. Frankl. *Orthogonal vectors in the  $n$ -dimensional cube and codes with missing distances*. Combinatorica **6**, 279–285 (1986).
- [189] F. Ihringer and H. Tanaka. *The independence number of the orthogonality graph in dimension  $2^k$* . Combinatorica **39**, 1425–1428 (2019).
- [190] J. J. Sylvester. *Thoughts on inverse orthogonal matrices, simultaneous sign successions, and tessellated pavements in two or more colours, with applications to Newton’s rule, ornamental tile-work, and the theory of numbers*. Philos. Mag. **34**, 461–475 (1867).
- [191] R. E. A. C. Paley. *On Orthogonal Matrices*. J. Math. Phys. **4**, 311 (1933).
- [192] H. Kimura. *Classification of Hadamard matrices of order 28*. Discrete Math. **133**, 171–180 (1994).
- [193] B. D. McKay. *Topics in Computational Graph Theory*. Ph.D. thesis, University of Melbourne (1980).
- [194] N. H. Valencia, V. Srivastav, M. Pivluska, M. Huber, N. Friis, W. McCutcheon, and M. Malik. *High-Dimensional Pixel Entanglement: Efficient Generation and Certification*. Quantum **4**, 376 (2020).
- [195] R. J. Nowakowski and D. F. Rall. *Associative graph products and their independence, domination and coloring numbers*. Discuss. Math. Graph Theory **16**, 53–79 (1996).

- [196] D. Geller and S. Stahl. *The chromatic number and other functions of the lexicographic product*. J. Comb. Theory Ser. B **19**, 87–95 (1975).
- [197] D. E. Roberson. *Conic formulations of graph homomorphisms*. J. Algebr. Comb. **43**, 877–913 (2016).
- [198] G. F. Royle and C. E. Praeger. *Constructing the vertex-transitive graphs of order 24*. J. Symb. Comput. **8**, 309–326 (1989).
- [199] B. D. McKay and G. F. Royle. *The Transitive Graphs with at Most 26 Vertices*. Ars Combin. **30**, 161–176 (1990).
- [200] D. Holt and G. Royle. *A census of small transitive groups and vertex-transitive graphs*. J. Symb. Comput. **101**, 51–60 (2020).
- [201] P. Potočnik, P. Spiga, and G. Verret. *Cubic vertex-transitive graphs on up to 1280 vertices*. J. Symb. Comput. **50**, 465–477 (2013).
- [202] K. Coolsaet, S. D’hondt, and J. Goedgebeur. *House of Graphs 2.0: A database of interesting graphs and more*. Discrete Appl. Math. **325**, 97–107 (2023).
- [203] Z.-P. Xu, X.-D. Yu, and M. Kleinmann. *State-independent quantum contextuality with projectors of nonunit rank*. New J. Phys. **23**, 043025 (2021).
- [204] A. Cabello. *Converting Contextuality into Nonlocality*. Phys. Rev. Lett. **127**, 070401 (2021).
- [205] A. Cabello, M. Kleinmann, and C. Budroni. *Necessary and Sufficient Condition for Quantum State-Independent Contextuality*. Phys. Rev. Lett. **114**, 250402 (2015).
- [206] S. Yu and C. H. Oh. *State-Independent Proof of Kochen-Specker Theorem with 13 Rays*. Phys. Rev. Lett. **108**, 030402 (2012).
- [207] A. Cabello, M. Kleinmann, and J. R. Portillo. *Quantum state-independent contextuality requires 13 rays*. J. Phys. A **49**, 38LT01 (2016).
- [208] E. G. Gilbert. *An Iterative Procedure for Computing the Minimum of a Quadratic Form on a Convex Set*. J. SIAM Control **4**, 61 (1966).
- [209] A. Montina and S. Wolf. *Discrimination of Non-Local Correlations*. Entropy **21**, 1–22 (2019).
- [210] J. Shang and O. Gühne. *Convex Optimization over Classes of Multiparticle Entanglement*. Phys. Rev. Lett. **120**, 050506 (2018).
- [211] P. Pandya, O. Sakarya, and M. Wieśniak. *Hilbert-Schmidt distance and entanglement witnessing*. Phys. Rev. A **102**, 012409 (2020).

- [212] C. Branciard, E. G. Cavalcanti, S. P. Walborn, V. Scarani, and H. M. Wiseman. *One-sided device-independent quantum key distribution: Security, feasibility, and the connection with steering*. Phys. Rev. A **85**, 010301(R) (2012).
- [213] N. Gisin. *Bell's inequality holds for all non-product states*. Phys. Lett. A **154**, 201–202 (1991).
- [214] J. Barrett. *Nonsequential positive-operator-valued measurements on entangled mixed states do not always violate a Bell inequality*. Phys. Rev. A **65**, 042302 (2002).
- [215] S. Popescu. *Bell's Inequalities and Density Matrices: Revealing "Hidden" Nonlocality*. Phys. Rev. Lett. **74**, 2619 (1995).
- [216] N. Gisin. *Hidden quantum nonlocality revealed by local filters*. Phys. Lett. A **210**, 151–156 (1996).
- [217] A. Peres. *Collective tests for quantum nonlocality*. Phys. Rev. A **54**, 2685 (1996).
- [218] C. Palazuelos. *Superactivation of Quantum Nonlocality*. Phys. Rev. Lett. **109**, 190401 (2012).
- [219] F. Hirsch, M. T. Quintino, J. Bowles, and N. Brunner. *Genuine Hidden Quantum Nonlocality*. Phys. Rev. Lett. **111**, 160402 (2013).
- [220] D. J. Saunders, S. J. Jones, H. M. Wiseman, and G. J. Pryde. *Experimental EPR-steering using Bell-local states*. Nat. Phys. **6**, 845–849 (2010).
- [221] P. G. Kwiat, S. Barraza-Lopez, A. Stefanov, and N. Gisin. *Experimental entanglement distillation and "hidden" non-locality*. Nature **409**, 1014–1017 (2001).
- [222] T. Pramanik, Y.-W. Cho, S.-W. Han, S.-Y. Lee, Y.-S. Kim, and S. Moon. *Revealing hidden quantum steerability using local filtering operations*. Phys. Rev. A **99**, 030101(R) (2019).
- [223] Y. Wang, J. Li, X.-R. Wang, T.-J. Liu, and Q. Wang. *Experimental demonstration of hidden nonlocality with local filters*. Opt. Express **28**, 13638–13649 (2020).
- [224] J. Wang, V. B. Scholz, and R. Renner. *Confidence Polytopes in Quantum State Tomography*. Phys. Rev. Lett. **122**, 190401 (2019).
- [225] M. Guță, J. Kahn, R. Kueng, and J. A. Tropp. *Fast state tomography with optimal error bounds*. J. Phys. A **53**, 204001 (2020).
- [226] J. O. de Almeida, M. Kleinmann, and G. Sentís. *Comparison of confidence regions for quantum state tomography*. arXiv:2303.07136 (2023).
- [227] A. Peres. *Separability Criterion for Density Matrices*. Phys. Rev. Lett. **77**, 1413 (1996).

- [228] T.-A. Ohst, X.-D. Yu, O. Gühne, and H. C. Nguyen. *Certifying Quantum Separability with Adaptive Polytopes*. arXiv:2210.10054 (2023).
- [229] D. Cavalcanti, L. Guerini, R. Rabelo, and P. Skrzypczyk. *General Method for Constructing Local Hidden Variable Models for Entangled Quantum States*. Phys. Rev. Lett. **117**, 190401 (2016).
- [230] F. Hirsch, M. T. Quintino, T. Vértesi, M. F. Pusey, and N. Brunner. *Algorithmic Construction of Local Hidden Variable Models for Entangled Quantum States*. Phys. Rev. Lett. **117**, 190402 (2016).
- [231] H. C. Nguyen, H.-V. Nguyen, and O. Gühne. *Geometry of Einstein-Podolsky-Rosen Correlations*. Phys. Rev. Lett. **122**, 240401 (2019).
- [232] M. Fillettaz, F. Hirsch, S. Designolle, and N. Brunner. *Algorithmic construction of local models for entangled quantum states: Optimization for two-qubit states*. Phys. Rev. A **98**, 022115 (2018).
- [233] H. C. Nguyen, A. Milne, T. Vu, and S. Jevtic. *Quantum steering with positive operator valued measures*. J. Phys. A **51**, 355302 (2018).
- [234] L. Tendick, H. Kampermann, and D. Bruß. *Activation of Nonlocality in Bound Entanglement*. Phys. Rev. Lett. **124**, 050401 (2020).
- [235] R. T. Rockafellar. *Convex Analysis*. Princeton University Press (1970).
- [236] D. A. Evans, E. G. Cavalcanti, and H. M. Wiseman. *Loss-tolerant tests of Einstein-Podolsky-Rosen steering*. Phys. Rev. A **88**, 022106 (2013).
- [237] T. J. Baker, S. Wollmann, G. J. Pryde, and H. M. Wiseman. *Necessary condition for steerability of arbitrary two-qubit states with loss*. J. Opt. **20**, 034008 (2018).
- [238] M. T. Quintino, T. Vértesi, D. Cavalcanti, R. Augusiak, M. Demianowicz, A. Acín, and N. Brunner. *Inequivalence of entanglement, steering, and Bell nonlocality for general measurements*. Phys. Rev. A **92**, 032107 (2015).
- [239] N. Tischler, F. Ghafari, T. J. Baker, S. Slussarenko, R. B. Patel, M. M. Weston, S. Wollmann, L. K. Shalm, V. B. Verma, S. W. Nam, H. C. Nguyen, H. M. Wiseman, and G. J. Pryde. *Conclusive Experimental Demonstration of One-Way Einstein-Podolsky-Rosen Steering*. Phys. Rev. Lett. **121**, 100401 (2018).
- [240] M. J. Holland and K. Burnett. *Interferometric detection of optical phase shifts at the Heisenberg limit*. Phys. Rev. Lett. **71**, 1355 (1993).
- [241] X. Zhou, D. W. Leung, and I. L. Chuang. *Methodology for quantum logic gate construction*. Phys. Rev. A **62**, 052316 (2000).

- [242] S. Bravyi and A. Kitaev. *Universal quantum computation with ideal Clifford gates and noisy ancillas*. Phys. Rev. Lett. **71**, 022316 (2005).
- [243] R. Raussendorf and H. Briegel. *A One-Way Quantum Computer*. Phys. Rev. Lett. **86**, 5188 (2001).
- [244] R. Raussendorf, D. E. Browne, and H. J. Briegel. *Measurement-based quantum computation on cluster states*. Phys. Rev. A **68**, 022312 (2003).
- [245] P. Hyllus, W. Laskowski, R. Krischek, C. Schwemmer, W. Wieczorek, H. Weinfurter, L. Pezzé, and A. Smerzi. *Fisher information and multiparticle entanglement*. Phys. Rev. A **85**, 022321 (2012).
- [246] G. Tóth. *Multipartite entanglement and high-precision metrology*. Phys. Rev. A **85**, 022322 (2012).
- [247] D. Braun, G. Adesso, F. Benatti, R. Floreanini, U. Marzolino, M. W. Mitchell, and S. Pirandola. *Quantum-enhanced measurements without entanglement*. Rev. Mod. Phys. **90**, 035006 (2018).
- [248] V. Vedral and M. B. Plenio. *Entanglement measures and purification procedures*. Phys. Rev. A **57**, 1619 (1998).
- [249] G. Vidal and R. Tarrach. *Robustness of entanglement*. Phys. Rev. A **59**, 141 (1998).
- [250] O. Giraud, P. Braun, and D. Braun. *Quantifying quantumness and the quest for Queens of Quantum*. New J. Phys. **12**, 063005 (2010).
- [251] A. J. Scott. *Multipartite entanglement, quantum-error-correcting codes, and entangling power of quantum evolutions*. Phys. Rev. A **69**, 052330 (2004).
- [252] P. Facchi, G. Florio, G. Parisi, and S. Pascazio. *Maximally multipartite entangled states*. Phys. Rev. A **77**, 060304(R) (2008).
- [253] R. Reuvers. *An algorithm to explore entanglement in small systems*. Proc. R. Soc. A **474** (2018).
- [254] G. Gour and N. R. Wallach. *All maximally entangled four-qubit states*. J. Math. Phys. **51**, 112201 (2010).
- [255] F. Huber, C. Eltschka, J. Siewert, and O. Gühne. *Bounds on absolutely maximally entangled states from shadow inequalities, and the quantum MacWilliams identity*. J. Phys. A **51**, 175301 (2018).
- [256] D. U. Contreras and D. Goyeneche. *Reconstructing the whole from its parts*. arXiv:2209.14154 (2022).



- [257] P. Horodecki, L. Rudnicki, and K. Życzkowski. *Five Open Problems in Quantum Information Theory*. PRX Quantum **3**, 010101 (2022).
- [258] S. A. Rather, A. Burchardt, W. Bruzda, G. Rajchel-Mieldzióć, A. Lakshminarayan, and K. Życzkowski. *Thirty-six entangled officers of Euler*. Phys. Rev. Lett. **128**, 080507 (2022).
- [259] A. Shimony. *Degree of Entanglement*. Ann. N. Y. Acad. Sci. **755**, 675–679 (1995).
- [260] H. Barnum and N. Linden. *Monotones and invariants for multi-particle quantum states*. J. Phys. A **34**, 6787 (2001).
- [261] T.-C. Wei and P. M. Goldbart. *Geometric measure of entanglement and applications to bipartite and multipartite quantum states*. Phys. Rev. A **68**, 042307 (2003).
- [262] P. Badziąg, Č. Brukner, W. Laskowski, T. Paterek, and M. Żukowski. *Experimentally friendly geometrical criteria for entanglement*. Phys. Rev. Lett. **100**, 140403 (2008).
- [263] A. Sen(De) and U. Sen. *Channel capacities versus entanglement measures in multi-party quantum states*. Phys. Rev. A **81**, 012308 (2010).
- [264] L. Chen, H. Zhu, and T.-C. Wei. *Connections of geometric measure of entanglement of pure symmetric states to quantum state estimation*. Phys. Rev. A **83**, 012305 (2011).
- [265] R. Orús and T.-C. Wei. *Visualizing elusive phase transitions with geometric entanglement*. Phys. Rev. B **82**, 155120 (2010).
- [266] O. Buerschaper, A. García-Saez, R. Orús, and T.-C. Wei. *Topological minimally entangled states via geometric measure*. J. Stat. Mech. p. 11009 (2014).
- [267] Q.-Q. Shi, H.-L. Wang, S.-H. Li, S. Y. Cho, M. T. Batchelor, and H.-Q. Zhou. *Geometric entanglement and quantum phase transitions in two-dimensional quantum lattice models*. Phys. Rev. A **93**, 062341 (2016).
- [268] A. Deger and T.-C. Wei. *Geometric entanglement and quantum phase transition in generalized cluster-XY models*. Quantum Inf. Process. **18**, 326 (2019).
- [269] P. Hayden, D. W. Leung, and A. Winter. *Aspects of Generic Entanglement*. Commun. Math. Phys. **265**, 95–117 (2009).
- [270] S. Gharibian and J. Kempe. *Approximation Algorithms for QMA-complete Problems*. SIAM J. Comput. **41**, 1028–1050 (2012).
- [271] A. Montanaro. *Injective tensor norms and open problems in quantum information*. <https://people.maths.bris.ac.uk/csxam/presentations/injnormtalk.pdf> (2012).

- [272] M. van den Nest, W. Dür, A. Miyake, and H. J. Briegel. *Fundamentals of universality in one-way quantum computation*. *New J. Phys.* **9**, 1–51 (2007).
- [273] M. J. Bremner, C. Mora, and A. Winter. *Are Random Pure States Useful for Quantum Computation?* *Phys. Rev. Lett.* **102**, 190502 (2009).
- [274] D. Gross, S. T. Flammia, and J. Eisert. *Most Quantum States Are Too Entangled To Be Useful As Computational Resources*. *Phys. Rev. Lett.* **102**, 190501 (2009).
- [275] M. Aulbach, D. Markham, and M. Muraio. *The maximally entangled symmetric state in terms of the geometric measure*. *New J. Phys.* **12**, 073025 (2010).
- [276] J. Martin, O. Giraud, P. A. Braun, D. Braun, and T. Bastin. *Multiqubit symmetric states with high geometric entanglement*. *Phys. Rev. A* **81**, 062347 (2010).
- [277] M. Hajdušek. and M. Muraio. *Direct evaluation of pure graph state entanglement*. *New J. Phys.* **15**, 013039 (2013).
- [278] E. Chitambar, R. Duan, and Y. Shi. *Tripartite entanglement transformations and tensor rank*. *Phys. Rev. Lett.* **101**, 140502 (2008).
- [279] C. J. Hillar and L. H. Lim. *Most tensor problems are NP-hard*. *J. ACM* **60**, 1–39 (2013).
- [280] L. Qi. *Eigenvalues of a real supersymmetric tensor*. *J. Symb. Comput.* **40**, 1302–1324 (2005).
- [281] L. Chen, A. Xu, and H. Zhu. *Computation of the geometric measure of entanglement for pure multiqubit states*. *Phys. Rev. A* **82**, 032301 (2010).
- [282] G. Ni, L. Qi, and M. Bai. *Geometric Measure of Entanglement and U-Eigenvalues of Tensors*. *SIAM J. Matrix Anal. Appl.* **35**, 73–87 (2014).
- [283] S. Hu., L. Qi, and G. Zhang. *Computing the geometric measure of entanglement of multipartite pure states by means of non-negative tensors*. *Phys. Rev. A* **93**, 012304 (2016).
- [284] A. Czaplinski, T. Raasch, and J. Steinberg. *Real eigenstructure of regular simplex tensors*. *Adv. Appl. Math.* **148**, 102521 (2023).
- [285] R. F. Werner and A. S. Holevo. *Counterexample to an additivity conjecture for output purity of quantum channels*. *J. Math. Phys.* **43**, 4353 (2002).
- [286] G. Aubrun and S. J. Szarek. *Alice and Bob Meet Banach: The Interface of Asymptotic Geometric Analysis and Quantum Information Theory*, chap. 8. American Mathematical Society (2017).

- [287] T.-C. Wei and S. Severini. *Matrix permanent and quantum entanglement of permutation invariant states*. J. Math. Phys. **51**, 092203 (2010).
- [288] L. Qi, H. Chen, and Y. Chen. *Tensor Eigenvalues and Their Applications*. Springer, Singapore (2018).
- [289] V. de Silva and L.-H. Lim. *Tensor rank and the ill-posedness of the best low-rank approximation problem*. SIAM J. Matrix Anal. Appl. **30**, 1084–1127 (2008).
- [290] O. Gühne, M. Reimpell, and R. F. Werner. *Estimating entanglement measures in experiments*. Phys. Rev. Lett. **98**, 110502 (2007).
- [291] A. Streltsov, H. Kampermann, and D. Bruss. *Simple algorithm for computing the geometric measure of entanglement*. Phys. Rev. A **84**, 022323 (2011).
- [292] S. Gerke, W. Vogel, and J. Sperling. *Numerical construction of multipartite entanglement witnesses*. Phys. Rev. X **8**, 031047 (2018).
- [293] G. K. Pedersen. *Analysis Now*, chap. Hilbert Spaces, pp. 79–125. Graduate Texts in Mathematics. Springer New York (1989).
- [294] T. Tilma and E. C. G. Sudarshan. *Generalized Euler angle parametrization for  $SU(n)$* . J. Math. Phys. **35**, 10467 (2002).
- [295] C. Jarlskog. *A recursive parametrization of unitary matrices*. J. Math. Phys. **46**, 103508 (2005).
- [296] C. Spengler, M. Huber, and B. C. Hiesmayr. *A composite parameterization of unitary groups, density matrices and subspaces*. J. Phys. A **43**, 385306 (2010).
- [297] S. Tamaryan, T.-C. Wei, and D. Park. *Maximally entangled three-qubit states via geometric measure of entanglement*. Phys. Rev. A **80**, 052315 (2009).
- [298] P. Krammer, H. Kampermann, D. Bruß, R. A. Bertlmann, L. C. Kwek, and C. Macchiavello. *Multipartite Entanglement Detection via Structure Factors*. Phys. Rev. Lett. **103**, 100502 (2009).
- [299] A. Higuchi and A. Sudbery. *How entangled can two couples get?* Phys. Lett. A **273**, 213–217 (2000).
- [300] E. Knill, R. Laflamme, R. Martinez, and C. Negrevergne. *Benchmarking Quantum computers: The Five-Qubit Error Correcting Code*. Phys. Rev. Lett. **86**, 5811 (2001).
- [301] M. Hayashi, D. Markham, M. Muraio, M. Owari, and S. Virmani. *Entanglement of multiparty-stabilizer, symmetric, and antisymmetric states*. Phys. Rev. A **77**, 012104 (2008).

- [302] D. Goyeneche, Z. Raissi, S. D. Martino, and K. Życzkowski. *Entanglement and quantum combinatorial designs*. Phys. Rev. A **97**, 062326 (2018).
- [303] W. Helwig. *Absolutely Maximally Entangled Qudit Graph States*. arXiv:1306.2879 (2013).
- [304] A. Burchardt and Z. Raissi. *Stochastic local operations with classical communication of absolutely maximally entangled states*. Phys. Rev. A **102**, 022413 (2020).
- [305] E. M. Rains. *Nonbinary quantum codes*. IEEE Trans. Inf. Theory **45**, 1827 (1999).
- [306] P. Mehta, M. Bukov, C.-H. Wang, A. Day, C. Richardson, C. Fisher, and D. Schwab. *A high-bias, low-variance introduction to Machine Learning for physicists*. Phys. Rep. **810**, 1–124 (2019).
- [307] Y. Nesterov. *A method for unconstrained convex minimization problem with the rate of convergence  $o(1/k^2)$* . Soviet. Math. Docl. **269**, 543–547 (1983).
- [308] K. Życzkowski, M. Kuś, W. Słomczyński, and H.-J. Sommers. *Random unistochastic matrices*. J. Phys. A **36**, 3425 (2003).
- [309] W. Hoeffding. *A Class of Statistics with Asymptotically Normal Distribution*. In S. Kotz and N. L. Johnson, eds., *Breakthroughs in Statistics*, Springer Series in Statistics, pp. 308–334. Springer, New York (1992).
- [310] K. Fitter, C. Lancien, and I. Nechita. *Estimating the entanglement of random multipartite quantum states*. arXiv:2209.11754 (2022).
- [311] M. Demianowicz and R. Augusiak. *From unextendible product bases to genuinely entangled subspaces*. Phys. Rev. A **98**, 012313 (2018).
- [312] S. Bravyi, G. Smith, and J. A. Smolin. *Trading Classical and Quantum Computational Resources*. Phys. Rev. X **6**, 021043 (2016).
- [313] D. Gottesman. *Theory of fault-tolerant quantum computation*. Phys. Rev. A **57**, 127 (1998).
- [314] M. Erhard, M. Krenn, and A. Zeilinger. *Advances in high-dimensional quantum entanglement*. Nat. Rev. Phys. **2**, 365–381 (2020).
- [315] T. G. Kolda and B. W. Bader. *Tensor Decompositions and Applications*. SIAM Review **51**, 455–500 (2009).
- [316] J. Eisert and H.-J. Briegel. *The Schmidt Measure as a Tool for Quantifying Multi-Particle Entanglement*. Phys. Rev. A **64**, 022306 (2001).
- [317] A. Uhlmann. *Fidelity and concurrence of conjugated states*. Phys. Rev. A **62**, 032307 (2000).

- [318] J. Ja'Ja'. *Optimal Evaluation of Pairs of Bilinear Forms*. SIAM J. Comput. **8** (1979).
- [319] N. Yu, E. Chitambar, C. Guo, and R. Duan. *Tensor rank of the tripartite state  $|W^{\otimes 2}\rangle$* . Phys. Rev. A **81**, 014301 (2010).
- [320] G. Vidal. *Efficient Classical Simulation of Slightly Entangled Quantum Computations*. Phys. Rev. Lett. **91**, 147902 (2003).
- [321] A. Nico-Katz and S. Bose. *Entanglement-complexity geometric measure*. Phys. Rev. Research **5**, 013041 (2023).
- [322] F. Verstraete and I. Cirac. *Matrix product states represent ground states faithfully*. Phys. Rev. B **73**, 094423 (2006).
- [323] H. J. Kimble. *The quantum internet*. Nature **453**, 1023 (2008).
- [324] N. Sangouard, C. Simon, H. de Riedmatten, and N. Gisin. *Quantum repeaters based on atomic ensembles and linear optics*. Rev. Mod. Phys. **83**, 33 (2011).
- [325] T. Kraft, S. Designolle, C. Ritz, N. Brunner, O. Gühne, and M. Huber. *Quantum entanglement in the triangle network*. Phys. Rev. A **103**, L060401 (2021).
- [326] T. Kraft, C. Ritz, N. Brunner, M. Huber, and O. Gühne. *Characterizing Genuine Multilevel Entanglement*. Phys. Rev. Lett. **120**, 060502 (2018).
- [327] L. Qi, G. Zhang, and G. Ni. *How entangled can a multi-party system possibly be?* Phys. Lett. A **382**, 1465–1471 (2018).
- [328] S. Friedland and T. Kemp. *Most boson quantum states are almost maximally entangled*. Proc. Am. Math. Soc. **146**, 5035–5049 (2018).
- [329] M. D. Schatz, T. M. Low, R. A. van de Geijn, and T. G. Kolda. *Exploiting Symmetry in Tensors for High Performance: Multiplication with Symmetric Tensors*. SIAM J. Sci. Comput. **36**, 453–479 (2014).
- [330] C. Battaglino, G. Ballard, and T. G. Kolda. *A Practical Randomized CP Tensor Decomposition*. SIAM J. Matrix Anal. Appl. **39**, 876–901 (2018).
- [331] S. Gao, G. Mishne, and D. Scheinost. *Nonlinear manifold learning in functional magnetic resonance imaging uncovers a low-dimensional space of brain dynamics*. Hum. Brain Mapp. **42**, 4510–4524 (2021).
- [332] L. Pachter and B. Sturmfels. *Algebraic Statistics for Computational Biology*. Cambridge University Press (2010).
- [333] K. V. Gandikota, J. Geiping, Z. Löhner, A. Czapliński, and M. Möller. *A simple strategy to provable invariance via orbit mapping*. In *Proceedings of the 16th Asian Conference on Computer Vision, Macao, China, December 4–8, 2022* (2023).

- [334] L.-H. Lim. *Singular values and eigenvalues of tensors: a variational approach*. In *Proc. IEEE Int., Workshop on Comput. Advances in Multi-Sensor Adaptive Processing*, pp. 129–132. IEEE (2005).
- [335] L. D. Lathauwer, B. D. Moor, and J. Vandewalle. *On the best rank-1 and rank- $(R_1, R_2, \dots, R_N)$  approximation of higher-order tensors*. *SIAM J. Matrix Anal. Appl.* **21**, 1324–1342 (2000).
- [336] E. Robeva. *Orthogonal Decomposition of Symmetric Tensors*. *SIAM J. Matrix Anal. Appl.* **37**, 86–102 (2016).
- [337] L. Oeding, E. Robeva, and B. Sturmfels. *Decomposing tensors into frames*. *Adv. Appl. Math.* **73**, 125–153 (2016).
- [338] T. Muller, E. Robeva, and K. Usevich. *Robust Eigenvectors of Symmetric Tensors*. *SIAM J. Matrix Anal. Appl.* **43**, 1784–1805 (2022).
- [339] A. Anandkumar, R. Ge, D. Hsu, S. Kakade, and M. Telgarsky. *Tensor decompositions for learning latent variable models*. *J. Mach. Learn. Res.* **15**, 2773–2832 (2014).
- [340] P. G. Casazza and G. Kutyniok. *Finite Frames*. Birkhäuser, Boston (2013).
- [341] P. G. Casazza, D. Redmond, and J. C. Tremain. *Real equiangular frames*. The 42nd Annual Conference on Information Sciences and Systems pp. 715–720 (2008).
- [342] M. A. Sustik, J. A. Tropp, I. S. Dhillon, and R. W. Heath. *On the existence of equiangular tight frames*. *Linear Algebra Appl.* **426**, 619–635 (2007).
- [343] T. G. Kolda and J. R. Mayo. *Shifted Power Method for Computing Tensor Eigenpairs*. *SIAM J. Matrix Anal. Appl.* **32**, 1095–1124 (2011).
- [344] D. T. Smithey, M. Beck, M. G. Raymer, and A. Faridani. *Measurement of the Wigner distribution and the density matrix of a light mode using optical homodyne tomography: Application to squeezed states and the vacuum*. *Phys. Rev. Lett.* **70**, 1244 (1993).
- [345] D. F. V. James, P. G. Kwiat, W. J. Munro, and A. G. White. *Measurement of qubits*. *Phys. Rev. A* **64**, 052312 (2001).
- [346] 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, C. Becher, O. Gühne, W. Dür, and R. Blatt. *Scalable multiparticle entanglement of trapped ions*. *Nature* **438**, 643–646 (2005).
- [347] C. Schwemmer, L. Knips, D. Richart, H. Weinfurter, T. Moroder, M. Kleinmann, and O. Gühne. *Systematic Errors in Current Quantum State Tomography Tools*. *Phys. Rev. Lett.* **114**, 080403 (2015).

- [348] M. Paris and J. Řeháček. *Quantum State Estimation*. Lecture Notes in Physics. Springer, Berlin (2004).
- [349] S. Aaronson. *Shadow Tomography of Quantum States*. SIAM J. Comput. **49**, 368–394 (2020).
- [350] C. Kokail, C. Maier, R. van Bijnen, T. Brydges, M. K. Joshi, P. Jurcevic, C. A. Muschik, P. Silvi, R. Blatt, C. F. Roos, and P. Zoller. *Self-verifying variational quantum simulation of lattice models*. Nature **569**, 355–360 (2019).
- [351] M. Cerezo, A. Arrasmith, R. Babbush, S. C. Benjamin, S. Endo, K. Fujii, J. R. McClean, K. Mitarai, X. Yuan, L. Cincio, and P. J. Coles. *Variational quantum algorithms*. Nat. Rev. Phys. **3**, 625–644 (2021).
- [352] A. Peruzzo, J. McClean, P. Shadbolt, M.-H. Yung, X.-Q. Zhou, P. J. Love, A. Aspuru-Guzik, and J. L. O’Brien. *A variational eigenvalue solver on a photonic quantum processor*. Nat. Commun. **5**, 4213 (2014).
- [353] H.-Y. Huang, R. Kueng, and J. Preskill. *Predicting many properties of a quantum system from very few measurements*. Nat. Phys. **16**, 1050–1057 (2020).
- [354] C. M. Bishop. *Pattern Recognition and Machine Learning*. Information Science and Statistics. Springer, New York (2006).
- [355] C. Hadfield. *Adaptive Pauli Shadows for Energy Estimation*. arXiv:2105.12207 (2021).
- [356] C. Hadfield, S. Bravyi, R. Raymond, and A. Mezzacapo. *Measurements of Quantum Hamiltonians with Locally-Biased Classical Shadows*. Commun. Math. Phys. **391**, 951–967 (2022).
- [357] A. Elben, R. Kueng, H.-Y. Huang, R. van Bijnen, C. Kokail, M. Dalmonte, P. Calabrese, B. Kraus, J. Preskill, P. Zoller, and B. Vermersch. *Mixed-State Entanglement from Local Randomized Measurements*. Phys. Rev. Lett. **125**, 200501 (2020).
- [358] A. Neven, J. Carrasco, V. Vitale, C. Kokail, A. Elben, M. Dalmonte, P. Calabrese, P. Zoller, B. Vermersch, R. Kueng, and B. Kraus. *Symmetry-resolved entanglement detection using partial transpose moments*. npj Quantum Inf. **7**, 152 (2021).
- [359] A. Rath, C. Branciard, A. Minguzzi, and B. Vermersch. *Quantum Fisher Information from Randomized Measurements*. Phys. Rev. Lett. **127**, 260501 (2021).
- [360] R. J. Garcia, Y. Zhou, and A. Jaffe. *Quantum scrambling with classical shadows*. Phys. Rev. Research **3**, 033155 (2021).

- [361] L. K. Joshi, A. Elben, A. Vikram, B. Vermersch, V. Galitski, and P. Zoller. *Probing Many-Body Quantum Chaos with Quantum Simulators*. Phys. Rev. X **12**, 011018 (2022).
- [362] H.-Y. Huang, R. Kueng, and J. Preskill. *Efficient Estimation of Pauli Observables by Derandomization*. Phys. Rev. Lett. **127**, 030503 (2021).
- [363] H.-Y. Hu, S. Choi, and Y.-Z. You. *Classical shadow tomography with locally scrambled quantum dynamics*. Phys. Rev. Research **5**, 023027 (2023).
- [364] H.-Y. Hu and Y.-Z. You. *Hamiltonian-driven shadow tomography of quantum states*. Phys. Rev. Research **4**, 013054 (2022).
- [365] T. Zhang, J. Sun, X.-X. Fang, X.-M. Zhang, X. Yuan, and H. Lu. *Experimental Quantum State Measurement with Classical Shadows*. Phys. Rev. Lett. **127**, 200501 (2021).
- [366] S. Chen, W. Yu, P. Zeng, and S. T. Flammia. *Robust Shadow Estimation*. PRX Quantum **2**, 030348 (2021).
- [367] R. Levy, D. Luo, and B. K. Clark. *Classical Shadows for Quantum Process Tomography on Near-term Quantum Computers*. arXiv:2110.02965 (2021).
- [368] J. Helsen, M. Ioannou, J. Kitzinger, E. Onorati, A. H. Werner, J. Eisert, and I. Roth. *Estimating gate-set properties from random sequences*. arXiv:2110.13178 (2022).
- [369] A. Acharya, S. Saha, and A. M. Sengupta. *Shadow tomography based on informationally complete positive operator-valued measure*. Phys. Rev. A **104**, 052418 (2021).
- [370] G. M. D'Ariano, P. L. Presti, and P. Perinotti. *Classical randomness in quantum measurements*. J. Phys. A **38**, 5979–5991 (2005).
- [371] K. Bu, D. E. Koh, R. J. Garcia, and A. Jaffe. *Classical shadows with Pauli-invariant unitary ensembles*. arXiv:2202.03272 (2022).
- [372] H. C. Nguyen, S. Designolle, M. Barakat, and O. Gühne. *Symmetries between measurements in quantum mechanics*. arXiv:2003.12553 (2020).
- [373] H. Zhu and B.-G. Englert. *Quantum state tomography with fully symmetric measurements and product measurements*. Phys. Rev. A **84**, 022327 (2011).
- [374] Y. I. Bogdanov, G. Brida, I. D. Bukeev, M. Genovese, K. S. Kravtsov, S. P. Kulik, E. V. Moreva, A. A. Soloviev, and A. P. Shurupov. *Statistical estimation of the quality of quantum-tomography protocols*. Phys. Rev. A **84**, 042108 (2011).
- [375] W. Fulton and J. Harris. *Representation Theory*. Graduate Texts in Mathematics. Springer, New York (2004).



- [376] Y. Chen, M. Farahzad, S. Yoo, and T.-C. Wei. *Detector tomography on IBM quantum computers and mitigation of an imperfect measurement*. Phys. Rev. A **100**, 052315 (2019).
- [377] J. Heinsoo, C. K. Andersen, A. Remm, S. Krinner, T. Walter, Y. Salathé, S. Gasparinetti, J.-C. Besse, A. Potočnik, A. Wallraff, and C. Eichler. *Rapid High-fidelity Multiplexed Readout of Superconducting Qubits*. Phys. Rev. Applied **10**, 034040 (2018).
- [378] S. Bravyi, S. Sheldon, A. Kandala, D. C. McKay, and J. M. Gambetta. *Mitigating measurement errors in multiqubit experiments*. Phys. Rev. A **103**, 042605 (2021).
- [379] R. Hicks, B. Kobrin, C. W. Bauer, and B. Nachman. *Active readout-error mitigation*. Phys. Rev. A **105**, 012419 (2022).
- [380] B. Nachman, M. Urbanek, W. A. de Jong, and C. W. Bauer. *Unfolding quantum computer readout noise*. npj Quantum Inf. **6**, 84 (2020).
- [381] K. Temme, S. Bravyi, and J. M. Gambetta. *Error Mitigation for Short-Depth Quantum Circuits*. Phys. Rev. Lett. **119**, 180509 (2017).
- [382] H.-S. Zhong et al. *Phase-Programmable Gaussian Boson Sampling Using Stimulated Squeezed Light*. Phys. Rev. Lett. **127**, 180502 (2021).
- [383] Y. Wu et al. *Strong Quantum Computational Advantage Using a Superconducting Quantum Processor*. Phys. Rev. Lett. **127**, 180501 (2021).
- [384] S. Ebadi et al. *Quantum phases of matter on a 256-atom programmable quantum simulator*. Nature **595**, 227–232 (2021).
- [385] R. Ott, T. V. Zache, F. Jendrzejewski, and J. Berges. *Scalable Cold-Atom Quantum Simulator for Two-Dimensional QED*. Phys. Rev. Lett. **127**, 130504 (2021).
- [386] Y. Cao, J. Romero, J. P. Olson, M. Degroote, P. D. Johnson, M. Kieferová, I. D. Kivlichan, T. Menke, B. Peropadre, N. P. D. Sawaya, S. Sim, L. Veis, and A. Aspuru-Guzik. *Quantum Chemistry in the Age of Quantum Computing*. Chem. Rev. **119**, 10856–10915 (2019).
- [387] B. Bauer, S. Bravyi, M. Motta, and G. K.-L. Chan. *Quantum Algorithms for Quantum Chemistry and Quantum Materials Science*. Chem. Rev. **120**, 12685–12717 (2020).
- [388] M. Motta and J. E. Rice. *Emerging quantum computing algorithms for quantum chemistry*. Wiley Interdiscip. Rev. Comput. Mol. Sci. **12**, 1580 (2018).
- [389] T. E. O'Brien et al. *Purification-based quantum error mitigation of pair-correlated electron simulations*. arXiv:2210.10799 (2022).

- [390] K. Bharti, A. Cervera-Lierta, T. H. Kyaw, T. Haug, S. Alperin-Lea, A. Anand, M. Degroote, H. Heimonen, J. S. Kottmann, T. Menke, W.-K. Mok, S. Sim, L.-C. Kwek, and A. Aspuru-Guzik. *Noisy intermediate-scale quantum algorithms*. *Rev. Mod. Phys.* **94**, 015004 (2022).
- [391] S. Endo, Z. Cai, S. C. Benjamin, and X. Yuan. *Hybrid Quantum-Classical Algorithms and Quantum Error Mitigation*. *J. Phys. Soc. Jpn.* **90**, 032001 (2021).
- [392] B. van Straaten and B. Koczor. *Measurement cost of metric-aware variational quantum algorithms*. *PRX Quantum* **2**, 030324 (2021).
- [393] G. Boyd and B. Koczor. *Training Variational Quantum Circuits with CoVaR: Covariance Root Finding with Classical Shadows*. *Phys. Rev. X* **12**, 041022 (2022).
- [394] H. H. S. Chan, R. Meister, M. L. Goh, and B. Koczor. *Algorithmic Shadow Spectroscopy*. *arXiv:2212.11036* (2023).
- [395] A. Seif, Z.-P. Cian, S. Zhou, S. Chen, and L. Jiang. *Shadow Distillation: Quantum Error Mitigation with Classical Shadows for Near-Term Quantum Processors*. *PRX Quantum* **4**, 010303 (2023).
- [396] B. Koczor. *Exponential Error Suppression for Near-Term Quantum Devices*. *Phys. Rev. X* **11**, 031057 (2021).
- [397] D. E. Koh and S. Grewal. *Classical Shadows With Noise*. *Quantum* **6**, 776 (2022).
- [398] H. C. Nguyen, J. L. Bönsel, J. Steinberg, and O. Gühne. *Optimizing Shadow Tomography with Generalized Measurements*. *Phys. Rev. Lett.* **129**, 220502 (2022).
- [399] A. Glos, A. Nykänen, E.-M. Borrelli, S. Maniscalco, M. A. C. Rossi, Z. Zimborás, and G. García-Pérez. *Adaptive POVM implementations and measurement error mitigation strategies for near-term quantum devices*. *arXiv:2208.07817* (2022).
- [400] K. Wan, W. J. Huggins, J. Lee, and R. Babbush. *Matchgate Shadows for Fermionic Quantum Simulation*. *arXiv:2207.13723* (2022).
- [401] C. Bertoni, J. Haferkamp, M. Hinsche, M. Ioannou, J. Eisert, and H. Pashayan. *Shallow shadows: Expectation estimation using low-depth random Clifford circuits*. *arXiv:2209.12924* (2023).
- [402] S. Endo, S. C. Benjamin, and Y. Li. *Practical Quantum Error Mitigation for Near-Future Applications*. *Phys. Rev. X* **8**, 031027 (2018).
- [403] A. Strikis, D. Qin, Y. Chen, S. C. Benjamin, and Y. Li. *Learning-Based Quantum Error Mitigation*. *PRX Quantum* **2**, 040330 (2021).

- [404] A. Lowe, M. H. Gordon, P. Czarnik, A. Arrasmith, P. J. Coles, and L. Cincio. *Unified approach to data-driven quantum error mitigation*. Phys. Rev. Research **3**, 033098 (2021).
- [405] E. van den Berg, Z. K. Mineev, A. Kandala, and K. Temme. *Probabilistic error cancellation with sparse Pauli–Lindblad models on noisy quantum processors*. Nat. Phys. (2023).
- [406] A. Montanaro and S. Stanisic. *Error mitigation by training with fermionic linear optics*. arXiv:2102.02120 (2021).
- [407] M. R. Jerrum, L. G. Valiant, and V. V. Vazirani. *Random generation of combinatorial structures from a uniform distribution*. Theoret. Comput. Sci. **43**, 169–188 (1986).
- [408] M. Lerasle. *Selected topics on robust statistical learning theory*. arXiv:1908.10761 (2019).
- [409] M. C. Tran, K. Sharma, and K. Temme. *Locality and Error Mitigation of Quantum Circuits*. arXiv:2303.06496 (2023).
- [410] M. C. Tran, C.-F. Chen, A. Ehrenberg, A. Y. Guo, A. Deshpande, Y. Hong, Z.-X. Gong, A. V. Gorshkov, and A. Lucas. *Hierarchy of Linear Light Cones with Long-Range Interactions*. Phys. Rev. X **10**, 031009 (2020).
- [411] X. Bonet-Monroig, R. Sagastizabal, M. Singh, and T. E. O’Brien. *Low-cost error mitigation by symmetry verification*. Phys. Rev. A **98**, 062339 (2018).
- [412] S. McArdle, X. Yuan, and S. Benjamin. *Error-Mitigated Digital Quantum Simulation*. Phys. Rev. Lett. **122**, 180501 (2019).
- [413] L. Botelho, A. Glos, A. Kundu, J. A. Miszczak, Ö. Salehi, and Z. Zimborás. *Error mitigation for variational quantum algorithms through mid-circuit measurements*. Phys. Rev. A **105**, 022441 (2022).
- [414] Z. Jiang, J. R. McClean, R. Babbush, and H. Neven. *Majorana Loop Stabilizer Codes for Error Mitigation in Fermionic Quantum Simulations*. Phys. Rev. Applied **12**, 064041 (2019).
- [415] J. R. McClean, Z. Jiang, N. C. Rubin, R. Babbush, and H. Neven. *Decoding quantum errors with subspace expansions*. Nat. Commun. **11**, 636 (2020).
- [416] B. Koczor and S. C. Benjamin. *Quantum analytic descent*. Phys. Rev. Research **4**, 023017 (2022).
- [417] B. Koczor and S. C. Benjamin. *Quantum natural gradient generalized to noisy and nonunitary circuits*. Phys. Rev. A **106**, 062416 (2022).

- [418] R. Nandkishore and D. A. Huse. *Many-Body Localization and Thermalization in Quantum Statistical Mechanics*. *Annu. Rev. Condens. Matter Phys.* **6**, 15–38 (2015).
- [419] D. J. Luitz, N. Laflorencie, and F. Alet. *Many-body localization edge in the random-field Heisenberg chain*. *Phys. Rev. B* **91**, 081103(R) (2015).
- [420] A. M. Childs, D. Maslov, Y. Nam, N. J. Ross, and Y. Su. *Toward the first quantum simulation with quantum speedup*. *Proc. Natl. Acad. Sci. U.S.A.* **115**, 9456–9461 (2018).
- [421] N. Friis, O. Marty, C. Maier, C. Hempel, M. Holzäpfel, P. Jurcevic, M. B. Plenio, M. Huber, C. Roos, R. Blatt, and B. Lanyon. *Observation of Entangled States of a Fully Controlled 20-Qubit System*. *Phys. Rev. X* **8**, 021012 (2018).
- [422] Y. Dai, Y. Dong, Z. Xu, W. You, C. Zhang, and O. Gühne. *Experimentally Accessible Lower Bounds for Genuine Multipartite Entanglement and Coherence Measures*. *Phys. Rev. Applied* **13**, 054022 (2020).
- [423] H.-Y. Huang, R. Kueng, G. Torlai, V. V. Albert, and J. Preskill. *Provably efficient machine learning for quantum many-body problems*. *Science* **377**, 1–10 (2022).
- [424] T. E. Baker and D. Poulin. *Density functionals and Kohn-Sham potentials with minimal wavefunction preparations on a quantum computer*. *Phys. Rev. Research* **2**, 043238 (2020).
- [425] S. Ma, C. Dasgupta, and C. Hu. *Random Antiferromagnetic Chain*. *Phys. Rev. Lett.* **43**, 1434 (1979).
- [426] C. Dasgupta and S. Ma. *Low-temperature properties of the random Heisenberg antiferromagnetic chain*. *Phys. Rev. B* **22**, 1305 (1980).
- [427] B. Koczor, S. Endo, T. Jones, Y. Matsuzaki, and S. C. Benjamin. *Variational-state quantum metrology*. *New J. Phys.* **22**, 083038 (2020).